From Knowledge Access to Economic Value

Essays on Digitalization, Innovation, and Patent Valuation

Ann-Christin Kreyer

Dissertation

Munich 2025

Max Planck Institute for Innovation and Competition

From Knowledge Access to Economic Value

Essays on Digitalization, Innovation, and Patent Valuation

Inaugural-Dissertation
zur Erlangung des Grades

Doctor oeconomiae publicae
(Dr. oec. publ.)

an der
Ludwig–Maximilians–Universität München

2025

vorgelegt von

Ann-Christin Kreyer

Referent: Prof. Dietmar Harhoff, Ph.D.
Korreferent: Prof. Fabian Waldinger, Ph.D.
Promotionsabschlussberatung: 16. Juli 2025

Datum der mündlichen Prüfung:	02. Juli 2025
Namen der Berichterstatter:	Prof. Dietmar Harhoff, Ph.D.
	Prof. Fabian Waldinger, Ph.D.
	Prof. Dr. Hanna Hottenrott

Acknowledgements

This dissertation would not have been possible without the support, guidance, and encouragement of many individuals and institutions, to whom I am deeply grateful.

First and foremost, I would like to thank my advisor, Dietmar Harhoff, for his excellent guidance, constructive feedback, continuous support, and the academic freedom he provided me throughout my doctoral journey.

I am also thankful to Fabian Waldinger for his insightful comments and encouragement at various stages of my research. Further, I am grateful to Hanna Hottenrott for completing my committee and for her valuable suggestions and support throughout my doctoral journey.

I am especially indebted to my co-author, Lucy Xiaolu Wang, who has mentored me and taught me so much over the past years. Her deep knowledge and enthusiasm for digitalization and innovation economics have been a constant source of motivation. Her guidance, her support, and her friendship were invaluable to the completion of this dissertation. I would also like to thank Herbert Henzler for his mentorship and the stimulating discussions that pushed me to think further and work harder.

I am grateful to the Max Planck Institute for Innovation and Competition for providing an outstanding research environment. In particular, I would like to thank the administrative and IT teams, as well as my research assistants, who were an invaluable support throughout my doctoral journey. I am also deeply grateful to Jeffrey Furman and the Questrom School of Business at the Boston University for hosting me while I worked on parts of this dissertation.

I am particularly thankful to my colleagues and friends at the Max Planck Institute, TUM, and LMU for their inspiring discussions, insightful feedback, and their support. My special thanks go to Erik Buunk, Marina Chugunova, Niko Dellner, Sebastian Erhardt, Jonathan Federle, Svenja Frieß, Mainak Ghosh, Klaus Keller, Anna-Sophie

Liebender-Luc, Felix Pöge, Myriam Rion, Albert Roger, Michael Rose, Cristina Rujan, Julian Schwierzy, Gisela Stingl, and Kathrin Wernsdorf.

Above all, I am deeply grateful to my family. My heartfelt thanks go to my mum, my sister, and both my grandmother and my honorary grandmother for their steadfast support and patience – this dissertation would not have been possible without their encouragement. I am especially thankful to Ernst, whose unwavering support and adventurous spirit, much like mine, brought balance to my doctoral journey. Thank you for being both my anchor and my escape, for exploring the world with me, and for always making me smile. I am also grateful to my friends for their continued support and for the many cherished memories we have created together along the way.

Finally, I am deeply grateful to my dad for inspiring my curiosity and encouraging me to explore the world with an open mind. I dedicate this thesis to him.

Contents

Preface

Since the early 20th century, scientists have studied innovation and knowledge creation (Mansfield, 1968; Soete & Freeman, 1997). As technological progress accelerates – most recently in the wake of the Fourth Industrial Revolution – this field of research continues to adapt, reflecting the dynamic interplay between innovation, society, and technology (Smith & Marx, 1994). With each innovation, the collective knowledge base expands, which implies that advancing the knowledge frontier is becoming increasingly complex and necessitates interdisciplinary approaches (Burt, 2004; Wuchty, Jones, & Uzzi, 2007; Jones, 2009; Bloom et al., 2020). The generation of new knowledge requires the integration of both codified (explicit) and tacit knowledge (Nelson & Winter, 1985; Nonaka, Takeuchi, & Umemoto, 1996). In particular, the transfer of tacit knowledge, which is inherently difficult to codify, is contingent on interpersonal interactions (Audretsch, 1998; von Krogh, 1998; Breschi, 2001; Gertler, 2003; Popadiuk & Choo, 2006).

However, this process presupposes that individuals have access to knowledge (Bell et al., 2019; Berkes & Nencka, 2020; Furman, Nagler, & Watzinger, 2021) and that there are environments where people can meet and exchange ideas (Breschi, 2001; Breschi, Lissoni, & Montobbio, 2005; Peschl & Fundneider, 2012). The pivotal role of institutions in providing access to knowledge and thereby fostering innovation has been well documented in the literature (Furman & Stern, 2011; Williams, 2013; Murray et al., 2016; L. X. Wang, 2022). The importance of such institutions is underscored by the innovations' public good characteristics. These characteristics imply that innovations' benefits often extend beyond the firm that creates it. As a result, markets tend to provide insufficient incentives for innovation (Nelson, 1959; Arrow, 1962). At the same time, firms are often reluctant to invest in long-term research due to its high uncertainty and delayed returns (Budish, Roin, & Williams, 2015).

Digitalization and technological progress have impacted the knowledge creation process on multiple levels (Bailey et al., 2022) and contributed to a democratization of innovation (von Hippel, 2005). First, the advent of digital technologies has significantly reduced barriers to knowledge access through the widespread availability of online libraries, massive open online courses (MOOCs), and online knowledge exchange platforms (Davenport & Prusak, 1998; Boudreau, 2010; Hwang, Singh, & Argote, 2015; Faraj et al., 2016). Second, it has transformed the way in which knowledge is created by facilitating online collaboration and promoting open and user innovation settings (von Hippel, 2005; Lanzolla et al., 2020; Bailey et al., 2022; Roche, Oettl, & Catalini, 2024).

Furthermore, technological advancements have altered the tools available for generating new knowledge. In the health sector, for instance, the advent of information technologies and the capacity to analyze large datasets has enabled the development of new treatments and improvements in the quality of patient care (Kreyer & Wang, 2022). In the field of neurosciences, technologies such as artificial intelligence and online platforms provide essential support for discovering new drugs and treatments (Falotico et al., 2017; Lou & Wu, 2021). Neurological disorders remain one of the leading causes of death worldwide and available treatments are often inadequate or too costly (Carroll, 2019; Cantarero-Prieto et al., 2020). The convergence of the physical and the digital is another important development, with ideas being designed digitally and then translated into physical products that allow for personal fabrication (Mikhak et al., 2002), thus further democratizing innovation (von Hippel & Krogh, 2003; Ludwig et al., 2014). Taken together, these changes have significant yet understudied implications for knowledge creation and innovation. As technology continues to evolve, understanding its impact on the innovation process remains critical to supporting the knowledge creation process and to informing innovation policymakers.

This dissertation contributes to the understanding of how access to knowledge and interdisciplinary knowledge exchange, in both digital and physical spaces, affect the creation of knowledge and, ultimately, innovation. Beginning with the impact of digital access in the context of neuroscientific research, Chapters 1 and 2 investigate the effects of the digital research infrastructure created by the European Union's Human Brain Project (HBP) on research productivity as well as on online discussion and problem-solving patterns among researchers. This analysis offers valuable insights into the impact of a top-down initiative that builds an online research platform for neurosciences on research. By offering access to comprehensive neuroscientific datasets and computational tools within a supercomputer-powered digital infrastructure, such initiatives can shape research productivity and foster interdisciplinary collaboration. Furthermore, the analysis of the HBP online discussion forum provides insights into how researchers discuss and solve problems online. In some cases, the acquisition of knowledge, especially tacit knowledge, and the mastery of new digital technologies depend on physical access to tools and machines, involving hands-on learning. (Gertler, 2003).

Moving from online to physical spaces, the third chapter of this dissertation focuses on the physical dimension of knowledge access and exchange by analyzing the impact of FabLabs on innovation in Germany. Community workshops, such as FabLabs, have the potential to democratize innovation by offering open access to digital fabrication tools, facilitating the acquisition of skills, and fostering the exchange of knowledge among diverse users, including tinkerers, inventors and entrepreneurs. Digital fabrication, the domain in which FabLabs specialize in, entails the process of designing artifacts digitally, followed by the fabrication of tangible objects using digital production machines (i.e., 3D printers). Digital fabrication tools are thus a prime example of how advances in technologies lead to a merge between physical and digital worlds, enabling rapid prototyping and thus for users to become innovators (Baldwin & von Hippel, 2011; Fleischmann, Hielscher, & Merritt, 2016; Franke & Lüthje, 2020).

While the democratization of knowledge access through digital platforms and community workshops plays a significant role in driving innovation, another key mechanism that incentivizes innovative activity is the patent system. By offering legal protection of the invention, patents provide economic incentives for innovation (Moser, 2005; Hall & Helmers, 2024). Patents represent one form of codified knowledge and are widely employed to protect inventions across technical fields in both the private and public sectors, and thus serve as an important measure of innovative activity (Griliches, 1990). As ubiquitous as patents are, their heterogeneity and complexity make them challenging to assess, as their value is contingent on both the technical contribution and the complementary assets owned by the patent holder (Teece, 1986). Assessing the value of patents remains a subject of ongoing debate among economists, with a range of approaches including surveys and analyses of licensing revenue (e.g., Harhoff et al., 1999; Bray & Lee, 2000; Giuri et al., 2007; Ziedonis, 2007).

Recent studies have increasingly relied on stock market reactions to patent grants issued to public firms as a straightforward, comprehensive, and cost-effective valuation method (Hall, Jaffe, & Trajtenberg, 2005; Hsu et al., 2021). The measure proposed by (henceforth KPSS Kogan et al., 2017) is one of the most widely used and cited approaches for assigning an economic value to patents based on stock market reactions to United States Patent and Trademark Office (USPTO) grant announcements. The widespread use of these KPSS-based economic patent values underscores the demand among innovation researchers to assign a value to patents. Building on the existing patent valuation literature, Chapter 4 addresses two major shortcomings of the KPSS approach and offers an alternative method for measuring patent value based on stock market reactions.

All studies of this dissertation are empirical and employ different and novel datasets. The first three chapters focus on the early stages of the innovation process by analyzing how the digitalization and democratization of innovation improve access to knowledge and drive knowledge creation and, ultimately, innovation. The final chapter examines the codified output of the innovation process, namely patents, and their economic valuation. Taken together, the analyses span both the creation of knowledge and innovation, as well as the valuation of its codified output.

Chapter 1, co-authored with Lucy Xiaolu Wang, provides an analysis of the impact of the European Union's Human Brain Project (HBP) on research productivity in the brain sciences. The HBP, which was launched in 2013 as a 10-year flagship project with a budget of €1 billion, was part of the European Commission's Future and Emerging Technologies program. Its aim was to advance brain sciences by building a comprehensive online research infrastructure, enabled by high-performance supercomputers, and by fostering interdisciplinary research. Motivated by the need to address neurological disorders and the paucity of treatments for these diseases, this chapter examines the hypothesis that participation in the HBP accelerated researchers' productivity, improved publication quality, and fostered interdisciplinary collaboration. We also analyze whether these effects vary by researcher's career stage, gender, and scientific subfields.

Participation can increase research productivity by enabling access to knowledge, both in terms of quantity and diversity, and by providing new opportunities for knowledge exchange. Conversely, the acquisition of novel skills and the exploration of new research domains can be arduous and time-consuming, potentially leading to a decline in productivity. The rate of productivity may remain unchanged if researchers have limited time resources and choose to switch research topics rather than take on additional ones. This study draws upon multiple sources to construct an original, individual-level dataset of 639 active HBP participants, encompassing the duration of their HBP participation, their research output over time including citations, gender, seniority level, and academic affiliation. To identify the effect of the HBP on researchers' productivity, we employ a staggered Difference-in-Difference model on the author-year level. To estimate the interdisciplinarity of the research output, we have employed a large language model to classify the publication topics.

The empirical findings reveal that active involvement in the HBP significantly increases individual productivity, measured by the number of publications and follow-on citations. Of particular note is the finding that junior scholars and female researchers show a significant increase in productivity. Furthermore, participation in the HBP has been shown to increase the probability of publishing in top-tier neuroscience journals and to foster interdisciplinary collaboration.

These findings underscore the value of digital research infrastructures in fostering high-quality, innovative scientific work. We thus expand the literature on institutions that support long-term R&D (Azoulay, Graff Zivin, & Manso, 2011; Furman & Stern, 2011) by analyzing the impact of digital infrastructure on research output. We also contribute to the existing literature on digitalization in health care by analyzing the impact of large-scale digital research infrastructure projects on research productivity. Furthermore, we make a contribution to the literature on large teams and innovation output (Tortoriello, McEvily, & Krackhardt, 2015) by examining the organizational structure and complementarity in large teams. Finally, we discuss the policy implications of these findings, suggesting that initiatives like the HBP not only accelerate the pace of knowledge creation in critical fields such as brain science but also provide a model for designing future institutional interventions aimed at catalyzing radical innovation.

Chapter 2, co-authored with Lucy Xiaolu Wang, investigates the role and effectiveness of the HBP's publicly accessible online forum as a mechanism for fostering knowledge sharing and collaborative problem solving in neuroscience. In a context where digital tools increasingly complement traditional laboratory-based approaches in STEM research, the study explores how actively the HBP forum is used, what factors encourage richer user interaction, and how effectively the forum enables problem-solving among neuroscientists.

To answer these questions, we construct a novel dataset that includes all user interactions and discussions from the public HBP Forum. We supplement the data with user information based on user profiles, including demographics, institutional affiliation, and scientific field. Using regression analysis, we analyze the determinants of richer interactions, measured by the number of replies per post. In order to assess the time taken to solve a problem, we use a Cox proportional hazards model and include several covariates that have an influence on the time taken to solve a problem. These covariates encompass the technical nature of discussion threads (such as the inclusion of explicit code), the diversity of participating users, and the role of HBP administrators in facilitating discussions.

We find that the HBP Forum is actively used and that activity levels are unaffected during periods of disruption, such as the COVID-19 pandemic. Posts with technical details and users from different countries tend to lead to more active discussions. The participation of administrators in the discussion, as well as the sharing of code snippets, speeds up problem resolution.

This chapter contributes to our understanding of digital collaboration and knowledge exchange and extends the literature on quantitative studies of proprietary or sub-national platforms (e.g., Faraj, Jarvenpaa, & Majchrzak, 2011; Morrison, 2017; Gallus, Jung, & Lakhani, 2019) by providing a systematic empirical analysis of an international, research-focused discussion forum supported by a supranational institution. We also extend the literature on digitalization in healthcare (e.g., McCullough et al., 2010; Falotico et al., 2017; Freedman, Lin, & Prince, 2018; L. X. Wang, 2021) by shifting the focus from downstream users of health information technologies to upstream researchers, and by shedding light on how neuroscientists collaborate online.

Our study thus highlights the importance of online platforms in overcoming physical barriers to collaboration and enriching interdisciplinary exchanges in a rapidly evolving scientific field.

Chapter 3 examines the role of FabLabs in stimulating local innovation in Germany. FabLabs, conceptualized at the Massachusetts Institute of Technology (MIT) in 2003, are a global network of community workshops with the common goal of offering a standardized set of machines and tools for digital fabrication to everyone, fostering knowledge exchange and skill acquisition. Since the first FabLab opened up in Germany in 2009, the number of FabLabs has grown steadily. By bringing diverse people together and enabling them to ideate, develop prototypes, exchange knowledge, and network, FabLabs can enable tinkerers to become inventors and thus act as catalysts for innovation. Considering the increasing complexity of knowledge creation and the need for hands-on experience in complex and developing technologies such as Additive Manufacturing (AM), the innovation-democratizing approach of FabLabs may be one way to approach the knowledge frontier.

This chapter therefore analyzes whether and how FabLabs enhance regional innovation and how their effects vary across different regional settings. In particular, I investigate whether FabLabs increase local patenting activity and enable first-time inventors, whether university-integrated FabLabs have a stronger impact due to their proximity to cutting-edge research, and whether their effects differ between urban and rural areas. I also explore the interaction between FabLabs and incumbent Makerspaces and Hackerspaces and how their presence affects FabLabs' impact on inventive activities.

To gain insights into the organization of FabLabs and to inform the analysis, I conducted two online surveys targeting FabLab managers and users. To identify the causal impact of FabLabs on local inventive activity, I leverage their staggered rollout across German cities since 2009. Using a difference-in-difference framework, I compare regions with and without a FabLab at two geographic levels.The analysis relies on a novel dataset of patent and utility model applications with at least one German-based inventor from 2003-2019, aggregated at the FUA-year and NUTS3-year levels. I examine FabLabs' effects on the extensive margin of patenting – measured by total applications, first-time inventors, organizational and individual applicants, and AM-related patent – as well as on the intensive margin, captured by the number of forward citations and the number of technical areas cited.

While FabLabs do not significantly boost overall patenting activity, they have localized effects, particularly in fostering AM-related patents and supporting first-time inventors. University-integrated FabLabs are especially effective in promoting interdisciplinary and AM-related patents, while utility model applications tend to decline. Their impact depends on the regional preconditions. In urban areas, FabLabs stimulate inventive activity – particularly in AM technologies – by providing access to advanced tools and facilitating first-time inventors. In rural areas, they serve as catalysts for innovation, likely driven by the access to tools and a collaborative environment, though their effects are weaker and more delayed compared to urban settings. Complementary community workshops further amplify FabLabs' impact. In urban areas, Makerspaces and Hackerspaces create synergies, while in rural areas, FabLabs are most effective when filling gaps in local innovation infrastructure. Overall, FabLabs foster regional innovation, but their effectiveness depends on local conditions and complementary infrastructure.

This study contributes to the literature by extending our understanding of how institutions, that provide open access to specialized knowledge and tools, and promote knowledge sharing, can enable tinkerers to become inventors. While prior studies of community workshops have relied primarily on qualitative or small-scale analyses or startups, this study offers one of the first comprehensive, quantitative assessments the economic impact of FabLabs on innovation. The study further advances the science of science literature by highlighting how local institutions foster tacit knowledge exchange and collaboration essential for disruptive innovation. In addition, the study contributes to the literature on open and user innovation by empirically demonstrating that open-access, collaborative environments like FabLabs can democratize innovation and promote grassroots creativity.

The findings highlight the heterogeneous impacts of FabLabs in different regional contexts and thus contribute to a more nuanced understanding of the role of FabLabs in fostering innovation. These insights underscore the importance of tailoring policy interventions to regional contexts, to ensure that FabLabs can effectively complement and strengthen innovation ecosystems by providing spaces for ideation.

Chapter 4, co-authored with Jonathan Federle and Dietmar Harhoff, addresses the challenge of obtaining reliable patent value estimates, which is important for a wide array of empirical economic analyses. While acknowledging the limitation that patents do not encompass all economic sectors and not all inventions are patented, the patent system is recognized as the most comprehensive information source of information for monitoring technological advances. However, patents are inherently heterogeneous and their skewed distribution (Scherer, 1965; Harhoff, Scherer, & Vopel, 1998) poses a challenge to their assessment.

In this chapter, we propose an alternative approach to estimating patent values using stock market reactions. Our approach is based on the influential measure by Kogan et al. (2017, henceforth KPSS). KPSS employ the stock market reactions of publicly listed U.S. firms to patent grants by the USPTO. They estimate the value of individual patents by isolating the patent-related component of stock returns and calculating the average capital gain per granted patent after adjusting for market expectations. However, their approach is subject to limitations. Specifically, their assumption of a constant signal-to-noise ratio (SNR) and their averaging of patent value across patents granted to a firm in a given week introduce systematic biases into the estimated patent values.

To address these biases, we introduce two refinements. First, we adopt a dynamic SNR approach, allowing the patent-unrelated error to vary at the firm-day level and the patent-related variance to depend on the ratio of granted patents to the firm's pre-grant market value. In addition, we allow the SNR to depend on the number of simultaneously granted patents. Second, instead of simply averaging weekly profits over the number of patents granted to a firm, we apply a hedonic patent value regression to assign weights within weekly patent grant bundles using publicly known quality characteristics. These weights are then applied to allocate weekly profits among individual patents granted simultaneously to a firm. This approach is implemented on a newly constructed dataset encompassing nearly 3 million USPTO patent grants, which provides a comprehensive and up-to-date measure of patent values.

Our empirical results confirm that the revised value estimates effectively reduce the biases present in the original KPSS estimates. Introducing a dynamic SNR that also accounts for the number of simultaneously granted patents reveals a heterogeneity that is masked by KPSS' constant SNR. Our hedonic regression approach yields patent value estimates that are significantly lower in magnitude than those of KPSS. Moreover, our newly derived patent value estimates reflect the the skewed distribution of patent values observed in practice more accurately. In addition, our approach significantly improves the correlation between patent value estimates and patent quality characteristics not included in the hedonic regressions, such as forward citations, standard essentiality, and firm size. Overall, our results contribute to the literature by clarifying the underlying assumptions of the widely used KPSS value estimates while introducing significant improvements. Additionally, our findings may help advance discussions on the use of stock market reactions in economic research on innovation and competition.

In summary, this dissertation offers new insights into how digitalization and knowledge-access initiatives influence knowledge creation, interdisciplinary collaboration, and innovation outcomes while advancing methods to economically evaluate the codified output of innovation processes.

By examining both large-scale digital research infrastructures and decentralized physical community workshops, it emphasizes the evolving landscape of knowledge access in the digital age. The findings indicate that structured, top-down initiatives, such as the HBP, can spur research productivity and interdisciplinary collaboration. At the same time, grassroots, bottom-up institutions, such as FabLabs, promote regional innovation by providing access to digital fabrication tools and by enabling first-time inventors, thereby democratizing innovation. These insights can inform policymakers on how to foster innovation through an infrastructure that is tailored to different contexts. Furthermore, this dissertation proposes an improved method for estimating the economic value of patents, offering more precise estimates that are more consistent with prior literature and internally coherent. This improved approach can benefit researchers across fields. By bridging research on knowledge access, digitalization, and economic valuation, the dissertation contributes to a deeper understanding of how innovation processes can be effectively supported and evaluated in a rapidly evolving technological landscape.

1

Megaprojects, Digital Platforms, and Productivity – Evidence from the Human Brain Project

1.1 Introduction

Neurological diseases are one of the leading causes of death globally (Carroll, 2019). Most brain diseases are not curable, and the limited treatments are expensive: costs for dementia treatment average at €32,000 in the EU and $42,000 in the US per patient-year (Cantarero-Prieto et al., 2020). The economic costs of dementia were estimated to be €800 billion in Europe in 2010 and $1.5 trillion in the US in 2013 (N. Rose,

*This chapter is based on joint work with Lucy Xiaolu Wang. For this chapter, I used ChatGPT and DeepL-Write to refine grammar and wording.

2014; Nager & Atkinson, 2016). New treatment progress has been slow, and the only Alzheimer's drug approved in almost two decades has generated much controversy.[1] Meanwhile, artificial intelligence (AI) is found to speed up pre-clinical drug discovery for drugs at medium levels of chemical novelty (Lou & Wu, 2021). Given the decline in productivity in biomedical science and increasing challenges to reach the frontier of knowledge (Jones, 2009; Bloom et al., 2020), creating non-market incentives to accelerate collaboration in multi-disciplinary research and pushing the frontiers are particularly important.

To enhance brain science with AI, the Human Brain Project (HBP) was launched in 2013 as a ten-year, €1 billion flagship initiative in the European Commission's Future and Emerging Technologies program. The HBP seeked to build cutting-edge information technology (IT) infrastructure to advance brain science with high-power computing and to facilitate collaborations between scientific and industrial researchers.[2] Among other things, the HBP built super-computer-powered online research platforms, provided funding for partnership institutions, and offered educational training and an online discussion forum (Kreyer & Wang, 2022). As of early 2023, more than 500 scientists and engineers at over 180 institutions (e.g., universities, hospitals, research centers) in more than 20 countries across Europe and beyond have engaged in HBP-relevant research activities, making it one of the largest life science research projects in the world and the largest research project in the EU (Lorents et al., 2023; Siva, 2023).

How does the HBP influence the rate and direction of R&D in brain science? Do researchers actively engaged in the HBP produce more, higher-quality, and/or more interdisciplinary work? And how does the impact of HBP differ across researchers' career stages, gender, and sub-fields within neuroscience/Computer Science (CS)? Intuitively, productivity can go in different directions. Researchers can become more productive by exploring new areas and expanding their network. Meanwhile, taking more risks and investing in new areas can take time and result in lower productivity. Alternatively, productivity may stay similar, as the time and energy is limited, but researchers substitute research areas. To investigate these questions, we construct a new dataset that tracks the set of individuals involved in the HBP, and infrastructure access, and research output (i.e., publications). We collect data from various sources, including HBP official websites, audits, EU's grant reporting systems, period reports,

[1] Source:https://www.medicalnewstoday.com/articles/progress-and-controversy-in-alzheimers-research-aducanumabs-fda-approval#Conflicting-trial-results-for-aducanumab.

[2] Source: https://www.humanbrainproject.eu/en/.

individual profiles, and Scopus.

Evaluating the impact of the HBP is challenging. First, finding a paper trail to measure the impact of research R&D is hard, especially with the spillover effects in knowledge. Second, identifying appropriate control groups is difficult, in the presence of cross-subsidy among researchers and unobservables. To address the challenges in measurement, we employ techniques developed in natural language processing (NLP), especially the neural, prompt-based Large Language Models (LLM). We exploit plausible exogenous variation based on the staged nature of the HBP, institutional details, and sudden revamp in the evolution of the HBP. We use a staggered difference-in-differences (DiD) method with two-way fixed effects to account for selection, and we discuss the interpretations given the specific contexts.

We compiled a novel and comprehensive individual-level dataset of 639 individuals actively engaged in the HBP during at least one of the phases, sourced from phase-specific deliverables and reports.[3] The data contain rich information on scholar names, seniority level by HBP phases, gender, organizational affiliation, and country of affiliation. Utilizing the Scopus database, we retrieved 39,524 publications from 2008-2022 for 639 active researchers with publication records. We then use NLP tools to classify each publication into four main topic categories: fundamental neuroscience, neurotechnology, AI-robotics, and patient care. In addition to citation data, we also use whether a publication is in a top neuroscience journal or CS/AI proceedings to proxy for high quality. Unique co-authors per author-year are computed to proxy for effective annual collaborative network size. We aggregate the publication data to the author-year level to allow for systematic comparison in balanced panels.

We observe an increase in new participation of junior faculty, graduate students, and female researchers joining the project over time. This coincides with the overall rise in active researchers in the HBP. We find that active involvement in the HBP increases individual scholars' productivity (number of publications per author-year) and follow-on citations, particularly driven by junior scholars (including junior faculty and graduate students). This suggests the significant benefits derived from the research infrastructure and collaboration opportunities facilitated by the HBP. Female researchers also experience a noteworthy boost in productivity, measured by per author-year publications and citations received, despite of lower estimated magnitudes.

Furthermore, an examination of the quality and publication outlets indicates that pub-

[3] We identified 824 active individuals from the reports and deliverables. Of these, 639 have a Scopus ID, enabling us to retrieve their publication history.

lications by those engaged in the HBP are more likely to appear in top neuroscience journals, predominantly within the areas of neurobiology and neurotechnology. Overall, engaging in the HBP appears to be the most beneficial for publications in the neurotechnology topic areas. Researchers with a background in natural/life sciences work more on neurotechnological topics areas. Analyzing the heterogeneity by the country of affiliation at the time of the first HBP participation unveils a positive correlation for all countries, with German, Italian, and Belgian-based scholars wielding a larger increase in the number of publications per author-year. This comprehensive analysis provides valuable insights into the multifaceted impacts of HBP participation on researchers and their scholarly output.

This paper contributes to three strands in the literature. First, our paper contributes to research on institutions and innovation by examining the design of new institutions that support long-term R&D. Despite the huge social value of innovation, markets tend to under-incentivize innovation given the public good nature of ideas (Nelson, 1959; Arrow, 1962), and private firms tend to under-invest in long-term research (Budish, Roin, & Williams, 2015). Institutions play an important complementary role to market incentives (e.g., tax, royalties) in the advance of sciences by enhancing open knowledge diffusion and reducing research costs (Furman & Stern, 2011; Williams, 2013; Murray et al., 2016; L. X. Wang, 2022). Institutions providing long-term grants can enhance the production of creative papers and successful trainees (Azoulay, Graff Zivin, & Manso, 2011). In addition, digital infrastructure removed the physical barriers and can boost human capital (Boudreau, 2010; Huang et al., 2022), but the largest effects from IT investments often take several years to materialize (Brynjolfsson & Hitt, 2003). Our paper offers the first systematic analysis of the effectiveness of the HBP, which provides digital infrastructure building besides regular grants to support research projects.

Second, this paper expands on the literature on the science of science by analyzing how the structure of large-scale megaprojects can enhance complementarity in organizations. As innovation often stems from knowledge recombination (Nelson & Winter, 1985), teams with inventors of diverse knowledge can benefit from complementarity (Tortoriello, McEvily, & Krackhardt, 2015), especially in teams with both generalists and specialists inventing in uncertain contexts (Melero & Palomeras, 2015). In addition, cross-firm R&D collaboration is well established to compensate for the lack of resources including knowledge (Hagedoorn, 2002). Knowledge transfers often happen among inventors from different organizations (Agrawal, Cockburn, & McHale, 2006), and complementarity between technological and commercial assets can drive successful high-tech alliances (Colombo, Grilli, & Piva, 2006). Furthermore, evidence shows that R&D productivity increased in firms that adopted open science practices to motivate and reward scientists (Cockburn & Henderson, 1994). This paper documents the within and cross-organization structure of scientific project teams and expands on the literature by examining organizational structure and complementarity in large teams.

Third, this paper extends the literature on digitalization in health care markets by investigating the impact of digitalization on upstream R&D in life sciences. Prior studies find that digitizing workforce through health information technologies helps physicians, improves clinical quality, and reduces mortality or morbidity among infants, elderly, and opioid users (Miller & Tucker, 2011; McCullough, Parente, & Town, 2016; Freedman, Lin, & Prince, 2018; Lu, Rui, & Seidmann, 2018; L. X. Wang, 2021; L. X. Wang & Bloch, 2023). The benefits come with increased costs (Agha, 2014), which can be mitigated with complementary assets in digital human capital and infrastructure (Dranove et al., 2014). However, health AI adoption remains relatively low in health care markets compared to other industries (Goldfarb, Taska, & Teodoridis, 2020). As a less explored area, combining neuroscience with AI tools on digital platforms can offer new channels to advance sciences (Falotico et al., 2017). Our paper extends the literature by examining the causal impact of a massive digital research infrastructure on research productivity.

Finally, this paper has broad policy implications. The HBP is one of the earliest and biggest brain initiatives in a global "brain race" since 2013, followed by various brain initiatives launched in the US, Israel, Japan, China (Grillner et al., 2016). Beyond brain sciences, digital laboratories have become increasingly important in offering advanced AI tools and building resilience during physical disruption, including the National Virtual Biotechnology Laboratory (NVBL) created during COVID-19.[4] Furthermore, there has been a global rise of "megaprojects" (i.e., projects that cost over $1 billion), including the US-led Human Genome Project, the EU's Quantum Flagship, and China Desert Project. Those projects often involve both infrastructure building and grant awards across a large amount of stakeholders. How to design these projects to maximize the long-term impact while evaluate intermediate progress remains largely an open question. Our study on the HBP can help inform policy related to such institutions with the potential to boost radical innovation.

[4]The NVBL is a consortium of national laboratories supported by the United States Department of Energy. Source: https://science.osti.gov/nvbl.

The paper proceeds as follows. Section 1.2 describes the background and conceptual considerations. Section 1.3 discusses data and dataset construction. Section 1.4 reports empirical strategies and results. Section 1.5 concludes.

1.2 Conceptual Considerations and Background

1.2.1 Conceptual Considerations

Innovation is harder to get with the slowdown in productivity and increasing burden to reach the frontier of knowledge (Jones, 2009; Bloom et al., 2020). These challenges can be mitigated by large research teams of members with complementary skills that combine specialists in diverse research areas (Jones, 2009). Team members from different social worlds also learn from exposure to a diverse source of information (Burt, 2004), and effective project teams benefit from strong ties within team members as well as ongoing relationships with members from other project teams (Reagans, Zuckerman, & McEvily, 2004). However, moral hazard with credit sharing often creates fundamental inefficiency in teams (Che & Yoo, 2001). In fact, evidence suggests large teams are becoming more prevalent in science, yet they are less likely to pursue highly disruptive research (Wu, Wang, & Evans, 2019). It remains an open question on the optimal team size for radical innovation.

In addition, most innovation projects are risky, unpredictable, long-term, labor-intensive, and idiosyncratic (Holmstrom, 1989), making effective contracting particularly challenging. Empirical evidence suggests that radical innovation projects require long-term investment, although such long-term support is often absent in standard funding systems (Azoulay, Graff Zivin, & Manso, 2011). In addition, private firms tend to under-invest in long-term, high-value projects (Budish, Roin, & Williams, 2015). In contrast, the international effort in a long-term megaproject, the Human Genome Project, shows the value of international collaboration and open knowledge sharing in spurring follow-on innovation (Williams, 2013). However, megaprojects often underperform and generate negative economic and social consequences (Denicol, Davies, & Krystallis, 2020), imposing concerns on project management.[5]

[5]Megaprojects are typically characterized as long-term endeavors, whether in the private or public sector, with a funding exceeding US$1 billion. These initiatives are overseen and managed by a newly established organization, bringing together diverse participants united by a shared objective. A megaproject may encompass multiple sub-projects.(Denicol, Davies, & Krystallis, 2020).

As the first large-scale megaproject in biology, the Human Genome Project (HGP, 1990-2003) is a success of international collaboration of researchers based in the United States, United Kingdom, France, Germany, Japan, and China. However, the HGP faced many obstacles at the early stage, including developing the scientific tools, competing for funding and other resources, centering on a small number of large groups, and much contention about how to carry out the work.As an extension project, HGP-Write, was proposed to spur the development and testing of human genomes in 2016, yet the initial bold plan was downsized and refocused. In the private sector, IBM Watson Health was deemed as "future of health care" in 2011 when established as an application of AI in the healthcare sector. However, the partnership with MD Anderson cancer center fell apart due to technical issues and cost overrun, resulting in a more than $1 billion sale off for parts.[6]

All those events lead to questions: How can we know if long-term large-scale megaprojects will succeed? Improving non-market institutions and the organization of sciences is critical for the innovation system. While some design issues can be adjusted at the initial stages of megaprojects, developing intermediate evaluation metrics is even more critical for long-term, high-risk, multi-stage projects. The HBP offers a valuable setting, where the initial goal is adjusted, the impact decades ahead is hard to predict yet, and the intermediate impact is extremely valuable to study.

1.2.2 Background of the Human Brain Project (HBP)

The European Commission (EC) called for proposals for two Flagship Projects under the Future and Emerging Technologies (FET) Program 2013 to enhance EU technology competitiveness. Each flagship project is funded €1 billion ($1.33 billion in 2013) over 10 years and the HBP was a winner.[7] The fund for HBP was dispersed over four project phases: ramp-up (phase 0, 2013.10-2016.3), followed by three special agreement phases (each lasts two years), with phase 3 ended in March 2023, and the whole HBP closing in September 2023 (see Fig. 1.1). The initial goal was to simulate the entire human brain in a super-computer to develop treatments for brain diseases (Markram et al., 2011). However, the singular goal, modeling approach, and lack of

[6]Source: 1) HGP: https://library.cshl.edu/oralhistory/topic/genome-research/challenges-hgp/; https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genome-project; 2) HGP-write: https://www.science.org/doi/10.1126/science.aaf6850; https://www.nature.com/articles/d41586-018-05043-x; 3) IBM Watson: https://slate.com/technology/2022/01/ibm-watson-health-failure-artificial-intelligence.html#:~:text=One%20of%20IBM's%20high%2Dprofile,was%20later%20audited%20and%20shelved.

[7]Sources: https://www.nature.com/articles/nrn3578; https://web.archive.org/web/20161116151222/http://cordis.europa.eu/fp7/ict/programme/fet/flagship/doc/press28jan13-02_en.pdf

transparency were criticized from the beginning, intensified into an open letter signed by over 800 neuroscientists by 2014, and resulted in the revamp of the management team and an adaption of the goal in 2015. From then on, the goal was to advance brain science with computational tools and build a European digital research infrastructure – “brain-inspired information technology”.[8]

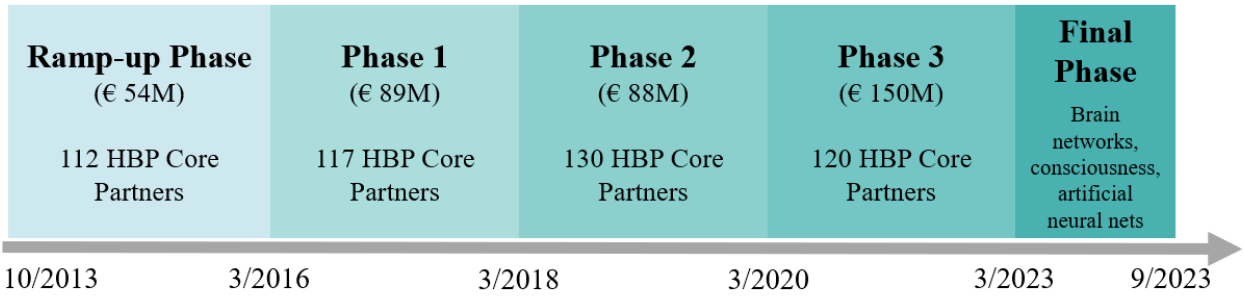


Figure 1.1: HBP timeline: phases, grants, partners

Notes: Author’s graph. The timeline depicts the project phases, the EU grants per phase and the number of HBP *core project partners* per phase.

The HBP consists of a *core project* funded by the EU and *partnering projects* funded mainly by the member states (Figure 1.2).[9] The *core project partners* coordinate and execute the HBP’s research plan and ensure the development of the digital infrastructure.[10] Current core project partners develop a work plan for the subsequent phase outlining the core direction and goals that form the basis for a “call for proposals”, which other institutions can apply to join. The work plan assigns tasks for existing core project partners continuing to the next phase, or terminates the activity and funding for some partners by the end of a phase. Institutions eligible to become core project partners must be in EU member states or eligible countries (e.g., Switzerland, Israel,

[8] Sources: 1) A background summary: https:/ / theconversation.com/ building-better-brain-collaboration-online-despite-scientific-squabbles-the-decade-long-human-brain-project-brought-measurable-success-to-neuroscience-collaboration-196473; 2) The NeuroFuture Letter and Mediation report: https:/ / web.archive.org/ web/ 20191207051843/ https:/ / www.neurofuture.eu/ ; https:/ / web.archive.org/ web/ 20160702065551/ http:/ / www.fz-juelich.de/ SharedDocs/ Downloads/ PORTAL/ DE/ pressedownloads/ 2015/ 15-03-19hbp-recommendations.pdf;jsessionid=05A0C89E3D1AAE2761604DA7ECB9CD1A?__blob=publicationFile

[9] Source: https://digital-strategy.ec.europa.eu/en/library/fet-flagship-model-implementation-and-governance-model-horizon-2020-short-overview-and-presentation

[10] HBP researchers work in 12 subprojects (SP): SP 1-4 focus on understanding the brain, SP 5-10 develop the digital infrastructure, SP 11-12 manage project management and ethical, social, and legal issues (Kreyer & Wang, 2022).

and United Kingdom post-Brexit).[11] A proposal must meet the criteria for unique research requiring tight integration across disciplines. Proposals are evaluated by three external experts and then by the EC for final decisions.

In addition, *partnering projects* complement the infrastructure development and research process by the core project partners. Partner project teams already conduct research in fields related to the HBP, and receive funding from non-HBP sources: national, regional, or other eligible sources. These projects contribute knowledge, skills, and resources to the core project for specific research tasks. Partnering projects can be proposed anytime and come from various sources, including collaborators of the core project partners, EC, funding agencies, private sector, or self-nomination. Partner projects will be reviewed and evaluated by core project partners based on alignment with subproject goals and potential contributions, and decided by the HBP science and infrastructure board on inclusion.[12]

Researchers in core and partnering project institutions collaborate to develop and use the online research infrastructure, which provides cloud-based data storage, software, computing tools, and communication channels. External researchers may apply for access. The current infrastructure, EBRAINS, hosts multiple advanced platforms powered by the EU neuroscience supercomputing center, and services as the foundation for a pan-European neuroscience research platform hosting HBP data, tools, and models securely accessible to researchers.

1.3 Data

1.3.1 Data on the HBP and Partnership Institutions

We obtained the details on HBP core project partners (department/unit, if available) regarding the timing of their engagement (entry/exit), the extent and expectation of

[11] Non-EU countries eligible for funding under Horizon 2020: Iceland, Norway, Albania, Bosnia and Herzegovina, North Macedonia, Montenegro, Serbia, Turkey, Israel, Moldova, Switzerland, Faroe Islands, Ukraine, Tunisia, Georgia, Armenia, Source:https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/3cpart/h2020-hi-list-ac_en.pdf. The United Kingdom remains eligible for EU funding after Brexit, Source: https://www.gov.uk/government/publications/continued-uk-participation-in-eu-programmes Sources on application, selection and funding process: https://ec.europa.eu/research/participants/data/ref/h2020/sgl/erc/h2020-erc-se-rules_en.pdf;https://ec.europa.eu/research/participants/data/ref/h2020/sgl/erc/h2020-erc-se-rules-amended2_en.pdf

[12] At least two core project partners evaluate the application. These are typically (deputy-)leaders of the subproject the applicant intends to contribute to. The HBP science and infrastructure board consists of the subproject leaders and one representative of the partnering projects. Source: HBP Framework Partnership Agreement.

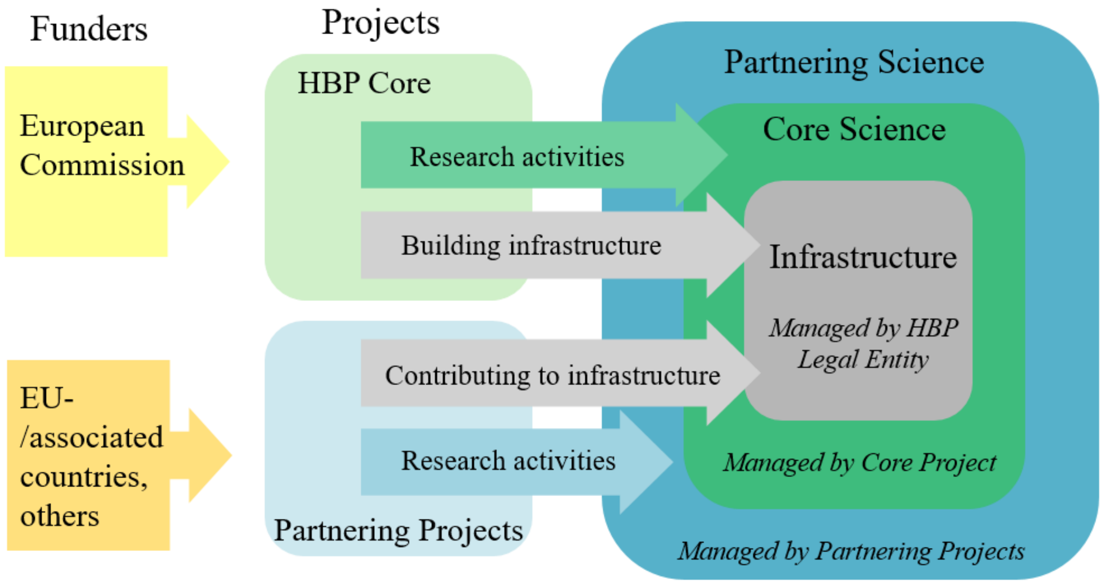


Figure 1.2: HBP partnership structure: funders, projects, and activities

Notes: This figure is adapted from HBP official Framework Partnership Agreement, released on 2015.10.30.

HBP-related activities, software and hardware evaluations, and major personnel. We collected these data from HBP official websites, end-of-phase project reports, framework partnership agreements and amendments, and European Commission's online manual for funding and tender opportunities.[13] Institutions are broadly defined in this context, including universities, teaching hospitals, public research institutes, and private corporations. We coded in the geographic location and associated institutional code of each institution. In addition to the core HBP partnership institutions, we also collected data on associate member institutions that indirectly benefited from the HBP, often as a collaborative institution with any HBP partnership institutions through HBP-relevant joint projects.[14]

In addition, we obtained HBP phase-institution-specific funding amounts, reports, and deliverables from the Community Research and Development Information Service (CORDIS), the European Commission's primary repository of EU-funded research and innovation projects, managed by the Publications Office of the European Union.[15]

[13] Source: https://www.humanbrainproject.eu/en/open-ethical-engaged/contributors/partners/; https://www.humanbrainproject.eu/en/about-hbp/project-structure/governance/framework-partnership-agreement/; https://webgate.ec.europa.eu/funding-tenders-opportunities/display/OM/Admissibility+and+eligibility+check.

[14] For example, the Allen Institute and McGill University are associate members but not core HBP partners. Source: https://www.humanbrainproject.eu/en/collaborate-hbp/partnering-projects/associated-members/.

[15] Sources: CORDIS HBP records by phase: 1) phase 0: https://cordis.europa.eu/project/id/604102; 2) phase 1: https://cordis.europa.eu/project/id/720270; 3) phase 2: https://cordis.europa.eu/project/id/785907; 4) phase 3: https://cordis.europa.eu/project/id/945539.

Most grants and core partner institutions are distributed widely across EU membership countries (Figure 1.3), joined by Switzerland and Israel. Germany leads in both grant amount and the number of partners, but the correlation between grants and partnership is not strong. For example, Spain and Turkey receive no grant but have active institutional participation.

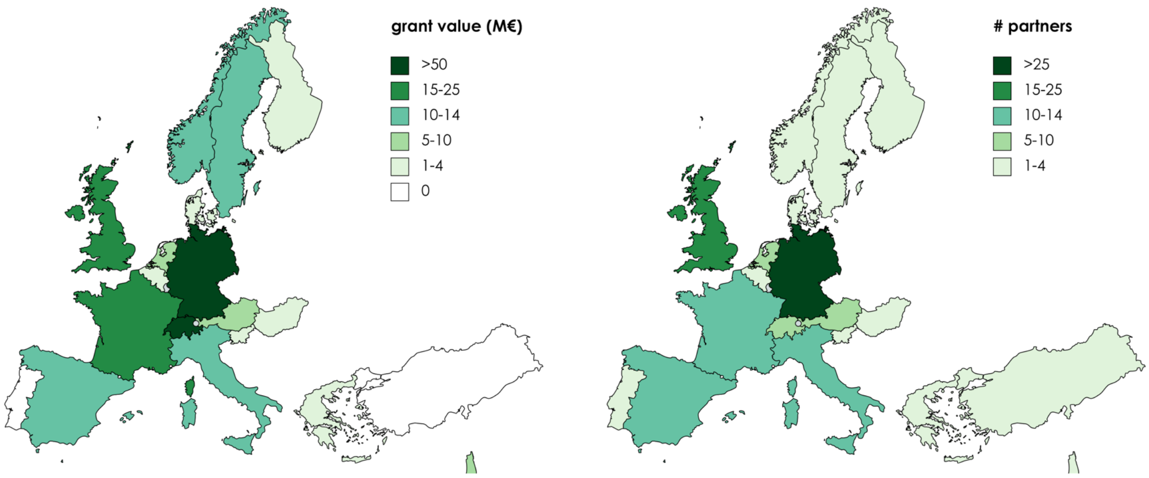


Figure 1.3: Geographic distribution of HBP grant allocations and partnership institutions

Notes: This figure depicts HBP grants and partner institutions as of Spring 2021, which includes complete information on grants and institutions that joined the HBP by the end of phase 2.

1.3.2 Individual-Level Data

We collected a comprehensive dataset of 639 individuals actively involved in the HBP during at least one of the first three phases. These individuals were collected based on HBP deliverables and reports of each sub-project/working package during each project phase, HBP PLUS user profile database maintained by the HBP, HBP websites, and HBP YouTube channels. For each individual, we collected details via online searches on gender, research areas, seniority level during each of the four HBP phases, organizational affiliation, and the country base of work.[16] Further details are collected from various sources, including institutional websites, Google Scholar, ResearchGate, LinkedIn, Xing, and Loop. We performed careful data curation to record each individual's occupation and institutional affiliation in each HBP phase, as some people

[16]When not listed, gender is coded based on first name and individual website information (e.g., image, statement).

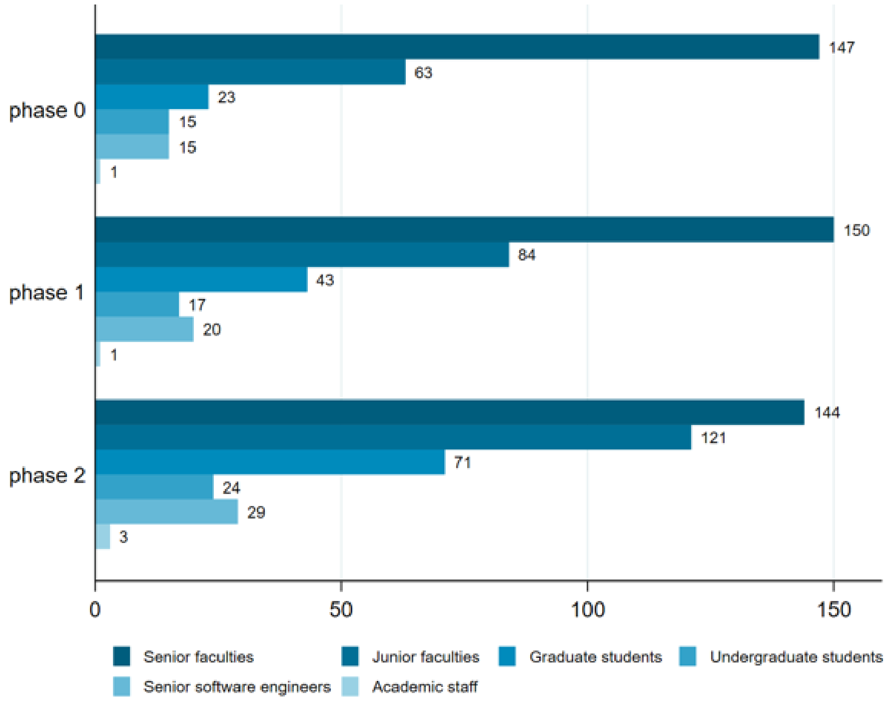
changed affiliations during our sample period, which is captured by our publication data discussed below.

We recorded seniority levels and fields to characterize the composition and diversity of scientific teams. We define seven levels of seniority, including four categories for individual researchers (undergraduate student, graduate student, junior faculty, and senior faculty), two categories for individuals involved in research but are not researchers themselves (senior software developer, academic staff), and non-academics for individuals working outside academia. For senior researchers outside academia, we also use their publication records to discern the level of research seniority.[17]

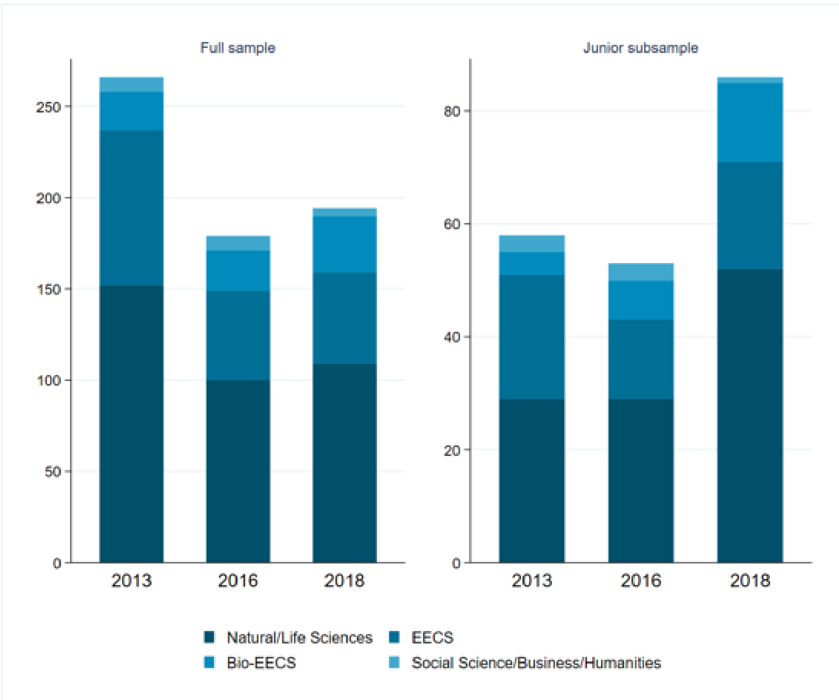
We recorded the highest degree field for each researcher. For 11 researchers, we used early-stage publications to determine their research focus areas. We then aggregated the 372 obtained fields into four broader categories, as shown in Table: *EECS*, *bio-EECS*, *natural and life sciences*, and *social science, business, humanities*.[18] The majority of HBP participants obtained their highest degree in natural and life sciences or EECS. The bar chart for junior researchers joining the HBP indicates that particularly in 2018, most juniors received their highest degree in natural/life sciences (see Fig. 1.4, Panel B). We use this information to form some qualitative basis of how many interdisciplinary areas are covered by researchers involved in the HBP.

[17] Junior: individuals with a post-doc or assistant professorship who have not been working at the same institution for more than 4 years and who have not had more than 4 career stages after finishing Ph.D. Senior: full or associate professors, research directors, or individuals that have either worked for more than 4 years at the same institute or have had more than four career stages after they finished their PhD. In some cases, we also categorized the seniority status based on their publication list. The seniority status always gives the level of seniority at the beginning of the given Phase. Specifically, all senior software developers are affiliated with a research institution; academic staffs are individuals working at research institutions but not directly involved in research, e.g., project management staff

[18] The detailed degree names of the 372 degrees we collected for the individuals are available upon request.



Panel A: Scientists by phase and seniority level



Panel B: Highest degree fields by individual entering phase

Figure 1.4: Scientists by seniority and field of training

Notes: Panel A of this figure depicts the seniority level of the sample of active researchers associated with the HBP during the ramp-up phase, the first two phases (SGA1 and SGA2). Panel B illustrates the highest degree fields for each individual entering phase into the HBP. The left-hand figure represents the full sample, while the right-hand figure focuses on the junior subsample.

Our final data contain 639 individuals involved in the HBP during phases 0, 1, and 2. In this sample, 264 researchers are active in phase 0, 315 in phase 1, and 392 in phase 2. 20% are female researchers. On average, 60% of the researchers are affiliated with either German, French, Swiss, or British research institutions. In phases 0 and 1, on average 52% are senior scientists and 25% are junior researchers. In phase 2, 37% are senior scientists and 31% are junior scientists. In all three phases, 6% of the active individuals are senior software engineers affiliated with a research institution. The percentage of participating graduate students increases from 9% in phase 0 to 18% in phase 2 (see Fig. 1.4, Panel A). 56.49% of researchers attained their highest degree in the field of *natural and life sciences*, 28.79% in *Electrical Engineering & Computer Science (EECS)*, 11.58% in *bio-EECS interdisciplinary degrees*, and 3.13% in *social sciences, business, or humanities*. We are also collecting individual information on the 815 individuals who signed the Neurofuture letter that led to the revamp of the HBP in 2016.

1.3.3 Publication Data

For our bibliometric analysis we use Scopus as the main source. Besides being one of the largest curated collections of abstracts and citations, Scopus also assigns unique author IDs to authors and provides their affiliation as well as supplementary information on institutional addresses (Mongeon & Paul-Hus, 2016; Baas et al., 2020). We have identified 639 author IDs in Scopus for HBP-relevant scientists and successfully calibrated the Scopus data extraction algorithms (API access key via MPI institutional license).[19]

For the 639 Scopus IDs we have recovered 39,524 publications between 2008 and 2022 from Scopus using the pybliometrics package (M. E. Rose & Kitchin, 2019). 63.25% are articles and 23.01% are Conference Papers. The other publication types are Reviews, Book Chapters, Editorials, and Letters. For 58.17 % of the publications funding information is reported by the authors. Based on these information, 58.17% received funding and out of these 3,746 (16.29%) publications received funding through the Horizon 2020 program, a European funding program of which the HBP is part of. For 41.83% of the publications, no funding is reported.

[19] Scopus' algorithm occasionally creates multiple profiles for a researcher. In our data, 13 active researchers have two author IDs assigned by Scopus. We merge these profiles in our dataset and check for duplicates in the publications.

1.4 Empirical Model and Results

1.4.1 Empirical Strategies

We first discuss our baseline outcome variables across empirical models, then we layout the set of empirical models we adopt, followed by additional measurement construction to tease out the mechanism in additional tests. Before the benchmark models, we plan to perform two-sample hypothesis testing at the publication-level analysis and test the hypotheses that HBP affiliation status (time-varying) is associated with higher level of output.

Baseline outcomes: We use several outcome variables (in raw numbers and in logarithm), including the number of publications, number of publications as first or last author, and number of other deliverables. We further distinguish citations by journal and conference proceedings, as the standard publication outlets for neuroscience and computer sciences are different. We identified top journal outlets for neurosciences and top conference outlets for computer sciences.[20] The dependent variables for our analyses are aggregated on author-year level (see Table 1.1). Aggregating publications at the author-year level resolves concerns about double-counting when co-authors are also HBP participants. Additionally, examining publications as first or last author helps mitigate this duplication within our sample.

[20] We classified the following neuroscientific journals as top-tier: Science, Nature, Nature Neuroscience, Neuron, eLife, Current Biology, Journal of Neuroscience, Journal of Neurophysiology , Proceedings of the Royal Society B, Proceedings of the National Academy of Sciences; Developmental Neuroscience, and Cortex. Top computer science outlets were categorized based on the A- and A* ranked journals provided by the CORE Conference DB.

Table 1.1: Summary statistics

Panel A: Individual level			*Panel B: Publication level*		
	Freq.	%		Freq.	%
Total	639	100	Total pubs	39,524	100
Female	129	20.19	*Type of publication*		
Male	511	79.81	Article	24,998	63.25
Active researchers per phase			Conference paper	9,095	23.01
Phase 0	266	41.63	Review	2,107	5.33
Phase 1	318	49.77	Book chapter	1,087	2.75
Phase 2	393	61.60	Other (e.g., editorial, letter)	2,237	5.66
Countries of affiliation			*Type of journal*		
Phase 0	23		Journal	29,091	73.60
Phase 1	20		Conference proceedings	7,129	18.04
Phase 2	21		Book / book series	3,294	8.33
			Trade journals	10	0.03

Panel C: Author-year panel					
Dep. Variable (author-year)	N	Mean	Std. dev.	Min	Max
Panel C1: Publications					
#pubs	9,585	4.12	7.24	0	131
#pubs as first author	9,585	0.35	0.79	0	15
#pubs as last author	9,585	1.38	3.29	0	61
#distinct co-authors	9,585	24.50	66.59	0	1105
#co-authors	9,585	48.93	330.12	0	12731
avg. #co-authors	9,585	5.37	8.96	0	100
#citations	9,585	271.55	677.56	0	11801
Panel C2: Journal quality					
Top Neuro	9,585	0.27	0.77	0	9
Top CS	9,585	0.11	0.58	0	11
CS A*	9,585	0.04	0.36	0	10
CS A	9,585	0.08	0.43	0	8
Panel C3: Probability of topic classification					
Neurobio	9,585	0.24	0.32	0	1
Neurotech	9,585	0.25	0.32	0	1
AI-Robotics	9,585	0.06	0.15	0	1
Clinical	9,585	0.10	0.20	0	1
Panel C4: Topic classification with the highest probability					
Neurobio	9,585	1.71	3.94	0	46
Neurotech	9,585	1.45	3.57	0	97
AI-Robotics	9,585	0.44	2.02	0	55
Clinical	9,585	0.73	2.27	0	44

Staggered DiD models To estimate the causal impact of the HBP on research productivity, we first use the standard specification for difference-in-differences (DiD) with staggered timing as the benchmark model. We exploit variation in the timing of individual access to HBP resources and compare individuals access to the HBP earlier to these later or not yet at the end of our sample period.[21] We estimate the following equation:

$$y_{it} = \beta HBP_{it} + \delta_i + \delta_t + (X_{it}) + \varepsilon_{it} \tag{1.1}$$

Subscripts i and t denote the individual and year, respectively. y_{it} is the outcome of interest. HBP_{it} is an indicator for whether an individual had ever participated in the HBP. δ_i and δ_t denote individual and year fixed effects, respectively. Standard errors are clustered at the individual level to allow for arbitrary autocorrelation within an individual. Because recent staggered DiD research cautions against time-varying covariates, we report main results with the key HBP variable and extensive fixed effects.

Two identifying assumptions are required: common trends and lack of common shocks. We test the common trends assumption using an event study model and we then elaborate with institutional knowledge on why the data generation process is driven by factors plausibly orthogonal to our outcomes of interest. The event study framework examines the validity of the common trends assumption that the trend in the control group (individual-years outside the HBP) is a valid counterfactual for the treated group (individual-years with access to HBP resources). Differential trends of outcomes between the treated and control groups in the pre-treatment period would suggest endogeneity HBP affiliation or potential correlation with other unobserved shocks. In addition, the event study reveals the dynamic responses of outcomes to HBP access – whether the effects build over time, stay constant, or fade away. We estimate the standard event study models as follows:

$$y_{it} = \sum_{k \in T} \beta_k 1\left\{HBP\ event\ time_k\right\}_{it} + \delta_i + \delta_t + \varepsilon_{it} \tag{1.2}$$

where β_k denotes the difference between treated and control units in the period k years relative to when an individual has access to HBP resources. The last period

[21] Those not-yet accessed individuals are individuals joined the HBP during phase 3, outside our main sample.

before access, $k = -1$, is omitted as the reference period. The event window T is specified to be #5 years before and #5 years after the HBP event.[22]

Additional measures However, direct publication count and citations do not capture deep nuances in the technological distance to HBP core areas, novelty, and disruptiveness of a research. For example, an interdisciplinary research may be pioneering but the initial citation may be scant, as evidence shows that many impactful interdisciplinary work (esp. done by junior faculty) has few citations early on, but the citation count catches up 10 years later (J. Wang, Thijs, & Glänzel, 2015). In the initial step, we utilize a combination of GPT-3.5turbo and GPT-4 to categorize publications into four topics: "Neurobiology" (abbr. Neurobio), "Neurotechnology" (abbr. Neurotech), "AI-Robotics," and "Clinical Research" (abbr. Clinical). The utilization of neural prompt-based Large Language Models (LLM) like GPT-3.5turbo and GPT-4 provides a significant advantage in generating topics that are interpretable for humans (Pham et al., 2023). While our goal is not to generate new topics but rather to classify, the use of these models proves immensely valuable, as LLMs have demonstrated great potential in text classification tasks (Clavie et al., 2023; Korinek, 2023; Viswanathan et al., 2023). Especially in few-shot classification tasks, where models are provided with examples of the desired outcome to learn within the context, LLMs like GPT-3.5turbo and GPT-4 perform very well(Brown et al., 2020; Chae & Davidson, 2023; Loukas et al., 2023).

In our approach, we provided a list of 31 keywords and incorporated this list, along with an example of a successful classification, into the prompt. For example, "neurobiology" topic area includes fundamental neurosciences (e.g., cognitive neurosciences) and neurobiologically-focused areas (e.g., neurogenetics). The "AI-robotics" areas including topics related to AI and machine learning. The interdisciplinary area termed "neurotechnology" ("neurotech" for short) covers areas in the intersection of technology and neuroscience, including computational biology, neuroinformatics, and neuromorphic computing. Table A.1 provides the details. We instructed the model to assign weights to each keyword for every abstract and title combination, with the cumulative weights per title and abstract summing up to one. The minimum probability assigned for any given keyword is 0.02. For our analyses, we aggregated the weights to the four topic area levels.

[22] Sample sizes fall small outside this range. Data before and after the event window are recoded to $k = -\#5$ and $k = -\#5$, respectively.

Different specifications In our main analyses, we use linear/log-linear specifications to best incorporate the recent methodological literature that mainly rely on linear models.

1.4.2 Results

Main Results on Research Productivity and Topic Involvement

Researchers actively involved in the HBP demonstrate, on average, higher productivity and publish more either as first or last author (Table 1.2, Panel A). Additionally, participation in the HBP expands researchers' networks and those researchers receive more citations. Specifically, the number of publications increases by 14% per author-year. Of these, 3% are publications as the first author, and 4% are publications as the last author. Concurrently, there is a 29% increase in the unique number of co-authors per author-year. This suggests that the increased productivity is, at least in part, explained by the expansion of the collaborative network. Additionally, the larger number of publications has garnered approximately 28% more citations. The coefficients for *#pubs*, *#pubs as first author*, *#co-authors*, and *#citations* in Panel B are notably larger for junior faculty and graduate students, suggesting that they derive greater benefits from the research infrastructure and collaboration facilitated by the HBP. Junior researchers have increased the number of publications by 31%, with 9% of these as first author. Additionally, their collaborative network has expanded by 63%, and their publications have received 63% more citations. Therefore, the positive impact of participation in the HBP on productivity appears to be primarily driven by junior researchers. We observe similar trends for females overall (Panel C), although the magnitude is less precise due to lower statistical power.

As illustrated by the event study estimates in Figures A.5-A.7, the parallel trend assumption holds for most specifications up to four years before HBP participation. The significant trend observed five years prior to HBP participation is likely due to limited statistical power. However, in the full sample, citation counts exhibit statistically significant pre-trends (see Figure A.5, Panel (e)). The estimates further suggest that the positive effects on publications, last authorship, and co-authorship are most pronounced in the first four years after researchers gained access to the HBP. In the junior and female subsamples, no significant pre-trends are observed for most variables. An exception in the junior subsample is the number of distinct co-authors, which exhibits a slightly positive and statistically significant anticipation effect, as well as citation counts, which display significant pre-trends (see Figure A.6, Panels (d) and (e) respectively). In the female subsample, the number of distinct co-authors also shows significant pre-trends, though these are accompanied by large standard errors which are likely due to limited statistical power (see Figure A.7, Panel (d)).

Table 1.2: Publications and collaborations

	(1) #pubs	(2) 1st author	(3) last author	(4) #co-authors	(5) #citations
Panel A: full sample					
HBP	0.140***	0.0322**	0.0430**	0.294***	0.281***
	(0.0255)	(0.0162)	(0.0178)	(0.0483)	(0.040)
LHS mean	4.1235	0.3528	1.3752	24.5028	271.553
Observations	9,585	9,585	9,585	9,585	9,585
#authors	639	639	639	639	639
Panel B: junior scholars (junior faculties, graduate students)					
HBP	0.307***	0.0864**	0.00211	0.630***	0.631***
	(0.0491)	(0.0351)	(0.0248)	(0.0947)	(0.0864)
LHS mean	2.0968	0.3723	0.3357	11.5662	108.9848
Observations	2,955	2,955	2,955	2,955	2,955
#authors	197	197	197	197	197
Panel C: female scholars					
HBP	0.112*	0.0170	0.0256	0.219*	0.348***
	(0.0577)	(0.0383)	(0.0339)	(0.107)	(0.0938)
LHS mean	2.9473	0.3571	0.9054	14.0677	140.8171
Observations	1,935	1,935	1,935	1,935	1,935
#authors	129	129	129	129	129

Notes: This table reports results estimating equation 1.1 on full sample (2008-2022). Outcome variables are log number of publications. One is added to all variables before taking the logarithm to include years without publications or publications without citations. #co-authors report the number of distinct co-authors per author-year.
Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1.3 reports additional results for quality and publication outlets, and topic direction metrics. Publications authored by individuals who were actively involved in the HBP at some point exhibit a higher probability of publishing in a top neuroscience journal. Notably, a majority of these distinguished neuro publications fall within the classifications of neurobiology and neurotechnology, representing two interdisciplinary fields of research (refer to Panel D). The engagement in the HBP markedly elevates the likelihood of publications being categorized as neurotechnology, as elucidated in Panel B and C. In summary, Table 1.3 suggest that the number of publications classified as neurotechnology, in particular, experiences a notable increase when authors participate in the HBP.[23]

For junior researchers, participation in the HBP increases the likelihood of publishing in the topics of neurotech (see Table A.4) and 1.4% of these are publications in leading neuroscientific journals. The results for the full sample regarding the impact on publication quality hold for females who were ever actively involved in the HBP, albeit with smaller statistical significance (see Table A.5).

[23]The related event studies confirm for most of the variables the parallel trend assumption (see Figures A.8-A.11). However, some variables exhibit significant pre-trends, i.e., the publications classified as top neuro journals.

Table 1.3: HBP participation and publications - full sample (author-year level)

	(1)	(2)	(3)	(4)
Panel A: Journal Quality (Top neuroscience/CS outlets)				
	Top Neuro	Top CS	CS A*	CS A
HBP	0.0266**	0.00746	-0.00279	0.0125*
	(0.0108)	(0.00820)	(0.00334)	(0.00686)
LHS mean	0.2678	0.1129	0.0371	0.0757
Panel B: Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0648***	0.140***	0.0784***	0.0522***
	(0.0220)	(0.0234)	(0.0158)	(0.0196)
LHS mean	2.6266	2.2341	0.8288	1.3723
Panel C: Topic Classification with the highest probability				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0378**	0.136***	0.0240**	0.0134
	(0.0192)	(0.0224)	(0.0110)	(0.0167)
LHS mean	1.7143	1.4548	0.4361	0.7322
Panel D: Topic Classification in Top Neuro Journals				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0268**	0.0152**	-0.000559	0.00683
	(0.0106)	(0.00665)	(0.00233)	(0.00675)
LHS mean	0.2476	0.0795	0.0099	0.0552
Observations	9,585	9,585	9,585	9,585
#authors	639	639	639	639

Notes: This table reports results estimating equation 1.1 on full sample (2008-2022). Outcome variables are log number of publications. One is added to all variables before taking the logarithm to include years without publications or publications without citations. Topic class with the highest probability contains 2,044 publications classified in multiple topics with equal maximum probabilities (2,037 cases with two topics, 7 cases with three topics), for which we count each topic as the max-likelihood topic (the results are qualitatively similar when analyzing disaggregated topic combinations). For those we count each topic equally as one. We also analyzed top CS outlets by topic classes, and none of the estimates are statistically significant nor economically meaningful (i.e., all estimates are very close to zero). Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Subsample Result by Highest Degree of Training

To what extent does the observed difference reflect HBP-active researchers trained in different areas contributing more to the productivity gain, as opposed to the platform provided encouraging collaboration? The former emphasizes the importance of training, and the latter stresses the importance of learning and knowledge-sharing. Ideally, we aim to capture details on all co-authors to gain insights into team composition and potential learning effects. However, since this is not feasible, we provide a proxy by utilizing hand-collected data on the highest degree of training of HBP-active authors. We included the highest-degree fields of training in the regression, using the field group of "business, social sciences, and humanities" as the baseline category.

In the full sample, researchers with the highest degree in natural and life sciences or bio-EECS publish significantly more, expand their collaboration network, and receive more citations (see Table 1.4, Panel A). Those trained in bio-EECS also benefit from the HBP by publishing more as the last author, compared to those trained in social sciences. Conditional on the highest-degree field, we observe much stronger results for junior researchers. This provides suggestive evidence that a combination of training and the expansion of the network at the beginning of one's career is most crucial for productivity gains. The only difference we find compared to the full sample is that junior researchers publish more as the first rather than as the last author (see Table A.6, Panel A). And this difference is to be expected. For female researchers, we do not observe any differences associated with the field of the highest degree.

When examining the topics in which researchers publish, we observe from Table 1.4 that those trained in bio-EECS have a higher probability of publishing in the topics of neurobio and neurotech, while those trained in natural sciences concentrate their publications on the topic of neurotech (see Table 1.4, Panel B). In the junior subsample, the results are similar but with larger coefficients (see Table A.6, Panel B). Female researchers trained in natural sciences appear to publish more in the topic of clinical research after joining the HBP.

Table 1.4: Publications and collaborations - field of training (author-year level)

	(1)	(2)	(3)	(4)	(5)
	#pubs	1st author	last author	#co-authors	#citations
Panel A: Publications and collaborations					
HBPxEECS	0.0623	0.0174	0.0222	0.111	0.260***
	(0.0407)	(0.0217)	(0.0265)	(0.0761)	(0.0852)
HBPxBio-EECS	0.215***	0.0388	0.152***	0.377***	0.272**
	(0.0510)	(0.0344)	(0.0463)	(0.0912)	(0.128)
HBPxNatural/Life Sc.	0.135***	0.00665	0.0252	0.317***	0.170***
	(0.0316)	(0.0186)	(0.0224)	(0.0584)	(0.0602)
Observations	9,585	9,585	9,585	9,585	9,585
#authors	639	639	639	639	639
	Neurobio	Neurotech	AI-Robotics	Clinical	
Panel B: Publication topics					
HBPxEECS	0.0187	0.0825**	0.102***	-0.00520	
	(0.0315)	(0.0354)	(0.0278)	(0.0283)	
HBPxBio-EECS	0.137***	0.185***	0.0948***	0.0930**	
	(0.0442)	(0.0432)	(0.0345)	(0.0394)	
HBPxNatural/Life Sc.	0.0801***	0.149***	0.0547***	0.0556**	
	(0.0255)	(0.0285)	(0.0187)	(0.0234)	
Observations	9,585	9,585	9,585	9,585	
#authors	639	639	639	639	

Notes: This table reports results estimating equation $y_{it} = \beta HBP_{it} * field_i + \delta_i + \delta_t + \varepsilon_{it}$. The highest degree field indicator for each individual was assigned based on the field in which they obtained their highest degree. The baseline category for this regression is the field of "business, social sciences, and humanities". One is added to all variables before taking the logarithm to include years without publications or publications without citations. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Geographical Diversity

Examining the coefficient plots that follow allows us to analyze the impact of the country of affiliation at the time of the first participation in the HBP on the dependent variables. As a baseline, we include all 17 countries for which the number of observations is below 400 publications.

Figure 1.5 illustrates a positive and significant correlation between the number of publications and all countries, compared to the base. However, being affiliated with a

German, Italian, Swedish, British, or Belgian research institute or university appears to exert a more substantial influence on the number of publications per author and year. The number of publications as the last author exhibits significant and positive coefficients for German, Swiss, Italian, and Israeli-based researchers. The results highlight regional disparities and specialized focuses on publication topics. For instance, participation in the HBP seems to have no discernible impact on researchers based in France, Spain, and Sweden regarding their publications in the AI-Robotics field. Instead, these researchers demonstrate a stronger inclination towards publishing within the neurotechnology field.

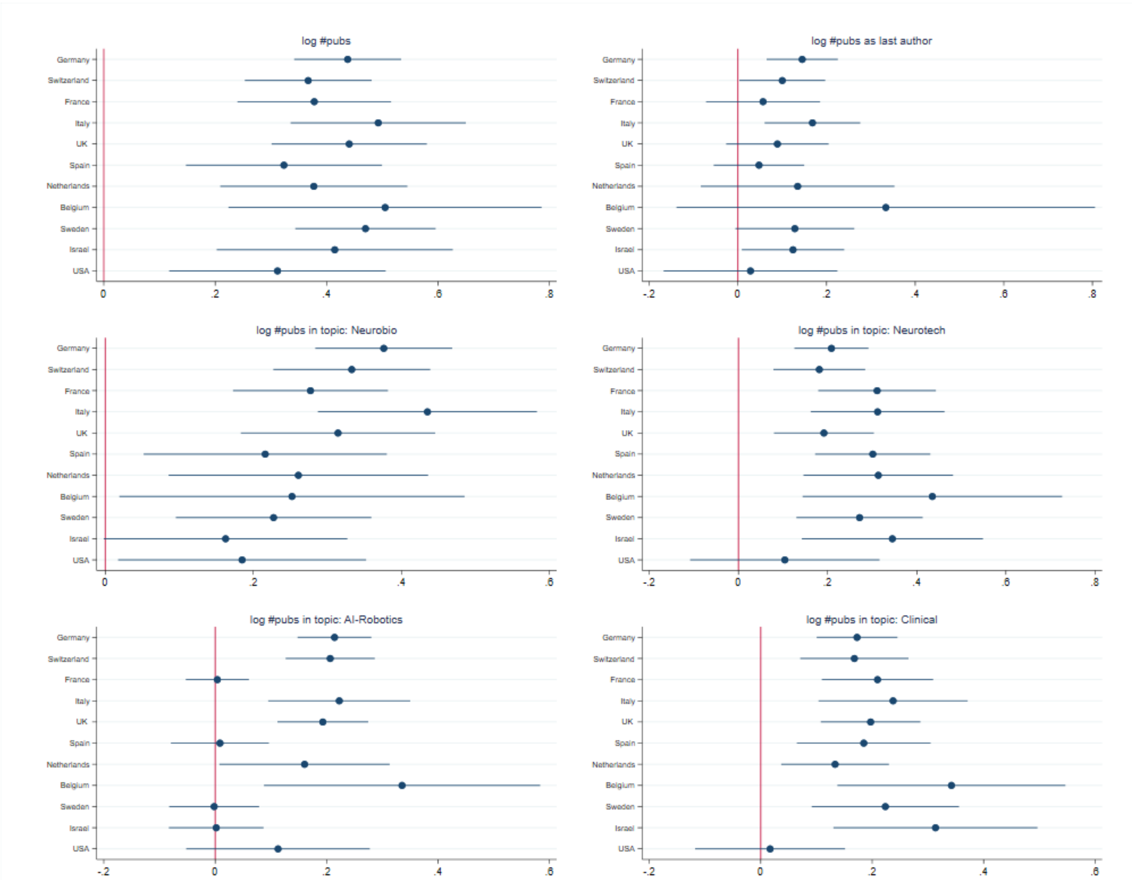


Figure 1.5: Regression results incl. country-specific interaction terms

Notes: This figure depicts the coefficients based on equation $y_{it} = \beta HBP_{it} * country_{it} + \delta_i + \delta_t + \varepsilon_{it}$. For each individual, we assigned the country indicator based on the country with which they were affiliated at their initial participation in HBP. The order of countries reflects the overall engagement of these countries. For #pubs, the base outcome is highly significant at the 1% level and equals 0.742. Similarly, for the number of publications as the last author, the base outcome is also statistically significant at the 1% level, amounting to 0.356. The baseline outcomes of the regressions of the four topic classes are statistically significant at the 1% level. Specifically, for publications in neurobiology, the coefficient is .560 for neurotechnology, it amounts to 0.462; for publications in AI-Robotics, it is 0.149; and for clinical topics, it equals 0.292. Confidence intervals are reported at the 95% level.

1.5 Conclusion

We analyze the intermediate effect of participating in the HBP, a 10-year megaproject funded by the EU to advance neuroscience with AI, on researchers' productivity and the direction of their research. To accomplish this, we constructed a novel dataset and employed two-way fixed effect model-based difference-in-differences analyses combined with NLP. Descriptive analyses reveal a consistent trend: over time, there is an increasing influx of junior faculty, graduate students, and female researchers into the HBP. Our empirical analyses further indicate a positive impact of HBP participation on researchers' productivity, as evidenced by an increase in the number of publications. Researchers also experience expanding collaborative networks and an uptick in citations received. Notably, researchers exhibit a heightened likelihood of publishing within the interdisciplinary field of neurotechnology. Also, our findings indicate that junior researchers derive the most substantial benefits from their involvement in the HBP.

While some event studies indicate empirical limitations, our findings provide valuable insights and should be interpreted as suggestive rather than strictly causal. To further strengthen our analysis, we are implementing a matching-based difference-in-differences approach. Specifically, we identify authors in Scopus that are similar to HBP-affiliated authors across key characteristics. For this purpose, we employ the *sosia* Python package developed by M. Rose and Baruffaldi (2020), which identifies matches based on year of first publication, publication outlets, citation count, publication volume, and prior co-authorship status. In addition, we are refining our topic classification approach by introducing an additional "Other" category to enhance the accuracy of our classification.

Although many breakthroughs require decades-long R&D, which are beyond the scope of our sample period, our study offers an intermediate-run perspective on the impact of a megaproject. Given the complicated nature of AI and brain sciences and the lack of existing classification, we cannot pinpoint the exact mechanism, but we have made the best attempt to trace the rate, direction, and quality of HBP-relevant publications. Future research would be valuable in unveiling the long-term, dynamic effects of the HBP.

Our investigation of the HBP has several practical policy implications. Although there have been different voices on whether "big sciences" projects like the HBP should be continued, many discussions are based on early-stage anecdotes. We provide intermediate-run analyses on the effectiveness of this modern megaproject using data and econometrics tools. Such intermediate-run evaluations are valuable as the long-run impact of the HBP is likely to unfold only after over a decade, but science policy needs to account for more immediate reactions in the scientific community. Overall, we find positive evidence in HBP for boosting research productivity and collaborative networks, particularly for junior scholars and interdisciplinary teams. However, further analysis is necessary to strengthen these findings.

2

Collaborating Neuroscience Online – The Case of the Human Brain Project Forum

2.1 Introduction

Neurological disorders are the leading cause of disability and the second leading cause of death worldwide, accounting for 9 million deaths (16.5% of total global deaths) and the loss of 276 million disability-adjusted life years in 2016 (Carroll, 2019). Most brain diseases have no cure, and many existing treatments are very expensive. Meanwhile, there is growing public and private investment in artificial intelligence (AI) for health care projects, with health care being the most invested sector by AI investors

*This chapter is based on joint work with Lucy Xiaolu Wang. A slightly adapted version of this chapter was published as Kreyer A-C, Wang LX (2022). Collaborating neuroscience online: The case of the Human Brain Project forum. PLoS ONE 17(12): e0278402.

(CB Insights, 2021). Although many life science areas, including neuroscience, are traditionally more laboratory- and experimentally-based, both COVID-19 and the massive global cost of neurological diseases heighten the need to harness digitalization in upstream health care markets. Since 2013, there have been different brain science initiatives launched in Europe, US, Israel, Japan, and China (Grillner et al., 2016) as well as an emerging international brain initiative (Adams et al., 2020), leading to a burgeoning "brain race" (Watson, 2013).

This paper studies the public-access online forum of one of the earliest brain initiatives, the Human Brain Project (HBP). Launched in 10/2013, the HBP is a flagship science initiative in the European Commission's Future and Emerging Technologies program and recipient of a 10-year, €1 billion grant. The HBP aims to advance brain research and improve treatments for neurological diseases by merging neuroscience with information and communication technology (ICT) including computational science, robotics, and artificial intelligence (Frégnac, 2017). As of 2021, a total of 179 institutions from over 20 countries had participated in the HBP. As well as making grants to research partners, the HBP allocates resources to build digital platforms, including a public-access online HBP Forum. As a major public discussion channel of the HBP, the utilization and collaborative problem-solving on the HBP Forum offers a valuable case study of institutional design to facilitate scientific collaboration.

We construct a novel dataset to examine whether and to what extent the HBP Forum is actively used, what factors are tied with richer online user interactions, and whether the HBP Forum offers an effective platform for problem-solving. We collected data from public sources to capture all user interactions and discussion content on the HBP Forum as well as characteristics of forum user profiles (e.g., demographics, institutions, scientific areas). We categorize the discussion threads based on the nature of topics, and further identify whether, when, and by whom a question has been solved. Data reveal that the HBP Forum is well-utilized and remains resilient during COVID-19, reflected in both the extensive margin of usage and intensive margin of user interactions. With the novel data constructed, this paper offers the first systematic empirical analyses of the utilization and performance of the HBP Forum.

We employ regression analyses to investigate what factors are associated with richer user interactions measured by the numbers of user replies per post within a quarter. We analyze covariates related to the content of discussed topics, the technical aspects, and the demographic and institutional profile of users who post the initial questions and users who reply. We further create a content-based measure of whether and when each question raised is solved effectively. We define a post as solved effectively if the asking user confirms the proposed solution; when direct confirmation is not available, we label the solution status based on the content and co-users' confirmation. We then utilize Cox proportional hazard models to analyze the time taken to solve a posted problem and covariates that accelerate problem-solving. We find that questions closely tied with HBP platforms and questions on programming issues with a higher share of explicit code in communications generate more discussions, especially when participating users are geographically more diverse. Questions posted on the Forum are solved faster when HBP administrators participate, and when code snippets are shared. Richness of interaction and likelihood of solution appear to be independent of participating users' HBP affiliation status.

This paper contributes to two strands in the literature. First, our paper contributes to studies about knowledge-sharing platforms by studying a large-scale, multinational scientific forum. As research specialization increases, knowledge-sharing and collaboration become increasingly essential for knowledge creation (Boudreau, 2010; Jones, 2021), which further promotes diversity in the process (Williams, 2013; McKiernan et al., 2016). Knowledge diffusion can be spurred by offline research institutions that further advance scientific discoveries (Furman & Stern, 2011; Murray et al., 2016). With the rise of remote work, online discussion forums spur knowledge sharing and creation with evolving communities and flat hierarchies (Faraj, Jarvenpaa, & Majchrzak, 2011; Grabher & Ibert, 2014; Randhawa et al., 2017). Prior studies have examined patterns and drivers of organizational sharing in sub-national and proprietary platforms (Haas, Criscuolo, & George, 2015; Gallus, Jung, & Lakhani, 2019; Gallus, Jung, & Lakhani, 2020; Di Stefano & Micheli, 2022). Recent qualitative studies suggest that life science platforms can function well by pooling resources from interdisciplinary areas (Morrison, 2017; Winickoff et al., 2021). We further offer a quantitative study of an online life science forum backed by a supranational organization.

Second, studies on digitalization in health care often focus on the adoption and utilization of health information technologies (HIT) among downstream users (e.g., health providers), and our study complements prior work by examining a research-oriented digital forum for upstream users (i.e., neuroscientists). Studies find HIT improve health outcomes mainly for complex conditions or specific populations (McCullough et al., 2010; McCullough, Parente, & Town, 2016; Freedman, Lin, & Prince, 2018), and that HIT can complement other programs in e.g. combating the opioid crisis (L. X. Wang, 2021). However, the increases in costs arising from HIT are also substantial

(Agha, 2014), although the cost burdens are less concerning in IT-intensive locations that provide complementary assets (Dranove et al., 2014). Given that brain-related diseases are mostly complex and effective treatments are rare, digitalization can provide new channels to spur global research collaborations for treatments. Some studies hypothesize that digital platforms can help advance neurosciences (Dario et al., 2005; Falotico et al., 2017), and our paper provides the first systematic analysis on how a digital forum is used by neuroscientists.

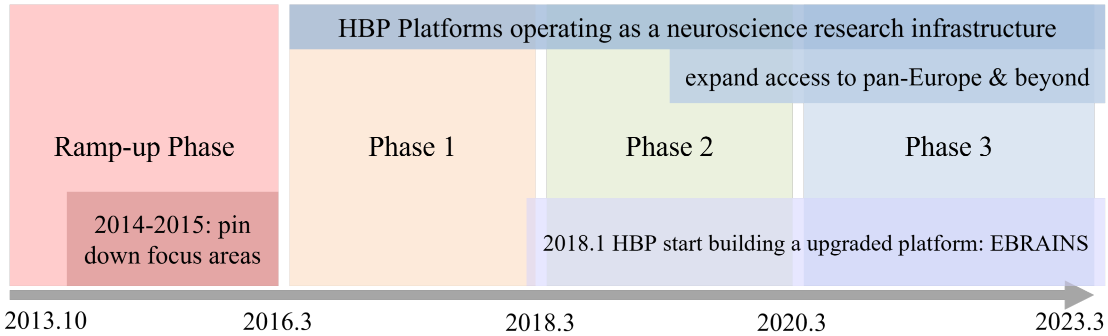
In addition, our study has policy implications for the design of online institutions for life sciences research. Both novel design and effective utilization of online institutions have become increasingly important given the disruption arising from COVID-19, which is reflected in the creation of the National Virtual Biotechnology Laboratory (NVBL, science.osti.gov/nvbl) by the US Department of Energy (DOE) as a consortium of DOE national laboratories. The pre-COVID-19 experience of the HBP offers a setting to understand how to proactively build institutional capacity that remains resilient during a disruptive period such as a pandemic. While it is difficult to evaluate the long-term impact of such projects, our analyses of contemporary performance on the forum can help inform policy regarding certain aspects of online institutional designs.

2.2 Background and Data Construction

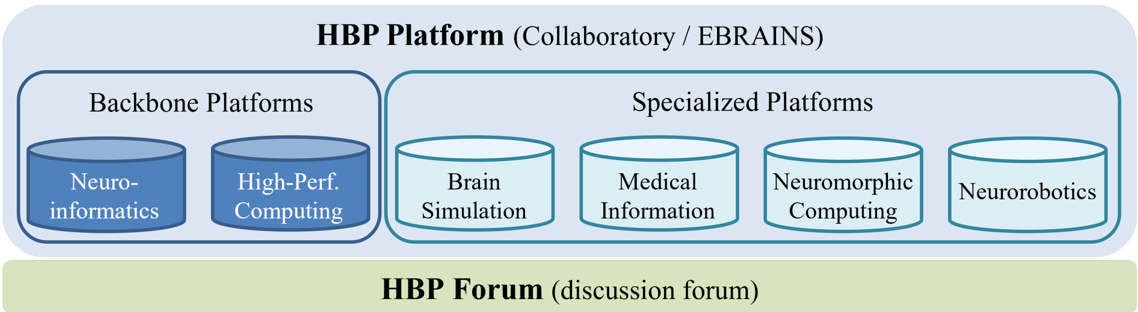
2.2.1 The Human Brain Project (HBP) and the HBP Forum

The HBP was launched in 2013, finished a ramp-up phase in 2016, and entered three special grant agreement phases in 2016, 2018, and 2020 (Fig 2.1(a)). The HBP includes 12 sub-project areas, including six generic topic areas (i.e., mouse brain organization, human brain organization, systems and cognitive neuroscience, theoretical neuroscience, management and coordination, ethics and society) and six platform-related sub-projects (i.e., neuroinformatics, brain simulation, high-performance analytics and computing, medical informatics, neuromorphic computing, and neurorobotics). Fig 2.1(b) shows the structure of the platforms. Access to the HBP platforms is granted to applicants from partner institutions; other users can request an account but are evaluated on a case-by-case basis. Besides general areas, research teams build project-specific private repositories, and such data are only accessible to related team members. From 01/2018, the HBP began building EBRAINS on the new

EU neuroscience supercomputing centers. The sub-platforms maintain the same focus and can perform better. The HBP platforms are hosted on a centralized access point: on HBP Collaboratory from 03/2016-09/2021, transitioning thereafter to EBRAINS as the new host, which remains current to date.



(a) HBP timeline and major events



(b) Structure of HBP Platforms and the Forum

Figure 2.1: The organization of the HBP platforms and the HBP Forum

Notes: Author's graph. Panel (a) depicts major milestones in the HBP timeline. Panel (b) provides an overview of the different AI-powered platforms built during the HBP and the broad relationship with the HBP Forum. While accessing certain areas within the HBP platforms requires additional authorization, the HBP Forum is publicly accessible and allows for decentralized knowledge sharing across the globe. Sources: https://www.humanbrainproject.eu/en/ and https://forum.humanbrainproject.eu/

The HBP Forum was launched in July 2015 as an integral connecting part of the HBP platform infrastructure. The Forum serves as a public discussion website about HBP-related topics, including questions on HBP-related activities in general, neuroscience progress, and programming challenges. Serving as the HBP's "Stack Overflow", topics raised and discussed in the Forum are public and can be read without registering an account, but only users with an account in the Forum can reply to or comment on topics raised in the Forum. Since the Forum is designed for public discussion, anyone interested in participating can create an account. Users do not have to be

HBP-affiliated to use the Forum, and users with an HBP account can use the same account for the Forum. Therefore, the HBP Forum is designed to facilitate informal collaboration and knowledge-sharing between researchers within and beyond the HBP community.

In the absence of detailed project-level data, user interactions in the public HBP Forum are the best available source to analyze HBP platform utilization and real outcomes of the Forum for the neuroscience research community. In addition, the HBP Forum retained the same structure and functionality during the transition of the HBP platform host starting in 09/2021, making the Forum a consistent measure of user activity. There were 534 posts on the Forum during 07/2015-03/2021., with 2,492 total replies and 550,175 total views. The collection and analysis method complied with the terms and conditions for the source of the data.

2.2.2 Post-Level HBP Forum Data Construction

We retrieved the full text of all posted threads available on the public HBP Forum between 07/2015 and 03/2021 (last accessed on 02/31/2021) and cross-checked the relationship database. We processed the rich text data and extracted information on the topics discussed in each post, the timestamp and content of each post and reply, and the number of total views for each post. We then merged in user-level data (see section 2.2.3) to the post-level data to capture the nuances on who interacts with whom, and how this differs by type of post. Based on the post-level data, we cleaned and organized the data at both the post-level including each complete thread of discussion following the initial question or topic posted and reply-level including each individual reply to a post. We have obtained approval and IRB exemption from the Data Protection Officer of the Max Planck Society and were confirmed that all data used are compliant with relevant sources and current regulation. During our sample period, a total of 534 posts were initiated by 208 of 283 total active users. We define active users as users who ever posted on the platforms, excluding registered users who never posted anything. On average, each post was viewed 1,030 times and received 3.7 replies. Across discussion topic categories, neurorobotics was the most popular, with 325 topics (60.8% of total) and 152 users (53.7% of total), followed by technical support (B.1 Fig).

To capture nuances in HBP Forum discussions, we categorized each Forum post in two independent ways. First, we obtained the content-based sub-categories tagged by the

Forum and grouped them into six major topic categories: neuromorphic, brain simulation/modeling, neurorobotics, technical support, organization, and others. Second, we analyzed all posts and manually categorized whether each post had a query to be solved, and if so, whether this query was solved by a HBP administrator/moderator or by users in the community. When multiple solutions were offered, we used the first solution timestamp to construct solving time. If an answer was provided first by a user and further clarified/confirmed by an administrator (2.5% of all posts), we classified it as administrator-solved. Posts that did not raise a question are classified as informational (i.e., were not question-oriented, and thus could not be solved). Two informational posts were re-categorized as questions as users, asked follow-up questions and were answered. Our results are robust to dropping or re-categorizing these two cases. Third, we recorded the timestamp when a query is solved. If an initial question was solved but inspired new questions and answers within the same post, we labelled such posts as multi-question posts and recorded information for each sub-question. In addition, we assigned two indicators for each post to capture if code snippets are included, and whether the topic being discussed is specifically related to HBP platforms (i.e., not a generic question).

To understand user diversity, equity, and inclusion on the platform, we further assigned indicators for whether a given post was created by a female user, whether it contained code snippets, whether the initial posting user was affiliated with an HBP partner institution, whether the initial post was created by a senior user, and the country where the user is based according to his/her main employer. Before aggregating the data to post- and post-quarter levels, we calculated the share of users who are female, share of replies that contain code, share of users who are more senior, and number of users affiliated with an HBP partner. In this way, we construct the dataset not only using the content of questions, but also to understand the type of questions being discussed and how diverse the Forum community is.

2.2.3 User-Level Data Construction

We further constructed a user-level database to capture details about who interacts with whom, and why. We combined information from multiple sources, including the HBP Forum, HBP PLUS (i.e., a user profile database maintained by the HBP for statistical reporting purposes for which users can opt in), HBP websites profiling key team members across project areas, and HBP YouTube channels with archived information on past team members. Specifically, we obtained the list of users, online usernames, real names, and institutional affiliations whenever publicly available for active users registered in the HBP Forum database. We used multiple matching algorithms to merge the Forum data with other sources based on details including full name, institutional affiliation, country of residence, contact details, gender, fields of experience, and highest level of education. We matched the users with the information we had collected from HBP promotional videos on YouTube and the internal HBP user database, HBP PLUS. The matching was based on the usernames and the real names using STATA's fuzzy matching algorithm "matchit" and the "merge" command. Each match was further verified manually. Where necessary, we supplemented these data with manual collection and disambiguation using Google Scholar searches, LinkedIn profiles, and institutional webpages. The dataset is anonymized and aggregated at the post-level.

2.2.4 Descriptive Statistics

In each complete calendar year during our sample period, there are on average 98 posts and 366 replies posted in the Forum (Table 2.1, Panel A). Per quarter, on average 23 posts are raised with a total of 85 replies, equivalent to about 3.7 replies per post (excluding the initial post). Most of the replies (i.e., about 90%) come within three months, and it takes 16 days on average to solve a question raised (Table 2.1, Panel B). Both numbers suggest active usage of the Forum within a fairly short time window. Furthermore, there is a fair amount of heterogeneity in post-level interactions. Some posts are discussed and solved quickly, while others generate lively follow-up questions and discussions that offer more learning and collaboration opportunities.

Table 2.1: Summary statistics

Panel A: Forum usage aggregate statistics					
Dependent variables	N	Mean	Std. Dev.	min	max
# posts per year (2016-2020)	5	98	54.09	56	190
# yearly replies (2016-2020)	5	366.2	219.02	178	717
# posts per year	7	76.29	57.67	21	190
# yearly replies	7	280.57	231.04	60	717
# post per quarter	23	23.22	14.4	6	57
# replies per quarter	23	85.39	62.54	16	231
# replies per quarter ($\leq$ 6 months)	23	79.78	60.68	16	230
# replies per quarter ($\leq$ 3 months)	23	77.04	58.03	13	215
# replies $\leq$ 3 months per posts	23	3.12	.9	1.85	4.67
Panel B: Post-level		Full Sample		Survival Sample	
		Freq.	%	Freq.	%
Total Posts		534	100	465	100
Information		69	12.92	0	0
Unsolved posts		84	15.73	84	18.06
Solved posts		381	71.35	381	81.94
Admin solved		312	58.43	312	67.10
User solved		69	12.92	69	14.84
Post w/ HBP platform relevance		456	85.39	443	95.27
Post w/o HBP platform relevance		78	14.61	22	4.73
Post w/ code		248	46.44	243	52.26
Post w/o code		286	53.56	220	47.74
Multi-thread posts		110	20.60	105	22.58
Average Solving Time		15.57 days		15.69 days	
Panel C: User statistics					
Users		260	100	233	100
Fully identified users		188	72.31	172	73.82
Non-admin users		211	81.15	188	80.69
Admin users		49	18.85	45	19.31
Male		211	86.48	194	88.18
Female		33	13.52	26	11.82
HBP affiliation		142	74.35	131	74.86
No HBP affiliation		49	25.65	44	25.14
User seniority		197	100	181	100
Undergraduate students		9	4.57	9	4.97
Graduate students		76	38.58	74	40.88
Junior researchers		28	14.21	24	13.26
Senior researchers		36	18.27	28	15.47
Senior software engineers		33	18.27	32	17.68
Non-academic		15	7.61	14	7.73

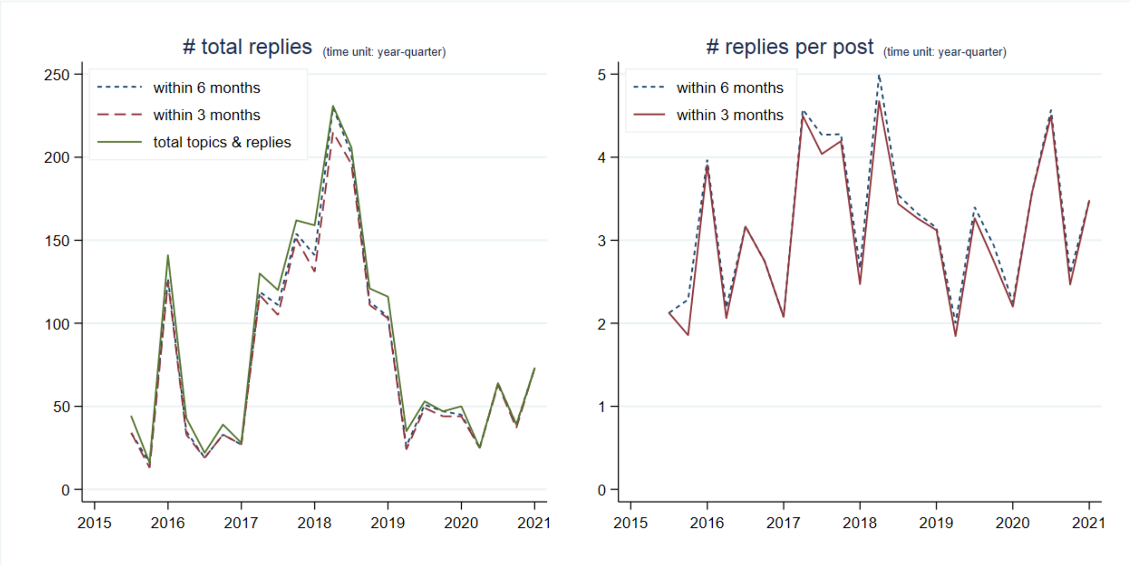
Notes: Panel (a) reports summary statistics of all discussion posts that appeared on the HBP Forum during the period 07/2015 and 03/2021. Panels (b) and (c) further disentangle characteristics for the full and focused samples at the post-level and user-level, respectively.

Out of the 534 total posts, 465 posts (87%) pose a question and the other 69 posts (13%) are informational posts (Table 2.1, Panel B). Informational posts address topics such as research diversity and explanations of different research tools. Of the 465 posts posing questions, 381 (82%) are solved, either by forum administrators (67%) or by users (15%). The share of unsolved posts on the Forum (18%) is smaller than that on Stack Overflow (29%) (Mondal et al., 2021). Considering all posts, most questions (85%) are directly related to the HBP platforms. Programming questions with code snippets account for 46% of all posts and 53% of question-oriented posts, including various levels of code intensity. Posts without code snippets often comprise organizational and application-related questions (B.2, B.3, B.4 Figs provide a few examples).

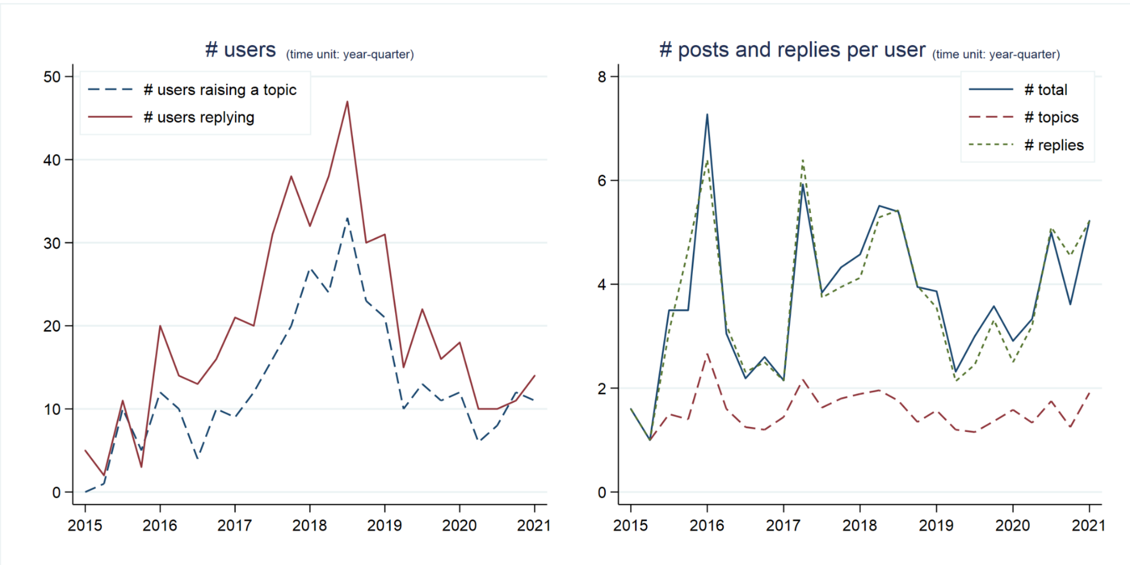
For the survival analyses, we focus on the initial question posted, excluding follow-up questions inspired within a given thread. We also exclude informational posts, as they pose no questions to be solved. These criteria result in our sample for the survival analyses with 465 observations. The survival sample has a similar distribution to the full sample regarding post-level and user-level characteristics, including share of posts with code, posts related to the HBP platforms, share of posted questions solved, and fully identified users (Table 2.1, Panel C).

There are 260 active users in our full sample, among which we fully identified 188 users (72%) regarding demographics, affiliation, and expertise. Unidentified users do not use their real names in their profile or in posts, and we do not have enough information to back up their identity. A large share of identified users is quite active and most of them are affiliated with an HBP partner institution. At the user level (Table 2.1, Panel C), most active users of the HBP Forum are male (86%), slightly lower than the male rate on Stack Overflow during 2015-2020 (>90%).The share of female users (14%) is higher than that in Stack Overflow (i.e., 6% in Europe), but lower than the share of female neuroscientists worldwide (25%) (source: https://insights.stackoverflow.com/survey/2020) (Metitieri & Mele, 2020). About 81% of Forum users are voluntary users (i.e., not administrators). 74% of users are affiliated with an HBP partner institution and are members (but not leaders) of HBP sub-projects. 92% of the active users are students or academic employees (e.g., researchers and software engineers). Non-academic users comprise computer scientists employed in private companies or self-employed entrepreneurs. Seniority level is coded for all active users. Nine users changed in seniority level during our sample period and are coded accordingly. Most Forum users are graduate students (39%). Junior researchers (post-doctoral scholars and assistant professors) make up 14% of users, while senior researchers (associate and full professors) make up 18%. Finally, 18% of active users are senior software engineers employed by a research institution. We group senior researchers and senior software engineers together in our analyses.

Geographically, active users are based in 27 countries, with the top six countries accounting for about 76% of the users: Germany (30%), Switzerland (21%), United Kingdom (8%), Italy (7%), United States (6%), and France (3%). On average, we observe users participated from 1.7 countries per post. The geographic diversity peaks in 2018 with 1.94 countries per post and is then stable throughout the observed window of time.



(a) Trends in Forum usage and per post user interaction



(b) Trends in active users on the Forum and user activities

Figure 2.2: HBP Forum usage: Trends in posts, replies, and users

Notes: Panel (a) depicts aggregate utilization of the Forum at a quarterly-level to the first quarter of 2021. Panel (b) offers a complementary view of the data from users' perspective and shows aggregate user activity over time, both at the extensive margin (i.e., # active users) and at the intensive margin (i.e., # posts and replies per active user) per quarter.

Fig 2.2 depicts the trends in Forum usage, measured by aggregate post-level interactions, total active user engagement, and corresponding disaggregated numbers per post. Forum usage peaked in 2018 with the largest volume of post-level interactions and active users; this is consistent with a "deadline effect" coinciding with the conclusion of HBP Phase One, when users may have been finalizing results for end-of-phase reports. The numbers stabilize to around 50 new posts and 15 active users per quarter in 2019. The number of replies stabilizes to 3-4 per post, and each active user replies 4-5 times per quarter. Despite a drop in usage during 2020, user interaction rose substantially in late 2020 and early 2021.

2.3 Empirical Analyses and Results

2.3.1 What Drives Richer Online User Interactions?

To examine the factors that drive richer online user interactions, we aggregated the data to the post-quarter level. The data structure is not a panel as we often have only one observation over time per post. The data set contains 599 observations of 534 posts, where 58 posts containing replies from more than a quarter (i.e., 3 months) later contribute to the number of observations exceeding posts. We perform regression analysis at the post-quarter level using the following equation:

$$y_{it} = \delta_t + \eta HBP_{platform_i} + \gamma X_{it} + \epsilon_{it} \tag{2.1}$$

Here y_{it}, is the number of replies for post i in quarter t. $HBP_{platform_i}$ indicates whether post i is related to the HBP platforms. X_{it} is a vector of post and user-level characteristics including geographic and gender diversity, posts including code snippets, and HBP partnership affiliation; we use shares instead of levels to account for the standardized user composition. δ_t denotes year-fixed effects. Heteroskedasticity robust standard errors are reported. Given that most post-level interactions happen within the first three months, we do not have within-post over-time variation to allow for post-level fixed effects. All variables are aggregated to the post-quarter level.

Table 2.2 reports the results from post-level analyses. Column (1) includes covariates related to post composition that vary at post-level. Column (2) comprises initial post characteristics. Column (3) combines the two sets of covariates. Column (4) further includes topic category fixed effects to account for differences in the underlying post topics. Throughout the specifications in columns (1)-(3) user interactions are significantly higher for posts related to the HBP platforms, and programming posts with code snippets. Posts with a higher share of code-embedded replies have on average more replies. In particular, the inclusion of code snippets in the initial question post is a strong predictor of more follow-up interactions, with the estimates positive and statistically significant in all specifications. This is consistent with prior studies on the importance of including code snippets in the initial question to receive more attention of fellow users, and thus a faster and more targeted solution (Asaduzzaman et al., 2013; Calefato, Lanubile, & Novielli, 2018).

Table 2.2: HBP Forum utilization analyses at the post-quarter level

Dependent variable:	# replies per post-quarter			
	(1)	(2)	(3)	(4)
HBP platforms related	1.085***	1.843***	0.872***	
	(0.240)	(0.310)	(0.260)	
% replies w/ programming code	0.910***		0.635*	0.599*
	(0.336)		(0.340)	(0.348)
% female users replying	0.983		1.171	0.984
	(0.709)		(0.753)	(0.752)
% replying users w/ HBP affil.	0.579*		0.516	0.542*
	(0.315)		(0.327)	(0.328)
% senior users replying	0.374		0.424	0.459
	(0.349)		(0.340)	(0.350)
% admin users replying	-0.674**		-0.601*	-0.691**
	(0.334)		(0.339)	(0.343)
# countries	1.965***		2.056***	2.022***
	(0.288)		(0.333)	(0.326)
Initial post w/ programming code		0.766***	0.633**	0.600**
		(0.274)	(0.254)	(0.254)
Initial post by user w/ HBP affil.		-0.304	-0.0118	-0.0693
		(0.237)	(0.235)	(0.235)
Initial post by female		0.350	-0.134	-0.187
		(0.326)	(0.301)	(0.308)
Initial post by senior		-0.323	-0.0931	-0.0552
		(0.414)	(0.392)	(0.398)
Category Brain Sim/Model				-1.841
				(1.163)
Category Neurorobotics				-0.163
				(1.075)
Category Tech Support				-0.278
				(1.102)
Category Organization				-1.161
				(1.086)
Category Others				-1.190
				(1.080)
LHS mean	3.86	3.86	3.86	3.86
Year Fixed Effects	YES	YES	YES	YES
Observations	599	599	599	599

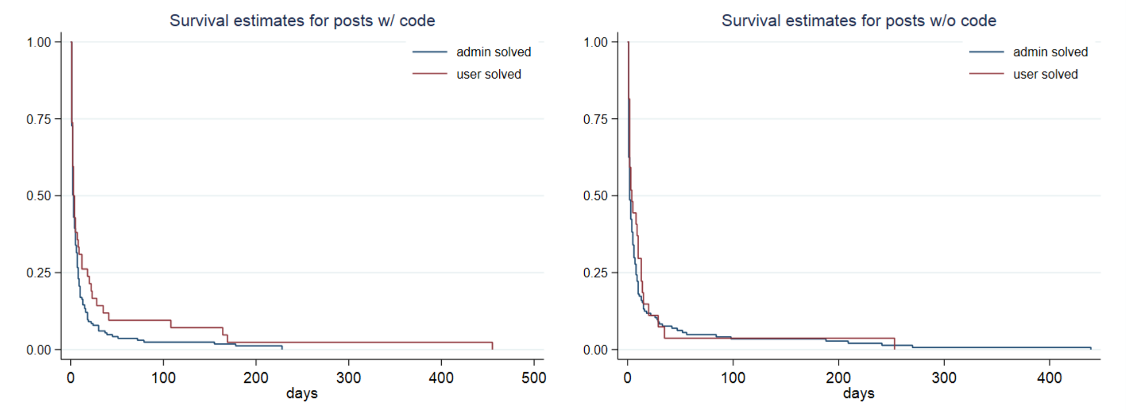
Notes: Each cell reports the coefficient of interest from a separate regression. Each unit is aggregated to the post-quarter level. We capture post-quarter user interaction by regressing the number of replies a post received on various covariates. Robust standard errors in parentheses. Robust p-values: *** p<0.01, ** p<0.05, * p<0.1.

Contrary to prior studies (Ford, Harkins, & Parnin, 2017; May, Wachs, & Hannák, 2019), we do not find statistically significant differences in user interaction patterns related to the gender of the user who raises a question. Similarly, whether a question is asked by a user with advanced access to the HBP platform or by a more senior user does not significantly alter the interaction on the Forum. These results likely partly reflect the usage of gender-neutral usernames and the lack of a direct information tag on user's HBP affiliation status or experience on their profile and posts. This design feature of the Forum is worth further investigation in future studies of online institution building. In contrast, Forum administrators are clearly marked on the profiles and a higher share of replies from administrators (i.e., more institutional support provided) associates with lower intensity of voluntary interactions. Further, greater geographic diversity in participating users is associated with significantly more replies to a post within a quarter, and this effect is stronger when controlling for both the post-level and the initial post characteristics in column (3).

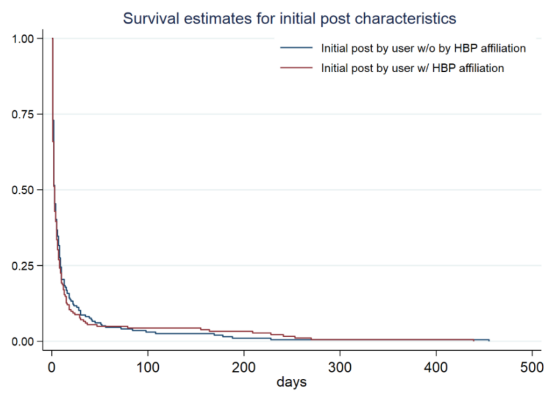
To account for differences between the underlying content discussed in each post, we further control for fixed effects at the content category level (column (4)). In this more demanding specification, the $HBP_{platform_i}$ variable is no longer present due to collinearity with the category fixed effects. After controlling for this content-level measure, more diverse country-level user participation (i.e., at the extensive margin) remains statistically significant and positively associated with active user interactions (i.e., at the intensive margin). Overall, we observe similar patterns in the main estimates. There is a higher level of user interactions for posts asking questions related to the HBP platforms, those including code snippets in the initial posts and in the replies, and those with geographically-diverse users.

2.3.2 What Factors Make Problem Solving Faster on the HBP Forum?

To further examine the factors that accelerate research problem solving in the HBP Forum, we start with Kaplan-Meier non-parametric survival estimates to visually represent differences in problem solving associated with a few key factors. The Kaplan-Meier survival estimates represent the probability of an event occurring after a certain point in time. Fig 2.3(a) suggests that HBP Forum administrators solve programming-related questions faster than users. However, the difference in problem solving time for non-programming related topics is minimal between administrators and users. The sharp drop in the share of unsolved posts to 25% within the first 25 days suggests that 75% of questions posted, programming-related or not, are solved by then. Fig 2.3(b) suggests that solving time does not differ much between posts initiated by users affiliated with an HBP partner institution or not, which reflects our results regarding the usage analysis in Section 2.3.1.



(a) Time to solution for questions with code snippets (left) and without (right) by user type



(b) Time to solution by initial post user's HBP affiliation status

Figure 2.3: Kaplan-Meier non-parametric survival analyses

Notes: This figure reports a set of main graphical results from the survival analyses to disentangle what factors are most relevant for effective problem solving in online platforms. Additional graphs on other factors in the hazard and time-to-event analyses are available upon request.

To systematically understand factors associated with online problem solving, we use a Cox proportional hazard model to analyze solving time (in days) of the question raised in post i. For the general model, we apply the following functional form:

$$h(d_i, s_i, X_i) = h_0(t) exp^{\lambda(\alpha d_i + \beta X_i)} \tag{2.2}$$

where $h_0(t)$ indicates the baseline hazard. s_i indicates whether the questions raised in post i is solved by the end of our sample period. d_i is the number of days between when question i is raised and solved on the HBP Forum. X_i contains an indicator of the relevance of a topic to the HBP platforms, as well as the set of time-invariant covariates about the post-level characteristics that we used in the previous analyses (section 2.3.1). The coefficient estimates represent hazard ratios, and a ratio greater (smaller) than one indicates a positive (negative) relationship with the probability of a post being solved (i.e., with s_i equals one). As a robustness check, we also conducted a time-to-event analysis that allows for a more flexible functional form of the baseline hazard by transforming the survival function using natural cubic splines as a link function (Royston & Parmar, 2002).

Columns (1)-(4) of Table 2.3 report the results of the Cox proportional hazard model. Column (1) includes the post profile covariates, column (2) the initial post characteristics, and column (3) a combination of both. Column (4) replaces the capture-all HBP Forum indicator with content-based topic category fixed effects. Column (5) shows the results of the time-to-event analysis using the same covariates as in Column (3). The results in columns (1)-(3) suggest that posts related to HBP platforms have a higher probability of being solved compared to non-platform related posts. The coefficient magnitudes are similar across specifications, and indicate that solving probability is on average twice as high for platform-related posts (e.g., Column (3): $1.956 - 1 = 0.956$). The estimate is similar and larger in magnitude in the time-to-event analysis.

When administrators participate, the probability that a question posted on the Forum is solved is 28-33% (e.g., Column (3) $1.302 - 1 = 0.302$) higher in the proportional Cox hazard model, and 37% higher in the time-to-event analysis. Combined with previous results that higher level of administrator participation associates with lower voluntary user interaction, our findings suggest that support from administrators solve questions effectively faster. In contrast to the wholly user-driven Stack Overflow experience (Asaduzzaman et al., 2013), institutional support by the administrators appears to be very beneficial for Forum users. Questions on programming issues with code and related to HBP platforms are also solved faster. In column (4), the estimates on the categories “Brain Simulation and Modelling” and “Neurorobotics” re-enforce the results that platform-related posts and application questions are solved faster.

Consistent with our usage analysis, neither the gender of the user raising the question nor that of those who reply is statistical-significantly associated with lower solving time. We find no evidence for the differences in online knowledge sharing between women and men as observed in other studies (Zolduoarrati & Licorish, 2021). The solving probability is also not affected if a higher share of senior researchers participate.

Table 2.3: Cox proportional hazard analyses and time-to-event analysis

	Cox regression hazard ratios				Time-to-event hazard ratios
	(1)	(2)	(3)	(4)	(5)
HBP platforms related	1.975**	1.837*	1.956**		2.168*
	(0.617)	(0.590)	(0.618)		(0.862)
Solved by admin	1.293*	1.283*	1.302*	1.331**	1.366**
	(0.177)	(0.165)	(0.180)	(0.186)	(0.199)
% replies w/ programming code	1.339**		1.265*	1.264*	1.398*
	(0.160)		(0.153)	(0.156)	(0.246)
% female users	0.949		0.830	0.928	0.868
	(0.335)		(0.359)	(0.422)	(0.506)
% users w/ HBP affil.	1.331		1.348	1.249	1.386
	(0.238)		(0.292)	(0.279)	(0.346)
% senior users per post	1.153		1.153	1.218	1.227
	(0.174)		(0.187)	(0.201)	(0.226)
# countries	0.792***		0.794**	0.768***	0.737***
	(0.0696)		(0.0720)	(0.0744)	(0.0666)
Initial post w/ programming code		1.191*	1.102	1.128	1.127
		(0.116)	(0.111)	(0.114)	(0.133)
Initial post by user w/ HBP affil.		1.127	0.973	0.979	0.970
		(0.105)	(0.112)	(0.113)	(0.129)
Initial post by female		1.125	1.110	1.169	1.087
		(0.149)	(0.189)	(0.214)	(0.235)
Initial post by senior		1.080	0.990	1.105	1.003
		(0.182)	(0.164)	(0.171)	(0.182)
Category Brain Sim/Model				2.238**	
				(0.851)	
Category Neurorobotics				1.922**	
				(0.574)	
Category Tech Support				1.519	
				(0.502)	
Category Organization				1.331	
				(0.523)	
Category Others				2.299*	
				(1.089)	
Year Fixed Effects	YES	YES	YES	YES	YES
Observations	381	381	381	381	381

Notes: We report the results from estimating a Cox proportional hazard model and a time-to-event analysis for the time in days to a first solution of the post. We report hazard ratios. Hazard ratios greater (smaller) than one indicate a positive (negative) relationship with the probability of solving the post. The percentages reported are obtained by subtracting 1 from the hazard ratio coefficients. Robust standard errors in parentheses. Robust p-values: *** p<0.01, ** p<0.05, * p<0.1.

In the previous analysis, larger geographical user diversity is associated with richer user interaction. The results in Table 3, however, suggest a 21-23% lower probability of the post being solved (e.g., Column (3): 0.794 – 1 = -0.206) as the number of user-participating countries increase. The results are similar using the Cox proportional hazard model and time-to-event analysis. This result could imply that more time is needed in the accumulation and harmonization of knowledge from various sources, and that questions attracting more geographically diverse user participation may be more complex. In the absence of a clear metric for question complexity, we leave a fuller investigation of this for future research.

2.3.3 Additional Discussions

In addition to providing an online community for neuroscientists, the HBP Forum also promotes open-source programming for back-end activities. While the HBP Forum provides direct user support on all HBP-related topics, some questions about the neurorobotics platform require further technical support and may be forwarded to an administrator-only repository – the neurorobotics Jira BitBucket, where more substantial issues and bugs are tracked and resolved. This repository is maintained by 14 out of the 50 HBP Forum administrators in our sample. In our sample, 18 out of 534 posts were sent to Jira BitBucket. In a robustness test, we control for whether a post was forwarded to BitBucket in the regressions to proxy for complexity. Our results remain unchanged, as the set of questions forwarded outside the Forum is small. In addition, most user interactions are voluntary instead of directed communications between users with prior ties: in only 25 posts we observe direct user tagging for additional support, and only 14 users were ever tagged directly (among which, nine were administrators).

Our result on female participation is in line with results of gender studies in online platform collaboration. Female users are in general under-represented in online technology-related Q&A platforms (Zolduoarrati & Licorish, 2021). Studies show that female users are more inclined to participate in posts if there are already other female users replying to a question (Ford, Harkins, & Parnin, 2017; May, Wachs, & Hannák, 2019). Similar to Stack Overflow, the gender and affiliation of active users are not revealed in the user profile and as such, users are left to guess gender (Ford, Harkins, & Parnin, 2017; Zolduoarrati & Licorish, 2021). This gender-neutral feature of platform design is worth further investigation and can potentially reduce gender-related biases.

2.4 Conclusion

We study the utilization and effectiveness of the HBP Forum, a platform that offers neuroscience researchers a public space to discuss issues and raise queries broadly related to the Human Brain Project. We construct a novel and comprehensive dataset to capture the usage of the HBP Forum, including the range of topics being discussed, plus whether, how quickly, and by whom research problems are solved. We find that the HBP Forum is actively used and remains active during COVID-19. On average, each forum post is discussed at a similar interaction level as those on general platforms (e.g., Stack Overflow), and Forum usage recovered fairly quickly after an initial drop of activity during COVID-19 first wave.

Our results provide encouraging evidence that the online community built through the HBP has generated active participation among users from different institutions and with different educational levels who may not have otherwise connected. The institutional support provided by forum administrators appears helpful in supporting the collaborative progress of the online neuroscience community, which may be especially important at the current time, when physical distance to peers is increased. Our analyses offer a first glimpse into the facets of a particular online collaboration infrastructure within a large, long-term life science project.

3

From Tinkering to Inventing – FabLabs as Catalysts of Innovations

3.1 Introduction

Access to relevant knowledge and tools and the opportunity to apply it are critical prerequisites for innovation (Breschi, 2001). While digitization and technological advancements have facilitated online knowledge sharing and acquisition, not all skills can be effectively learned through digital means alone (Davenport & Prusak, 1998; Faraj et al., 2016; Rayna & Striukova, 2021). Creating new knowledge and driving innovation require the exchange and combination of both tacit and explicit knowledge across disciplines and individuals (Nonaka, Takeuchi, & Umemoto, 1996; von Krogh, 1998; Popadiuk & Choo, 2006). In particular, tacit knowledge often requires direct interpersonal exchange and hands-on experience (Gertler, 2003; Breschi & Lissoni,

*For this chapter, I used ChatGPT and DeepL-Write to refine grammar and wording.

2009). However, as the global knowledge frontier expands, reaching and advancing it has become increasingly challenging (Jones, 2009).

Certain developing and complex technologies necessitate physical access to specialized tools and collaborative environments (Ludwig et al., 2014). A prime example is Additive Manufacturing (AM), or 3D printing, where objects are designed digitally and physically fabricated (Gebhardt & Hötter, 2016). Mastering AM thus requires both software and hardware expertise and integrates the digital and physical worlds (Ludwig et al., 2014). By enabling rapid prototyping, customization, and flexible product development, AM facilitates innovation, making it particularly relevant for entrepreneurs and early-stage innovators (Berman, 2012; Weller, Kleer, & Piller, 2015; Mortara & Parisot, 2018). While declining costs have made AM more accessible, barriers remain both in terms of affordability and the technical skills required for effective use (Fleischmann, Hielscher, & Merritt, 2016; European Patent Office, 2023).

Digital Fabrication Laboratories (FabLabs) address these challenges by providing open-access community workshops equipped with AM technologies and other digital fabrication tools. By lowering entry barriers, they foster experimentation, collaboration, and both tacit and explicit knowledge exchange, and thus have the potential to act as catalysts for user-driven innovation (Mikhak et al., 2002; Maravilhas & Martins, 2019; Oppedisano, 2024). Unlike incubators and accelerators, which support the scaling of existing ventures (Assenova, 2020; Hallen, Cohen, & Bingham, 2020), FabLabs provide an open environment focused on early-stage ideation. Their inclusive and open-access model allows individuals to transform ideas into tangible products. By broadening access to cutting-edge technologies and fostering knowledge exchange and skill acquisition, FabLabs can unlock latent inventive potential and foster grassroots innovation (Gershenfeld, 2007). While similar to other types of community workshops, e.g. Makerspaces, FabLabs are distinguished by their standardized global structure, emphasis on openness, exclusive focus on digital fabrication, and educational mission (Rayna & Striukova, 2021; Oppedisano, 2024).

This paper investigates the impact of FabLabs on regional innovation in Germany measured by patents.It examines whether FabLabs stimulate local invention activity and how their effects vary across different regional contexts. Specifically, it assesses whether FabLabs increase local invention activity and facilitate first-time inventors, whether university-integrated FabLabs have a stronger effect due to exposure to cutting-edge research, and whether impacts differ between urban and rural areas. In addition, it explores how FabLabs interact with pre-existing Makerspaces and

Hackerspaces and whether their presence enhances or mitigates FabLabs' effectiveness. These questions are particularly relevant for Germany, historically renowned for its engineering-driven innovation but facing declining innovation activity in recent years (Klöckner & Specht, 2023; Koppel, 2023).

To better understand the organizational structures and functions of German FabLabs FabLabs and inform my empirical strategy, two online surveys were conducted. The first targeted FabLab managers, gathering insights on founding motivations, organizational structure, funding, integration into local innovation ecosystems, and services offered. The second survey, aimed at FabLab users, explored their engagement, usage patterns, outputs, and perceived benefits.

To identify FabLabs' causal impact on local patenting activity, this study exploits the staggered rollout across German cities since 2009 and analyzes the effect at two regional levels: OECD Functional Urban Areas (FUA) and NUTS-3 regions.[1] The empirical strategy exploits variation in the timing of FabLab establishment, comparing regions that have a FabLab to those that do not have one yet or never will within the sample period. The analysis is based on a novel dataset of patent and utility model applications with at least one Germany-based inventor from 2003–2019, aggregated at the FUA-year and NUTS3-year levels. The impact of FabLabs on the extensive margin of patenting is examined through patent and utility model application counts, patent applications by first-time inventors, patent applications by individuals and by organizations, and AM-related patent applications. The effects on the intensive margin are assessed using forward citations and the number of technical areas cited per patent, capturing knowledge spillovers and technological breadth of inventions.

The study reveals that FabLabs do not significantly increase overall patenting activity but rather exert localized effects, particularly by fostering AM-related patents and enabling first-time innovators. University-integrated FabLabs are especially effective in promoting interdisciplinary and AM-related patents, whereas utility models tend to decline. FabLabs' impact depends on regional preconditions. In urban regions, they amplify inventive activity by fostering especially AM-related patents and enabling new (accidental) inventors, with access to advanced tools likely being the key mechanism. In rural areas, they can act as catalysts of innovation, albeit with delayed and weaker impacts compared to urban areas. In rural regions, both access to tools and

[1]FUAs, which include both an urban core and its commuting zone, better capture the functional economic area of a city, thus accounting for a broader pool of potential inventors and users (Dijkstra, Poelman, & Veneri, 2019). Meanwhile, NUTS-3 regions offer a finer-grained view of localized innovation dynamics, enabling differentiation between rural and urban areas.

collaborative environment may play a role. The presence of complementary community workshops further enhance FabLabs' impact, particularly in urban areas, where Makerspaces and Hackerspaces create synergy effects and FabLabs address previously unmet demand for advanced tools. In contrast, in rural regions, FabLabs are most effective when filling gaps in innovation infrastructure by being the first community workshop, although adaptation periods tend to be longer. Overall, FabLabs foster regional innovation, but their impact remains highly contingent on the local conditions and and complementary infrastructures.

This paper contributes to four key strands of literature that collectively explore how institutions that facilitate access to technology influence innovation. Specifically, it examines the role of FabLabs in providing access to advanced tools, fostering knowledge exchange, enabling technological experimentation, and stimulating innovation. While FabLabs are often recognized as hubs of grassroots innovation, empirical evidence on their broader economic impact remains limited. This study addresses this gap by assessing how FabLabs influence local innovation dynamics.

The first strand of literature investigates the impact of FabLabs on innovative activity, yet much of this research remains qualitative or based on small-scale case studies. Savastano et al. (2017) highlight that nearly half of FabLab studies focus on education, emphasizing their role in schools and universities Other studies describe the evolution of FabLabs and successful examples of FabLabs in developing countries (Diez, 2012; Stacey, 2014). Research on knowledge sharing and skill development emphasizes FabLabs' role in fostering collaboration (Schrape, 2020) and propose mechanisms to enhance knowledge exchange (Troxler & Wolf, 2010; Troxler & Schweikert, 2016). A recent global survey finds that FabLabs prioritizing digital manufacturing research and industry collaboration generate more innovation-linked projects than those focused on education (Garcia-Ruiz, Lena-Acebo, & Rocha Blanco, 2023). The entrepreneurial impact of FabLabs has been analyzed qualitatively by Mortara and Parisot (2018), and quantitative evidence for related community workshops, such as Makerspaces (Bao, 2025) and Hackerspaces (Cuntz & Peuckert, 2023), suggests positive local entrepreneurial effects. Despite these insights, there remains a critical need for comprehensive, quantitative analyses of FabLabs' broader economic impact.

This paper contributes to the literature on institutions that facilitate access to specialized technological knowledge and their impact on innovation. Existing literature largely analyzes how access to codified scientific knowledge, such as patents, publications, or scientific data, affects innovation. Reducing barriers to codified knowledge, whether through patent disclosure (Graham & Hedge, 2015; Gross, 2019; Furman, Nagler, & Watzinger, 2021), lower patent licensing costs (L. X. Wang, 2022), or open access to patents, scientific publications, and other research outputs (Furman & Stern, 2011; Williams, 2013; McKiernan et al., 2016), has been shown to accelerate knowledge diffusion and increase patenting activity. However, codified knowledge often remains inaccessible to non-experts, limiting its broader impact. Many potential innovators could benefit from more accessible, hands-on knowledge (Bell et al., 2019). Local access to general, practical knowledge fosters patenting activity among both non-experts and inventors (Berkes & Nencka, 2020), while open access to high-tech tools stimulates innovation among diverse actors in fields like pharmacology (Murray et al., 2016) . Despite these insights, empirical research on institutions offering physical access to high-tech technologies and knowledge remains limited. This paper aims to close this gap by examining how FabLabs—akin to public libraries but for fabrication technologies—impact local innovation.

The third contribution relates to the science of science literature by examining how local institutions shape knowledge exchange and innovation. Proximity enhances tacit knowledge exchange as this requires face-to-face interactions and a willingness to share insights (Breschi, 2001; Breschi, Lissoni, & Montobbio, 2005). Universities' R&D activities generate spillovers to local firms and laboratories (Jaffe, 1989; Agrawal & Cockburn, 2002; Alcacer & Chung, 2007; Furman & MacGarvie, 2007), and their co-location and interaction with innovative industries foster dynamic innovation ecosystems (Asheim & Gertler, 2009; Audretsch, Hülsbeck, & Lehmann, 2012; Granstrand & Holgersson, 2020). Beyond inter-organizational spillovers, local collaborative spaces enable intra-organizational knowledge exchange that drive knowledge diffusion and catalyze disruptive innovations (Singh & Fleming, 2010; Catalini, 2018; Roche, Oettl, & Catalini, 2024). Innovation depends on collaboration among actors with complementary capabilities, as no single entity possesses all necessary knowledge (Harhoff, Henkel, & von Hippel, 2003). Networks of diverse individuals foster the conditions for participants to become innovators themselves (Tortoriello, McEvily, & Krackhardt, 2015), and even brief, informal exchanges enhance productivity and spur future innovation (Berends et al., 2006; Sandvik et al., 2020; Andrews, 2023; Baruffaldi & Poege, 2025). This paper extends this literature by examining FabLabs as local institutions that foster knowledge exchange and collaboration, particularly in high-tech technologies.

Fourth, this paper contributes to the literature on open and user innovation. The role of users as innovators in open, collaborative settings has been widely recognized as an alternative to traditional producer-driven innovation models (Harhoff, Henkel, & von Hippel, 2003; Baldwin & von Hippel, 2011). Numerous studies have demonstrated successful product innovations originating from users (Baldwin & von Hippel, 2011; Gambardella, Raasch, & von Hippel, 2017; Boutillier et al., 2020). FabLabs, as open-access, collaborative spaces, are theorized to promote user-driven innovation, knowledge exchange, and grassroots creativity (Boutillier et al., 2020; Oppedisano, 2024). However, empirical evidence on the extent and nature of user innovation within FabLabs remains scarce. While qualitative research suggests that FabLabs democratize innovation and facilitate grassroots creativity, there is little systematic, large-scale analysis of their impact on innovation outcomes. This study seeks to address that gap by providing empirical evidence on how FabLabs enable technological development.

The remainder of this paper is structured as follows: Section 3.2 provides theoretical considerations and background on FabLabs. Section 3.3 describes the survey and its results, and section 3.4 describes the dataset. Section 3.5 outlines the empirical strategy, and Section 3.6 presents the findings. Section 3.7 concludes.

3.2 Conceptional Considerations and Background

3.2.1 Conceptional Considerations

The accumulation of knowledge over time necessitates increasingly interdisciplinary approaches, making it more complex for individuals to push the knowledge frontier and innovate (Burt, 2004; Wuchty, Jones, & Uzzi, 2007; Jones, 2009; Bloom et al., 2020). While digitization offers extensive online access to both tacit and explicit (codified) knowledge (Davenport & Prusak, 1998; Faraj et al., 2016; Rayna & Striukova, 2021), access itself does not guarantee its effective application in creating new knowledge (Mokyr, 2002). In fields like complex technologies, which require hands-on learning and specialized training (Ludwig et al., 2014; Giorcelli, 2019), mere access is insufficient. Knowledge creation depends on combining tacit and explicit knowledge across fields (Nonaka, Takeuchi, & Umemoto, 1996), but tacit knowledge is difficult to codify and often necessitates personal interaction (Audretsch, 1998; von Krogh, 1998; Breschi, 2001; Gertler, 2003; Popadiuk & Choo, 2006). Creating spaces that facilitate such exchanges among diverse individuals can foster innovation (Peschl & Fundneider,

2012) and strengthen the innovation ecosystem (Asheim & Gertler, 2009). However, it is an ongoing debate how to effectively design such spaces.

User-driven innovations offer another pathway to advancing the knowledge frontier. Users play a key role in open and collaborative innovation processes, often leading to radical product developments (Harhoff, Henkel, & von Hippel, 2003; Henkel & von Hippel, 2004; Baldwin & von Hippel, 2011; Gambardella, Raasch, & von Hippel, 2017; Boutillier et al., 2020; Preißner, Raasch, & Schweisfurth, 2024). User innovators freely share their ideas without seeking compensation (Franke & Shah, 2003; Jong et al., 2015). These voluntary information spillovers enhance the broader innovation ecosystem and promote collaborative innovation (Harhoff, Henkel, & von Hippel, 2003; Baldwin & von Hippel, 2011; Boutillier et al., 2020). Digital fabrication technologies, such as AM, further empower users to become innovators by allowing for rapid prototyping and fast improvement cycles (Baldwin & von Hippel, 2011; Fleischmann, Hielscher, & Merritt, 2016; Franke & Lüthje, 2020). Despite this willingness to share, barriers to the diffusion and development of ideas persist (Svensson & Hartmann, 2018; Jeppesen, 2021).

Given the critical role of user innovation, community workshops offer essential infrastructure and community support to enhance these activities. These spaces, including FabLabs, Hackerspaces, and Makerspaces, foster knowledge sharing and democratize innovation by providing access to tools and a collaborative environment (von Hippel, 2005; Halbinger, 2018; Bell et al., 2019). Such workshops are communities that prioritize openness, mutual learning, and creativity (Jong et al., 2015; van Holm, 2017; Browder, Aldrich, & Bradley, 2019; Zakoth, Mauroner, & Emes, 2023). While researchers recognize the potential of these spaces to boost local innovation (Gertler, 2003; Harhoff, Henkel, & von Hippel, 2003; van Holm, 2017), their actual impact or integration within local innovation ecosystems remains unclear due to challenges in quantifying outcomes such as economic benefits or the quality of innovations.

Combining accessible knowledge, user-driven innovation, and collaborative environments like FabLabs can thus play a crucial role in pushing the knowledge frontier and fostering innovation. Analysing how FabLabs influence local innovation, particularly through measurable outcomes such as patent filings, provides an unique opportunity to gain valuable insights into their impact on fostering and democratizing innovation. This might also inform policy makers how to effectively design innovation policies to support innovation originating in user communities.

3.2.2 Background on FabLabs

FabLabs were the first organized community workshops that promoted and enabled the fostering of skills in personal and digital fabrication (Blikstein, 2013). The first FabLab was established at the MIT Center for Bits and Atoms (CBA) by Neil Gershenfeld in 2003 as a result of Gershenfeld's course "How to Make (Almost) Everything" (Gershenfeld, 2012; Garcia-Ruiz, Lena-Acebo, & Rocha Blanco, 2023). The original idea of FabLabs is to transition from personal computation to personal fabrication, empowering individuals to produce their own customized technology rather than consuming mass-produced technology (Mikhak et al., 2002; Karagianis, 2006). For this, FabLabs are designed as physical spaces providing both regular and digital fabrication tools for everyone to use.

Digital fabrication, defined by Neil Gershenfeld, involves using computer-controlled tools to design and produce physical objects on demand (Gershenfeld, 2012). It is the combination of designing artifacts digitally using computer-aided-design software (CAD) and materializing these artifacts on demand using digital production machines for printing, milling, cutting, and shaping (Savastano et al., 2017).[2] Unlike traditional fabrication, which depends on molds and assembly lines, digital fabrication allows the direct production of the digitally designed object. Digital fabrication thus enables personal fabrication, fosters creativity, rapid prototyping, and thus innovation (Cutcher-Gershenfeld, Gershenfeld, & Gershenfeld, 2018; European Patent Office, 2020). By democratizing access to these technologies, FabLabs serve as platforms that empower local makers, inventors, and entrepreneurs while also functioning as educational hubs for learning, creating, and mentoring in digital tools, fabrication, and STEM subjects at large (Fab Foundation, 2023).

The values and principles of FabLabs are written in the "Fab Charter" and must be acknowledged by every FabLab. The Fab Charter mandates openness to everyone[3], a standardized set of tools[4], and the promotion of knowledge sharing by actively participating in the global FabLab network (FabLab Foundation, 2012). Today, FabLabs have

[2]The digital production machines include, for example, computerized numerical controlled mills (CNC mills), tools of AM and laser cutters (Savastano et al., 2017; Cutcher-Gershenfeld, Gershenfeld, & Gershenfeld, 2018).

[3]FabLabs may cater to the following communities: schools, artists, university students, entrepreneurs, governments, companies, and the general public

[4]The list includes the following (but can be extended according to the needs of the FabLab): 3D printers and scanners, laser cutters (2D/3D), CNC milling machines, suites of electronic components, and programming tools for low-cost, high-speed micro-controllers and onsite rapid circuit prototyping. For details see http://inventory.fabcloud.io/

evolved into a global network of over 1,750 labs in more than 100 countries. With the numbers constantly increasing, the FabFoundation was established in 2009 to support the global spread and provide an institutional structure for the global network of FabLabs. While the FabFoundation offers a digital fabrication education program, the "Fab Academy",[5] it does not provide funding. FabLabs are nonprofit organizations and receive their funding either from institutions such as universities or donor organizations. As such, they usually offer their tools and resources at no cost or at material costs. Alternatively, they may operate as social clubs with subscription models.

The broad access to cutting-edge technologies provided by FabLabs for entrepreneurs, students, tinkerers, hobbyists, and interested citizens (Troxler & Wolf, 2010) has the potential to unlock inventive capabilities across a diverse user base and foster grassroots innovation (Gershenfeld, 2007). While not all tools available in a FabLab are extremely expensive, having all the different tools accessible in one place generates added value for users (van Holm, 2017) and enables creativity and prototyping (Bogers et al., 2017).

Innovation in FabLabs is driven not only by democratized access to tools but also by the active promotion and fostering of knowledge sharing within the community (Troxler et al., 2020). Although not all users of a FabLab have the initial intention to invent and commercialize ideas (Halbinger, 2018), the fertile environment offered by FabLabs creates a space where people with diverse backgrounds, but often similar values, meet and interact. These interactions can range from meeting at the coffee machine and chatting, to mentoring, exchanging ideas, and collaborating on projects, and can be enough to exchange knowledge and foster innovation (Berends et al., 2006; Sandvik et al., 2020; Andrews, 2023; Baruffaldi & Poege, 2025).

University-integrated FabLabs may further foster knowledge spillovers and innovation, as their users have access to the latest research which can create valuable knowledge spillovers among FabLab users and university researchers (Gertler, 2003; Furman et al., 2006). In rural areas they provide the necessary infrastructure that enable tinkerers. Experimenting with the tools available at the FabLab can transform tinkerers into "accidental entrepreneurs" or "accidental innovator" (Shah & Tripsas, 2007; Halbinger, 2018). Standardized tools worldwide facilitate knowledge transfer by enabling

[5] Designed by MIT's CBA, users of FabLabs can enroll, and their home FabLab serves as a "classroom" (Fab Foundation, 2023; FabAcademy, 2024).

makers to replicate and learn from others' ideas.[6] FabLabs can thus act as incubators of ideas (Zakoth, Mauroner, & Emes, 2023). Promising ideas and inventions could then be selected and, once ready, channeled to a connected incubator for further development and preparation for commercialization.[7]

FabLabs strike a balance between promoting open-source innovation and allowing commercialization. While the Fab Charter encourages knowledge sharing, it also allows for commercialization of inventions, provided they remain accessible for others to learn from (Fab Foundation, 2023). However, this does not necessarily conflict with intellectual property (IP) rights. As Moritz et al., 2024 point out, even in the context of user innovation, patents and utility models can foster knowledge diffusion by allowing inventors to define the conditions under which their invention may be used by others. FabLabs have produced both open-source projects (Troxler & Wolf, 2010), entrepreneurial ventures (Mauroner, 2017), and patented inventions (Mortara & Parisot, 2018). However, not everything developed in a FabLab can be commercialized or patented (Franke & Lüthje, 2020).

While FabLabs are not the only community workshops today, their structure,central goal, and their role in fostering innovation make them unique. Three types of community workshops with similar values of knowledge sharing, learning, making, and inventing are prevalent in the so-called "Maker movement", namely Hackerspaces, Makerspaces, and FabLabs, on which this paper focuses.[8] Although these types differ in their focus and organization, the names are often used interchangeably due to a lack of standardized definitions (Mersand, 2021).

While FabLabs focus on digital and personal fabrication, the origin of Hackerspaces dates back to the 1990s when the first computers emerged, and hackers and tinkerers experimented with computer software and hardware in collaborative workshops (Moilanen, 2012; van Holm, 2014; Mersand, 2021). Similar to FabLabs, Hackerspaces are rather homogeneous in structure and equipment (Cuntz & Peuckert, 2023), but, contrary to FabLabs, their focus lies on software, although there are no limitations to the tools and crafts they offer (Moilanen, 2012; Cuntz & Peuckert, 2023).

[6]One example of knowledge sharing among FabLabs is a wireless network improvement project that originated in Boston, was enhanced in Norway, tested in South Africa, deployed in Afghanistan, and is now commercially active in Kenya (Mikhak et al., 2002; Gershenfeld, 2012).

[7]The TU Ilmenau, for example, aims to do this with its university-funded FabLab and incubator Technische Universität Ilmenau, 2024.

[8]See Zakoth, Mauroner, and Emes (2023) and van Holm (2014) for details on Makerspaces and Hackerspaces.

In contrast, Makerspaces offer a broader range of crafts, including glasswork, sewing, woodworking, and digital fabrication (Browder, Aldrich, & Bradley, 2019). Echoing Dale Dougherty's assertion that "everyone is a maker" (Dougherty, 2012), Makerspaces support diverse activities rather than standardized tools. Unlike FabLabs, their openness and the offered tools vary greatly, and they are not organized in a global institutional network like FabLabs. For example, in Germany, both community repair shops and those offering woodworking and digital fabrication are listed as Makerspaces.

Makerspaces obtained their name from the "Maker Movement" although the community workshops existed prior to the movement. The expiration of AM technology patents in 2006 and the 2005 announced open-source project to build a consumer 3D printer (RepRap Project) sparked a renewed interest in the "Do-it-yourself" (DIY) culture (West & Kuk, 2016; Moritz et al., 2024). At the same time, Dale Dougherty published the first MAKE Magazine, which offers DIY guides on a variety of topics (Maker Media GmbH, n.d.), and with this, the "Maker Movement" was born, and community spaces began calling themselves "Makerspaces" (van Holm, 2014). Individuals in the Maker Movement, the makers, share a passion for making, repairing, and hacking (Dougherty, 2012).

FabLabs stand out within the Maker Movement due to their institutional support from the FabFoundation, the global network of labs, standardized set of tools, their focus on digital fabrication and education, and their goal to foster knowledge sharing and innovation. They may therefore influence local innovation and patenting more decisively compared to other community workshops. The knowledge diffusion within FabLabs fosters a culture of continuous learning and creativity combined with access to advanced tools that enable rapid prototyping and testing, which is crucial for developing new products. University-integrated FabLabs may further enhance this effect due to knowledge spillovers between academic research and FabLab users. The social interactions and networking among diverse makers encourage collaborative problem-solving and innovation further promote the positive effect. The emphasis on standardization and knowledge sharing in FabLabs makes them particularly suited for analyzing the impact of community workshops on local innovation outputs.

3.3 FabLab Survey

3.3.1 Survey Data

To gain deeper insights into the operational dynamics and mechanisms of German FabLabs, I conducted two anonymous online surveys using the Qualtrics survey platform.[9] The first survey was directed at the managers of German FabLabs, with the aim of gathering information on the initial motivations behind founding a FabLab, organizational aspects (including funding sources and structure), collaborations with and integration into local innovation ecosystems, user statistics, additional services provided (such as teaching, training, and competitions), and the outputs generated by FabLab users. The second survey targeted the users of German FabLabs, focusing on their perspectives and usage patterns within the FabLab environment. The questionnaires for both surveys are included in the appendix (see C.3).

The surveys were distributed to the FabLab managers via email on November 11, 2024, with anonymous responses collected through until December 19, 2024. Managers were asked to share the user survey with their respective users. Among all 86 active German FabLabs (including those founded after 2019), 23 managers responded to the survey, yielding a response rate of approximately 27%. Though this rate may seem rather low, it is close to the average response rate of online surveys of 31% (Nulty, 2008). Responses by FabLab users were more scarce, amounting to 17.

3.3.2 Survey Results

To gain deeper insights into the operations and dynamics of the German FabLabs, I conducted an online survey among the managers and the users of the FabLabs. Detailed figures are reported in Appendix C.3.3.

Results of the Manager Survey

Among all 86 active FabLabs in Germany, 23 managers responded, yielding a response rate of 27%.[10] Most of the FabLabs represented in the survey were founded prior to 2020, predominantly motivated by a commitment to the FabLabs' overall mission

[9]The survey passed the ethical approval from the Ethics Commission, Department of Economics, University of Munich in October 2024.

[10]This aligns with the average response rate of online surveys of 31% (Nulty, 2008).

of providing open access, fostering knowledge sharing and digital fabrication skills (80%). Other motivations to found a FabLab comprised the FabLabs' focus on digital fabrication (40%), the teaching opportunities provided through the FabAcademy (38%), and being part of the global FabLab network (20%). These findings confirm the assumption that FabLabs offer a unique added value that is well recognized and appreciated by supporters.

FabLabs are funded through a diverse channels, with membership fees (48%) and private funding (43%) being the most common, followed by university (35%) and government support (32%). Other sources include municipal governments, projects with local firms, and local foundations. Funding partners are often also collaboration partners. Notably, 52% of the managers aim to secure long-term funding through local collaborations with small-medium sized enterprises (SMEs) and startups (50%), NGOs (40%), large enterprises (30%), political entities (20%), and others such as local cultural and educational institutions (25%). However, these collaborations serve purposes beyond funding: 72% aimed at the exchange of knowledge and resources, 68% focused on organizing joint workshops, 53% engaged in joint research projects, and 40% provided internships and project opportunities for students.

Despite these collaborations and interactions, 76% of the managers reported that their FabLabs were not part of a local innovation ecosystem. However, 24% reported being part of ecosystems such as FabCity Hamburg, UnternehmerTUM, and regional and supra-regional innovation networks. Considering their extensive collaborations with local firms, governments and universities, FabLabs may play a more significant role in local innovation networks than managers recognize.

FabLabs are typically managed by both paid staff and volunteers and promote their activities mostly via social media (75%), local newspapers (70%), fairs and conferences (52%), collaborations with local firms (80%) and universities (35%), and collaborations with local social clubs and educational institutions. FabLabs are open to the general public, although some report restrictions, such as age limits (e.g., no unsupervised children under 14) or limited hours of operation. The majority of FabLabs are open more than three days per week or more, with 75% reporting over 20 users per week, 35% more than 40 users per week, and 20% over 50 weekly users.

FabLabs are primarily used by hobbyists, inventors, entrepreneurs, students, faculty members, and firms which is consistent with prior findings that community workshops attract both hobbyists and innovators (Halbinger, 2018). Some FabLabs also collaborate with schools. The outcomes of FabLab users reported by managers are divers with almost all reporting that users mostly work on personal projects, collaborate on non-profit initiatives, and develop functional prototypes. About half of the FabLabs reported that users complete qualification papers (e.g., theses or dissertations), launch businesses or products, and develop software development. 15% of the managers report patent or utility model applications as user outcomes. The high rate of prototype development (80%) and entrepreneurial activity underscores FabLabs' role in early-stage innovation, providing spaces for ideation, experimentation, and prototyping. FabLabs contribute to academic knowledge creation and promote open innovation and knowledge sharing, which is highlighted by the prevalence of qualification papers and non-profit collaborations. Although only 15% of FabLabs report patenting activity, prior research suggests that some open-source innovations transition to proprietary innovations over time (West & Kuk, 2016). Moreover, not all ideas are actually patentable (Franke & Lüthje, 2020).

FabLabs actively promote collaborations among users through regular networking events, workshops, and informal exchanges (e.g., casual conversations around coffee machines), and by inviting users to document and share their projects. While 35% use internal sharing platforms, others rely on wikis, chat platforms, social media, websites, or newsletters. 25% reported limited or nonexistent knowledge-sharing processes which is consistent findings by Troxler et al. (2020). Although FabLabs promote knowledge sharing, its implementation at the individual lab level is not always consistent.

FabLabs offer mostly technical courses (90%) and school programs (75%), followed by contract work (40%), entrepreneurship courses (35%), co-working spaces (20%), FabAcademy courses (10%), and patenting support (5%). Other activities comprise internal product development programs, participation in social club projects, and the support of university courses. While introductory workshops are generally free of charge, specialized courses may involve fees. These offerings highlight FabLabs' educational mission and their role in skill acquisition.

The main challenges reported by managers included securing volunteer support to keep labs operational and offer courses (50%), ensuring long-term funding (50%), and attracting and retaining users (37%). Other challenges included space constraints, financing staff, and identifying effective platforms to support community engagement. Informal interviews with managers during the 2024 FABUniverse Meeting indicate that many FabLabs initially relied on short-term government grants, necessitating the pursuit of alternative, sustainable funding sources over time.

Despite managers' perceptions, FabLabs exhibit strong engagement with local innovation ecosystem through their local collaborations and their efforts to become integrated in local communities efforts. Innovation ecosystems, broadly defined as dynamic networks of actors, activities, resources, and institutions — along with their complementary and competitive interconnections — that influence the innovative capacity of individual actors and groups of actors (Asheim & Gertler, 2009; Carayannis & Campbel, 2009; Granstrand & Holgersson, 2020), encompass the very actors with which FabLabs engage. Their long-term sustainability depends on successful integration into these ecosystems, particularly for securing funding and attracting users. At the same time, FabLabs contribute to local innovation by providing access to digital fabrication tools, fostering knowledge exchange, and promoting user-driven innovation. This bidirectional relationship underscores the role of FabLabs as both beneficiaries and enablers of local innovation dynamics.

Results of the User Survey

A total of 17 users across nine FabLabs responded to the user survey, revealing insights into their motivations, activities, and demographics.

The majority of the users use a single FabLab within 10 kilometers of their residence, underscoring prior findings that users value local access to tools and that this fosters user-driven innovation (Svensson & Hartmann, 2018). Users learned about the FabLab mostly through internet searches and personal recommendations, with additional sources including newspaper articles, events, and deliberate searches for such facilities.

All respondents work on personal projects, with 80% also aiming to acquire new skills and to learn about new technologies. Other motivations included engaging with the FabLab community (69%), collaborations with other users (60%), developing functional prototypes (32%), and pursuing entrepreneurial activities (13%). These findings reinforce FabLabs' educational role in fostering digital fabrication and through it innovation and entrepreneurship (Baldwin & von Hippel, 2011). While usage frequency varies, most users visit the FabLab at least once per week, indicating FabLabs' function as consistent resources for iterative experimentation and learning.

Collaboration among users is common, but intensity levels vary. Users working on their projects frequently exchange ideas with other users or receive assistance, while other users predominantly work in collaborative projects. These responses underscore the culture of knowledge sharing prevalent in FabLabs

Users value most the access to tools and technologies, the community and networking opportunities, and skill acquisition. Most users had participated in FabLab workshops. The outcomes of their activities range from skill acquisition and personal projects to prototypes, networking, and product launches. Projects are mostly shared through platforms such as GitHub and Thingiverse. Projects were predominantly categorized into topics of engineering and technology, followed by art/design, sustainability, and software. As areas of improvement users listed more advanced tools, additional workshops and community events, and improved knowledge management technologies. Nonetheless, the respondents were overall very satisfied with their FabLab experience.

Demographically, respondents were predominately male, aged between 35 and 44 years, with additional representation from the 25–34 and 45+ age groups. A master's degree was the most commonly reported educational attainment, and most users were employed by large companies. Professional backgrounds were primarily in information technology and mathematics, followed by electrical engineering and education. These findings indicate that while FabLabs foster innovation, barriers to broader participation appear to exist — particularly in gender and educational diversity.Informal interviews at 2024 FABUniverse Meeting confirmed these challenges, but also highlight the commitment of the FabLab community to address these issues.

3.4 Data and Descriptive Analysis

To analyze the impact of FabLabs on regional innovation, I build a regional panel dataset between 2003 and 2019 that combines the following data.[11]

3.4.1 Data on FabLabs and Other Community Workshops

I obtained data on the German FabLabs, including location, and funding scheme, from self-reported sources such as the FabFoundation website and the German FabLab community blog "Fabrikationslabor". For control purposes, I also collected data on Mak-

[11] I restrict the panel to the years prior 2020 to avoid biases in the results due to the COVID-19 pandemic which led to the closure of public spaces in Germany starting in March 2020 (Jungblut, 2020; Bundesministerium für Gesundheit, 2024; Hochschulrektorenkonferenz, 2024). The accompanying restrictions made establishing new FabLabs risky and, in many cases, unfeasible. Therefore, my analysis excludes data from 2020 onward.

erspaces and Hackerspaces. The main source for Makerspaces was the self-reported list of German Makerspaces provided by the the "Verbund offener Werkstätte" (VoW), while the data on Hackerspaces originates from the Hackerspace website' self-reported list of Hackerspaces worldwide.[12] The founding year of each community workshop was added through manual research.[13]

Relying on self-reported information may entail that the data does not fully capture the landscape of FabLabs, Makerspaces, and Hackerspaces in Germany, or that community workshops report on multiple platforms. However, founding a FabLab is a deliberate decision and typically implies that the founders are aware of the Fab Foundation and its benefits.[14] I am thus confident that the data I collected from the Fab Foundation and the German FabLab community covers the majority of German FabLabs.

To deal with possible multiple reporting and to ensure a clean dataset, I applied specific classification rules to differentiate FabLabs, Makerspaces, and Hackerspaces. A community workshop is classified as a FabLab if it is listed on any FabLab website Community workshops listed only on the Hackerspace website are classified as Hackerspaces, and those listed on the VoW website are classified as Makerspaces. In cases where a Makerspace also reported itself on the Hackerspace website, I classify it as Makerspace only.

I identified 77 FabLabs in Germany which were founded by the end of 2019 (see Figure C.1). Of these, 19 (25%) FabLabs are university-integrated, 47 (62%) operate as social clubs, and the remaining 11 labs are affiliated with research institutes or a nonprofit limited liability company. 132 Hackerspaces are reported on the Hackerspace website with a founding year before 2020, of which 72 Hackerspaces remain after applying the classification rules. Out of 324 Makerspaces founded before 2020, 223 are counted as Makerspaces in this study.[15]

For this study, I matched the locations of the community workshops with Germany's Functional Urban Areas (FUA), as defined by the OECD. FUAs identify urban centers

[12]Sources: fablabs.io,http://www.fabrikationslabor.de/fablabs_in_deutschland/, offene-werkstaetten .org, wiki.hackerspaces.org

[13]Sources encompassed individual community workshop websites, blog posts, social media platforms, the German commercial register, and local newspaper reports.

[14]The results of the manager survey confirm this assumption. For details on the manager survey see Section 3.3.2

[15]Out of the 132 total Hackerspaces, 42 also report as Makerspace but not as FabLab, 16 report also as FabLab and Makerspace. Two Hackerspaces report as also as FabLab but not as Makerspace. The 324 reported Makerspaces include 24 repair cafes and 19 purely artistic Makerspaces which were dropped. 58 are known FabLabs and were thus not counted as Makerspace.

(areas with over 50,000 inhabitants) independent of administrative boundaries and define their commuting zones capturing travel for work as well as travel for education, health, and other services. FUAs therefore define an area encompassing both the urban center and the relevant commuting zone, which allows for a more comprehensive consideration of a city's functional and economic area.[16] This approach is particularly relevant for analyzing FabLabs' impact on local innovation activities, as it accounts for potential users and inventors who commute to the community workshops (van Holm, 2017), enabling a broader analysis of innovation spillovers and collaborations that extend beyond administrative boundaries.

However, based on user survey responses indicating that many users are located near the FabLab, I also conducted analyses at the NUTS3 level. NUTS3 regions offer a more granular regional definition compared to FUAs, enabling the analysis of localized effects and allowing to focus on rural areas[17]. This granularity can reveal variations within smaller areas that might be averaged out in FUA-level analyses. Since NUTS3 regions often correspond to areas with specific local government policies, analyzing at this level can provide insights that are directly relevant to regional policymakers aiming to foster innovation through FabLabs. An additional benefit of NUTS3 analysis is its comprehensive coverage of all of Germany, capturing all community workshops, unlike FUAs, which do not encompass all such workshops. Figure 3.1 depicts the spatial distribution of community workshops by type across both FUAs (first row) and NUTS3 regions (second row). In total, 69 of 77 FabLabs, 65 of 72 Hackerspaces, and 205 of 223 Makerspaces are matched with the FUAs.

[16] For details on FUAs and their construction, see Dijkstra, Poelman, and Veneri (2019)

[17] The Nomenclature of Territorial Units for Statistics (NUTS) was developed by Eurostat to segment the EUs economic territory into standardized regions for regional statistics and policy interventions. NUTS3 regions, which correspond to districts (Kreise) and cities in Germany, typically have populations ranging between 150,000 and 800,000. I use the definition of NUTS3 regions of 2022 and respective shapefiles (Eurostat, 2024a, 2024b; Federal Office for Cartography and Geodesy (BKG), 2024; Federal Statistical Office, 2024).

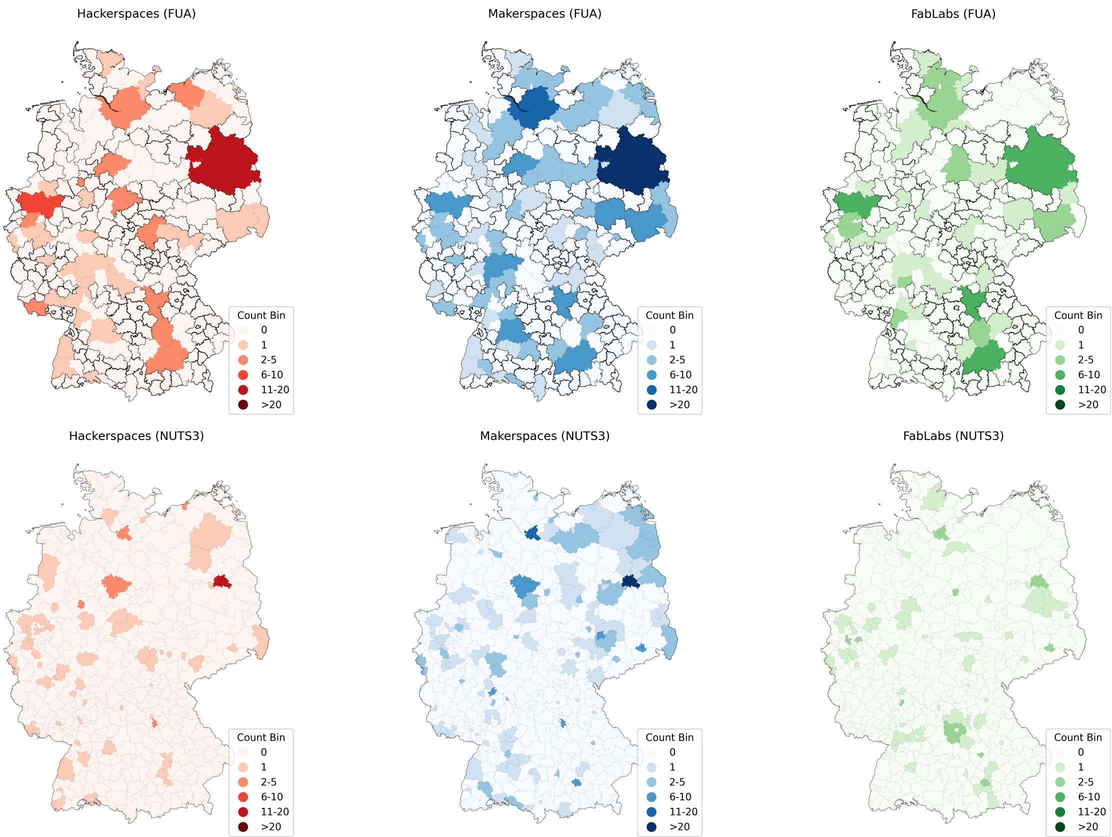


Figure 3.1: FUAs and NUTS3 with community workshops in Germany

Notes: This figure shows the the spatial concentration of community workshops in Germany. The first row displays their distribution across FUAs, while the second row shows them in NUTS3 regions. Columns represent different types: Hackerspaces (first), Makerspaces (second), and FabLabs (third). Workshop counts are categorized into six bins, with the darkest color indicating the highest concentration (above 20).

The map also reveals that community workshops are more likely to be located in the vicinity of larger cities, e.g., Munich, Berlin, Hamburg, Cologne, Dresden. However, there are also more rural areas that show community workshop activities, e.g. the FabLab in Wismar. The first FabLab in Germany was established in 2009 at RWTH Aachen. On average, six FabLabs were founded per year in the FUA-dataset and seven in the NUTS3-dataset, following a wavelike pattern with peaks (e.g. 11 new FabLabs in the FUA-dataset in 2012) and subsequent declines. (see Figures 3.2 and C.2, Panel a)). The sharp decline in 2019 can be attributed to the COVID-19 pandemic. Makerspaces show a similar wave-like founding pattern with peaks in 2011 and 2015. In contrast, Hackerspaces have remained relatively stable, with an average of three Hackerspaces founded annually in both datasets and a brief peak in 2014.

In total, 76 out of 96 FUAs have at least one type of community workshop. Out of the

76 FUAs with some type of community workshop, 37 FUAs have at least one FabLab, 67 FUAs have at least one Makerspace, and 38 FUAs have at least one Hackerspace. Notably, only three FUAs host only a FabLab, while the majority of FabLabs were founded in FUAs where either a Makerspace or a Hackerspaces already existed (see Figure 3.3, Panel a)). In fact, there are 31 FUAs where either a Makerspace or a Hackerspace existed prior to the FabLab. In 12 FUAs, both a Hacker- and a Makerspace preceded the FabLab, in 12 FUAs a FabLab was preceded only by a Makerspace, and in six FUAs the FabLab was the first community workshop.

At the NUTS3 level, 158 out of 401 NUTS3 have at least one community workshop, with 62 hosting a FabLab, 57 a Hackerspace, and 113 a Makerspace (see Figure 3.3, Panel b)). In 88% of the 41 NUTS3 where multiple types of community workshops exist, a Makerspace or a Hackerspace was founded before the FabLab. In 10 out of 15 regions with all three types both a Makerspace and a Hackerspace preceded the FabLab. In 26 NUTS3 was a FabLab the first community workshop to be founded. In 15 out of 20 NUTS3 with both a Makerspace and a FabLab the Makerspace preceded the FabLab.

These differences in the founding order and co-location of community workshops suggest heterogeneity in local conditions—specifically, in the structure of the local innovation ecosystem—that may influence the impact of FabLabs on regional innovation activities.

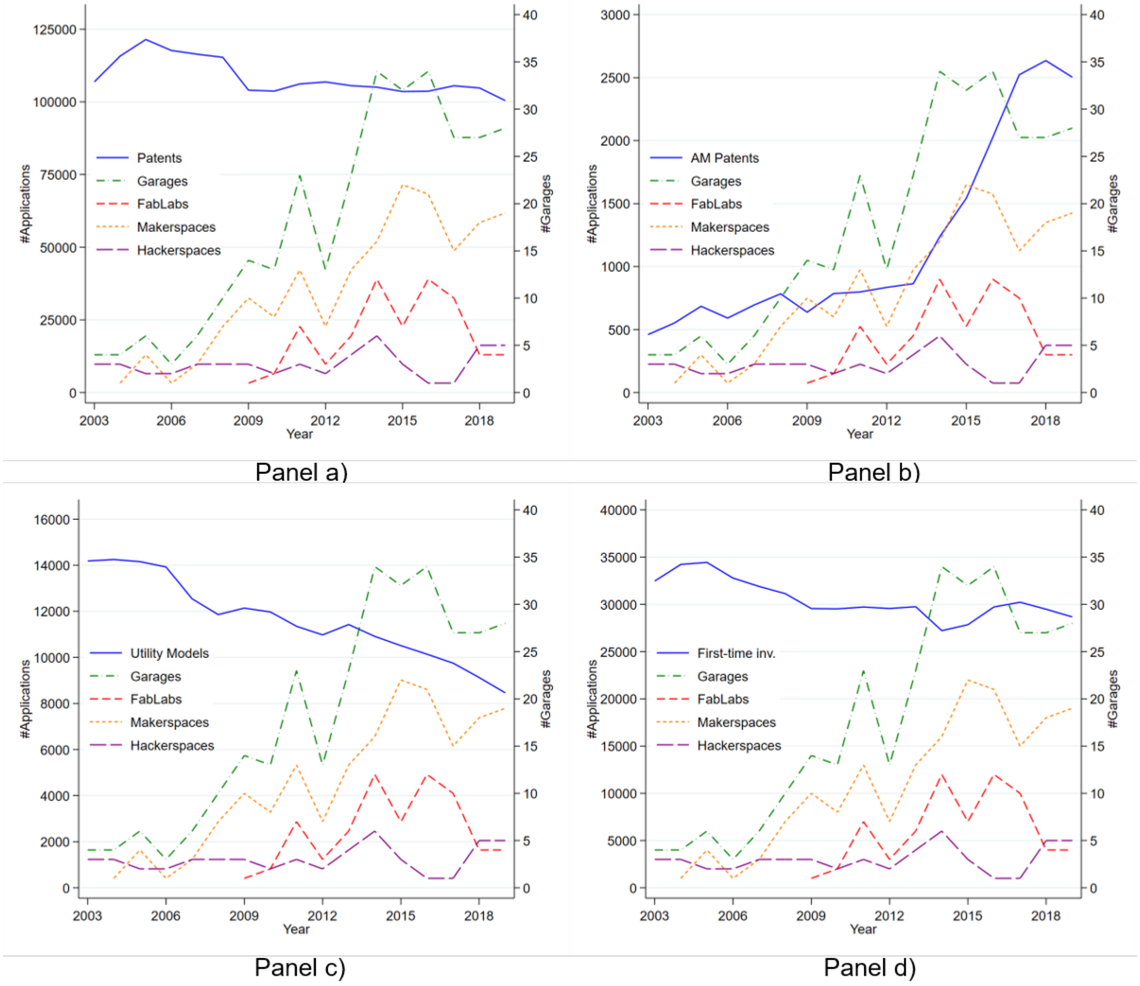


Figure 3.2: Trends in patenting activity and community workshops across FUAs (2003–2019)

Notes: This figure presents the temporal development of patenting activity and community workshops across all FUAs from 2003 to 2019. Panel (a) shows the total number of patent applications, while Panel (b) focuses on patent applications related to AM. Panel (c) displays the number of utility model applications, and Panel (d) illustrates the number of applications filed by first-time inventors. For comparison, Figure C.2 provides the corresponding time trends at the NUTS3 regional level.

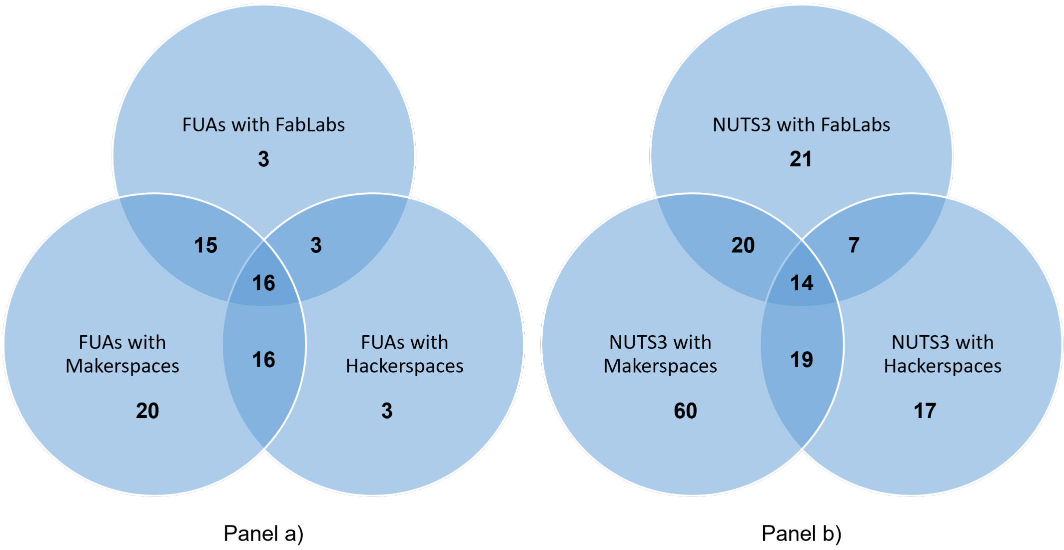


Figure 3.3: FUAs and NUTS3 with community workshops

Notes: The left Venn diagram (Panel a)) shows the number of FUAs with community workshops: 37 have at least one FabLab, 38 have at least one Hackerspace, and 67 have at least one Makerspace. The right Venn diagram (Panel b)) presents the distribution across NUTS3 regions: 62 contain at least one FabLab, 57 at least one Hackerspace, and 113 at least one Makerspace.

3.4.2 Data on Patents and Utility Models

To measure the effect of the FabLabs on local innovation activities, this study compiles a new and unique dataset of patent and utility model applications with at least one German-based inventor for the years 2003-2019 and aggregates the applications to the FUA-year and NUTS3-year level.

Data on Patents

To capture the effect of FabLabs on local innovation output and knowledge creation, I follow the literature and use the count of patent applications per year in a region (Acs, Anselin, & Varga, 2002; Lee, Florida, & Acs, 2004; Moreno, Paci, & Usai, 2005). While patents do not capture the entirety of innovative output, their comprehensiveness and availability make them a widely used measure of innovation activity. To assess whether FabLabs increase patent applications by individuals rather than by organizations, I include the separate counts of patents filed by organizations and by individuals. To analyze the potential of FabLabs to enable tinkerers to become inventors, I use the number of applications by first-time inventors per region-year, as well as applications by individual applicants as dependent variables. To assess the impact of FabLabs on related AM technologies, I include the number of patent applications related to AM as dependent variables. For a qualitative measure, I use the forward citations and the number of technical areas cited (Harhoff et al., 1999; Hall, Jaffe, & Trajtenberg, 2005; Higham, de Rassenfosse, & Jaffe, 2021).

I obtained data on patent applications and addresses of inventors using the fall 2023 version of PATSTAT.[18] to measure innovative activity. Besides focusing on application counts, I am also interested in the number of applications by first-time inventors. Although, PATSTAT assigns a unique person ID to each inventor based on their name, address and country code, spelling errors or missing addresses can result in multiple IDs for the same inventor. To resolve this, I disambiguated inventor names and locations, creating a dataset that uniquely identifies individuals and their locations (for details see Appendix Section C.4).[19]

This disambiguation also allowed me to partly fill in the missing addresses for inventors. For the years 2003 to 2019, I matched 1,844,030 unique patent applications filed by 371,739 German-based inventors to FUAs, and 2,136,287 patent applications[20] filed by 457,289 inventors to NUTS3 regions. As the FUAs represent the urban centers of Germany, the difference of 292,257 patent applications between NUTS3 and FUAs datasets implies that 20% of the patent activity originates from rural areas.

For applications with multiple German-based inventors, I followed the fractional counting approach of Pfister et al., 2021 and Lehnert et al., 2022, and assign the respective fraction of the total number of inventors for a patent to a region.[21] The regional distribution of patent applications per capita is shown in Figure 3.4, Panel a). Notably, regions in southern and western Germany (e.g., Munich, Heidelberg, Stuttgart, Cologne, Nuremberg, and Frankfurt) exhibit a particularly high patent activity per capita. To account for the specific focus of FabLabs on digital fabrication rather than chemicals, I include a count variable that excludes patent applications in the chemical area.[22] Since FabLabs provide access to tools for AM, this study also distinguishes AM-related patent applications by matching a dataset from the Euro-

[18]PATSTAT is the comprehensive patent database of the European Patent Office (EPO).

[19]I am grateful to Patrick Lehnert for providing me with the dataset to conduct the inventor disambiguation.

[20]The total number of observations in the FUA-dataset is 3,893,699, and 4,800,922 in the NUTS3 dataset.

[21]For example, for a patent with three inventors—one located in region A, one in region B, and the third based abroad—the patent is allocated one-third to region A and one-third to region B. The share attributed to the inventor abroad is excluded from the analysis.

[22]For this, I use the main technological areas defined by Schmoch (2008), which is based on the IPC classification of patents.

pean Patent Office (EPO) covering patents associated with AM technologies between 2000 and 2019 (European Patent Office, 2020, 2023).[23] The match resulted in 20,158 matches in the FUA dataset and 22,840 matches in the NUTS3 dataset.

The final set of dependent variables used to assess the extensive margin of FabLabs' impact is aggregated by application year and region, and includes total patent applications, non-chemical patents, patents filed by first-time inventors, patents filed by individual applicants, and patent filed by organizational applicants (e.g., firms). To evaluate the intensive margin, I calculate the average number of forward citations and the number of technological areas cited per patent by application year and region.

Data on Utility Models

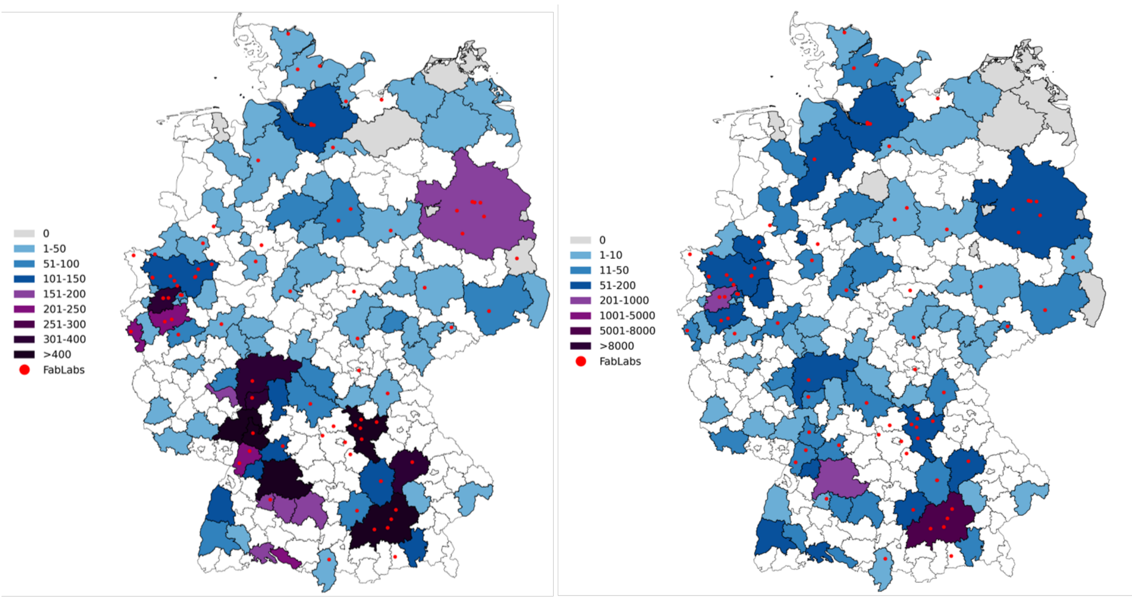
Utility models are a distinct form of intellectual property rights in Germany granting exclusive rights for up to 10 years (Deutsches Patent- und Markenamt, 2025a). They are easier and faster to obtain than patents, as the application and granting requirements are less stringent, and the application costs are lower (Moritz et al., 2024; Deutsches Patent- und Markenamt, 2025b). To apply for a utility model, the invention must be new, involve an inventive step, and be commercially applicable. However, the granting process checks only formal requirements and the novelty is assessed only if the utility model is challenged (Deutsches Patent- und Markenamt, 2025b). Utility models protect only technical inventions and exclude processes, plants or biotechnological inventions (Deutsches Patent- und Markenamt, 2025b). As cost-effective and fast way to protect inventions, utility models may thus be particularly suitable for inventors in FabLabs. Especially, because they can serve as prior art if the inventor later files for a patent (Moritz et al., 2024), utility models are frequently employed as initial step in strategic patenting (Heikkilä & Verba, 2018). For these reasons, utility models are also include as dependent variable in the analyses.

To identify utility models filed for in Germany, I filtered applications using the relevant kind code in PATSTAT. I collected applicant addresses directly from online database of the German Patent and Trademark Office (DPMA).[24] Unlike patents, utility model applications require only one address for communication purposes and thus cannot be fractionally assigned to the inventors' location in case of multiple inventors. Instead, they are assigned to a single region. Between 2003 and 2019, the DPMA received

[23] I am grateful to the EPO and Ilya Rudyk for providing the dataset.

[24] For utility models the inventors' addresses are not included in PATSTAT. Data source of addresses: DPMARegister, https://register.dpma.de/DPMAregister/pat/basis

199,498 utility model applications for which an address was recovered. The regional distribution of utility model applications is shown in Figure 3.4, Panel b), and reveals a spatial concentration of utility models similar to patent applications, especially in southern and western Germany.



Panel a) Panel b)

Figure 3.4: Regional variation in patenting activity per capita in FUAs

Notes: This figure illustrates the geographic distribution of patenting activity in FUAs from 2003 to 2019, normalized by population. Panel (a) displays the number of patent applications per capita, while Panel (b) presents the number of utility model applications per 100 inhabitants. For corresponding data at the NUTS3 regional level, see Figure C.3.

3.4.3 Regional Data

To control for regional differences which might influence local activity, I use data from the INKAR database, provided by the by the German Federal Institute for Building, Urban and Spatial Research.[25] INKAR offers comprehensive regional statistical data on a wide range of socially relevant topics, including education, demographics, employment, economy, housing, transportation, and the environment. The data is provided on yearly basis and either on municipality (Gemeinde) level or district (Kreis) level.

For this study, I use the following variables: population, GDP, and unemployment. To control for urban and rural areas in the NUTS3 dataset, I use the INKAR database,

[25]Source: https://www.inkar.de/

which classifies municipalities as rural or urban.[26] To link municipalities to FUAs and aggregate the data to the FUA-year level, I use the list provided by the OECD, which maps each FUA to its constituent municipalities (OECD, 2024).To link districts to FUAs, I utilize the list provided by INKAR, which maps districts to municipalities (Federal Office for Building and Regional Planning (BBSR), 2024). For the NUTS3 dataset, explanatory variables were added using the direct NUTS3 code mapping, as the INKAR file already includes the NUTS3 codes.

3.4.4 Descriptive Analysis

Panel (a) of Figure 3.2 illustrates that, apart from an increase between 2004 and 2009, the annual number of patent applications within FUAs remains relatively stable at approximately 100,000. In contrast, AM-related patent applications grow significantly (see Figure 3.2, Panel b)). Panel (c) of Figure 3.2 illustrates the overall negative trend for utility model applications in FUAs which aligns with DPMA reports (Deutsches Patent- und Markenamt, 2022). The decline is non-linear, with stable numbers of 14,000 annual applications until 2006, followed by a drop and subsequent decreasing numbers. Between 2003 and 2005, the applications by first-time inventors initially increase, coinciding with a rise in the number of newly founded Makerspaces. (see Figure 3.2, Panel d)). However, from 2005 to 2009, the numbers decrease further, stabilizing in 2009, the year the first FabLab opened in Germany. After remaining constant for five years, there is a drop in 2014 from 30,000 to 27,500, followed by a recovery to pre-2014 levels. The patterns are similar for the NUTS3 regions but differ slightly in terms of numbers with the NUTS3 regions having higher numbers (details see Figure C.2).

Table 3.1 presents the distribution of the main dependent variables in Panel A and the control variables in Panel C for the FUA-year dataset. On average, 875 patents are filed annually in a FUA. There are 116 patent applications by first-time inventors. 105 Patents are filed by individual applicants, and 753 are filed by organizations. Each patent averages 9 forward citations and cites six different technological areas. On average, 10 patent applications are related to AM, and 123 utility model applications are filed annually within a FUA.

[26]Based on the Regionalstatistischen Raumtyp 7 classification by the Bundesministerium für Digitales und Verkehr (BMDV), 2021, I calculated the share of each regional type within a NUTS3 region. Based on these shares, I assigned an indicator of whether a NUTS3 region is urban or rural.

For the NUTS3-year dataset, the dependent variables are generally smaller, as reported in Table 3.1, Panel B. On average, 238 patent applications originate from a NUTS3 region. NUTS3 regions see an average of 25 patent applications by first-time inventors and 31 patents filed by an individual applicant. Patents in NUTS3 regions average 9 forward citations and cite five different technological areas. On average, three patent applications are related to AM, and 24 utility model applications are filed annually in a NUTS3 region.

In the FUA dataset, treated and control FUAs differ primarily in size, with treated FUAs having larger populations and a higher GDP compared to the control group. Additionally, treated FUAs show more intense patent activity, with an average of 1,433 patents per year, compared to 525 patents per year in the control group (see Table C.1 Panel A and Table C.2) Panel A.

The pattern is reversed for patent and utility model activity in the NUTS3 regions. On average, treated NUTS3 regions file 238 patents per year, while the control group files 268. Patents in both groups are similar in terms of average forward citations and areas cited. Treated NUTS3 regions have a lower GDP and population, and are less urbanized (see Table C.2 Panel B, and Table C.1 Panel B). To account for these differences, I use never treated and not yet treated observations as control groups.

Table 3.1: Summary statistics

	N	Mean	Median	SD	Min.	Max.
Panel A: dependent variables FUA dataset						
Patents	1,632	874.78	355.81	1,529.95	3.50	10,676
NoChem. Pat.	1,632	720.95	299.03	1,350.38	3.50	10,259
First-time	1,632	115.90	50.40	188.71	0.00	1,375
Pats. by ind.	1,632	105.15	36.15	219.13	0.00	2,177
Pats. by orgs.	1,632	753.21	307.61	1,339.39	3.50	9,366
Forward cit.	1,632	8.68	8.65	4.27	0.52	27
Areas cited	1,632	5.77	5.61	1.43	1.88	12
AM patents	1,632	9.59	2.11	25.57	0.00	373
Utility models	1,613	122.56	39.00	341.62	1.00	3,620
Panel B: dependent variables NUTS3 dataset						
Patents	6,800	263.07	133.57	421.77	2.25	5,892
NoChem. Pat.	6,800	219.43	116.46	367.73	0.50	5,458
First-time	6,800	29.32	11.24	57.14	0.00	863
Pats. by ind.	6,800	31.76	13.39	62.73	0.00	1,145
Pats. by orgs.	6,800	226.36	114.54	368.77	1.08	5,523
Forward cit.	6,800	8.65	8.27	5.21	0.23	121
Areas cited	6,800	5.30	5.03	1.64	0.95	32
AM patents	6,800	2.80	0.57	7.78	0.00	168
Utility models	6,592	34.47	10.00	155.62	1.00	3,317
Panel C: explanatory variables FUA dataset						
Population	1,632	573,910	286,198	857,607	88,793	5,281,728
GDP	1,632	30,957	16,558	38,930	2,651	246,981
Panel D: explanatory variables NUTS3 dataset						
Population	6,800	205,005	151,683	233,158	33,944	3,669,491
GDP	6,800	6,860	4,209	10,734	833.83	157,131
Urban areas	6,800	0.44	0.00	0.50	0.00	1.00

Notes: The table presents summary statistics for the dependent variables in the FUA-year and NUTS3-year datasets over the period 2003–2019. GDP is expressed in €1,000.

3.5 Empirical Strategy

3.5.1 Empirical Strategy of the Main Analysis

To identify the effect of FabLabs on patenting activities, this study leverages the staggered rollout of FabLabs in Germany since 2009. In doing so, I follow the literature that analyzes the causal effects of research institutions (Jaffe, 1989; Toivanen & Väänänen, 2016; Pfister et al., 2021; Lehnert et al., 2022) and of knowledge access (Berkes & Nencka, 2020) on innovation activities. I use a Difference-in-Difference (DID) approach to estimate the effect of FabLabs on local innovation activities. Key assumptions for a causal interpretation are parallel trends in treated and control groups and no treatment anticipation.

In recent years, the DID approach with staggered treatment adoption has garnered significant attention in the literature and several limitations of the Two-Way Fixed Effects (TWFE) regression with staggered treatment have been highlighted. de Chaisemartin and D'Haultfœuille (2020) and Goodman-Bacon (2021) underscore the issue of "bad comparisons", where newly treated units are compared to already treated units, leading to distorted causal interpretations and potentially misleading negative estimates. Additionally, TWFE estimates a weighted average of average treatment effects (ATT), but the weights are complex and difficult to interpret. Solutions to these issues have been proposed by Callaway and Sant'Anna (2021), Goodman-Bacon (2021), Sun and Abraham (2021), Wooldridge (2021), and Borusyak, Jaravel, and Spiess (2024).

In the main analyses of this paper, I use the solution proposed by Callaway and Sant'Anna (henceforth CS 2021), which facilitates the calculation of group-time average treatment effects. This is particularly useful for analyzing the effects of FabLabs across the different years of establishment. The group-time average treatment effects can be aggregated into overall treatment effects. The average effect of treatment participation for units in group g is identified by comparing the actual change in outcomes for that group between period $g-1$ and t with the change in outcomes for a comparison group. Assuming parallel trends, this comparison reflects the outcomes that units in group g would have experienced had they not undergone the treatment. In this study, groups g represent regions in which a FabLab was established in a given year. The comparison groups comprise regions that either do not have a FabLab yet or will not have one within the sample period.

Furthermore, the CS approach allows to include pre-treatment covariates to control for non-parallel outcome dynamics due to different observed characteristics, yielding identification under conditional parallel-trend assumptions (Callaway & Sant'Anna, 2021).

To identify group-time average treatment effects (ATTs) non-parametrically using covariates, CS provide three approaches: modeling conditional expectations with regression adjustment (Heckman, Ichimura, & Todd, 1997; Heckman et al., 1998), propensity score modeling for treatment likelihood (Abadie, 2005), and a doubly-robust estimator that combines the first two approaches, requiring only that either the outcome model or the propensity score model is correctly specified. In the study at hand, I employ the doubly-robust estimator, which is also the default of the CS approach.

The group-time average treatment effects are defined by:

$$ATT(g,t) = E\left[Y_t - Y_{g-1} | G_g = g\right] - E\left[Y_t - Y_{g-1} | G_g = \varrho\right] \tag{3.1}$$

where Y_t is the post-treatment outcome, and Y_{g-1} refers to the outcome one period before the treatment occurs for treatment group g, i.e. FUAs or NUTS3 in which a FabLab was established. ϱ represents the comparison group, which can either consist of never treated regions or not-yet treated regions. CS then run the following regressions for each treatment group (conditional on covariates):

$$Y = \alpha_1^{g,t} + \alpha_2^{g,t} \cdot G_g + \alpha_3^{g,t} \cdot 1\{T = t\} + \beta^{g,t} \cdot \left(G_g \cdot 1\{T = t\}\right) + \gamma \cdot X + \epsilon^{g,t} \tag{3.2}$$

where $\beta^{g,t}$ is the average treatment effect on group g at time t, and X represents the covariates.

The approach by CS allows for different aggregation schemes to analyze heterogeneity in treatment effects. CS aggregate the $ATT(g,t)$'s to reflect overall treatment effect heterogeneity over time, resembling the event study design while avoiding the aforementioned pitfalls concerning comparison groups. Additionally, they calculate the cumulative average treatment effects across all groups up to time t and provide the overall average treatment effect across all groups. Using the provided event-study approach, I test for the parallel trends and anticipation effects.

I use the OECD's Functional Urban Areas (FUA) as the level of my main analysis and complement it by the more granular level of NUTS3 regions. For both analysis levels, I include as controls the annual regional population and GPD. Standard errors are clustered at the regional level to allow for arbitrary autocorrelation within the region.

To control for outliers and to obtain the percentage change in the outcome variable, I transform the dependent variables using the natural logarithm after adding one to the value in the regressions.

3.5.2 Empirical Strategy for Analyzing Heterogeneity in the Presence of Pre-Existing Community Workshops

To examine the effects of the pre-existing community workshop ecosystem on FabLab outcomes, I adopt the DiD imputation approach proposed by Borusyak, Jaravel, and Spiess (2024). This method imputes untreated potential outcomes by fitting unit and period fixed effects using untreated observations. Using these imputed untreated outcomes, the weighted treatment effect on the treated is estimated. This approach enhances transparency and provides more accurate treatment effect estimates, particularly when heterogeneity is present (Borusyak, Jaravel, & Spiess, 2024). The error terms are clustered at the regional and year and regional fixed effects are included. I apply this method to analyze the heterogeneous effects resulting from varying pre-existing community workshops ecosystems in a region before the establishment of a FabLab.

The composition of community workshops varies across regions (see Figure 3.3). In an innovation ecosystem, actors interact through both competitive and complementary relations (Asheim & Gertler, 2009; Carayannis & Campbel, 2009). Thus, the coexistence of multiple community workshops in a region can either enhance or hinder FabLabs' effect on a region's innovation potential.

Competition effects may arise if several community workshops compete for users and resources. FabLabs may face competition from incumbent Makerspaces or Hackerspaces, especially if these community workshops have an established user base and are well-integrated within the local community. While FabLabs are differentiated by their focus on digital fabrication, global networks, and standardized tools, it can be difficult to communicate these benefits effectively to potential users. Overlapping tools an machinery may further exacerbate the competition for users. Furthermore, prior research and my user survey underline the importance of the social network within a community workshop (Moilanen, 2012; van Holm, 2017; Cuntz & Peuckert, 2023). This implies that FabLabs face significant network effects and switching costs when competing for users. However, without a user base signalling the demand for a FabLab, newly founded FabLabs may struggle to secure long-term funding.

As well as competing for users, FabLabs also compete for funding from local authorities, educational institutions and firms, which are the main sponsors of German FabLabs (see 3.3.2). While urban areas typically offer more resources, i.e. more firms locate in urban areas, rural areas face resource constraints, which may intensify competition and undermine the sustainability of all workshops in the region. Ultimately, competition may reduce knowledge exchange and innovation if there are fewer or less diverse users of FabLabs, or if, for example, opening hours are reduced due to financial constraints. In extreme cases, competition could stifle the innovation capacity of the whole region by weakening all community workshops.

Conversely, multiple community workshops may complement one another by offering specialized services and increasing regional innovation ecosystem diversity. This dynamic may resemble the positive effects on local innovation capacities observed when multiple research institutions co-exist within a municipality (Lehnert et al., 2022). Moreover, regions with multiple community workshops may generally have a more mature and also more open innovation ecosystem, capable of supporting multiple community workshops and thus conducive to user innovation. The decision to open a FabLab will thus not be entirely arbitrary, but will also depend on regional opportunities.

To analyze these heterogeneous effects, I define four indicator variables. The first, “MS or HS”, equals one if the FabLab was preceded by either a Makerspace or a Hackerspace. The second, “MS & HS”, equals one if both a Makerspace and a Hackerspace existed before the FabLab. As Makerspaces account for the largest number of community workshops in Germany, the third indicator variable, ‘MS‘”, equals one if only a Makerspace existed before the FabLab. To assess the impact of FabLabs alone, the final variable “FL” equals one if the FabLab was the first community workshop space in the region.

The coefficients τ_1 to τ_5 in the regression tables represent the treatment effects at the different time horizons after the FabLab establishment. *Base* denotes the coefficient when the indicator variable equals 0, while "ind.=1" refers to the coefficient when the condition for that column is true. The coefficients $Pre1$ to $Pre4$ capture pre-treatment trends in the dependent variable. All regressions use the log of the dependent variable, controlling for log population and log GDP.

3.6 Results

In this section, I present the results of the empirical strategy outlined in Section 3.5. I evaluate whether the establishment of a FabLab in either a FUA or a NUTS3 region affected inventive activities. I later differentiate by university affiliation, rural areas, as well as existing community workshops prior to the FabLab.

3.6.1 Impact of the FabLabs on Inventive Activity

This section examines the impact of the establishment of FabLabs in FUAs and NUTS3 regions on inventive activity based on the empirical methodology outlined in section 3.5. To control for outliers and to obtain the percentage change in the outcome variable, the outcome variables are log-transformed. Population and GDP per region are included in the analysis as control variables to control for omitted variable bias and selection effects related to the location of the FabLab.

Overall, the average treatment effects on the treated (ATT) are statistically insignificant and small for all the dependent variables (see Table 3.2). Event studies confirm the absence of pre-trends for all dependent variables, and do not indicate any significant long-term effects of FabLabs (see Figure 3.5). The initially negative impact on first-time inventors weakens over time, though the large standard errors suggest considerable variability. In contrast, AM-related patents exhibit a positive long-term trend. Further investigation is needed to determine whether these trends reflect genuine effects or statistical noise.

The ATTs for total patent applications, non-chemical patents, first-time inventors, and areas cited are negative but statistically non-significant, while the ATTs for organizational patents, AM-related patents, utility models, and forward citations are positive yet insignificant (see Table 3.2).

Focusing on individual cohorts, the 2009 cohort exhibits a significant positive effect on the total patent applications, especially by organizational applicants, while non-chemical patent applications are unaffected. Following the first FabLab in 2009, the number of patents by individual applicants declines by 41%, accompanied by a significant decrease in patent quality (measured by forward citations). In the 2015 cohort the first-time inventors significantly increase, while other cohorts show negative yet insignificant effects. Utility model applications increase by 24% in the 2016 cohort. AM-related patents increase significantly only in the 2010 cohort.

In summary, the results show a very mixed picture with hints on overall positive effects on AM related patents and predominantly negative impacts on first-time inventors on the larger, urban level of FUAs. However, this broader regional area might also mask some local impacts of the FabLabs.

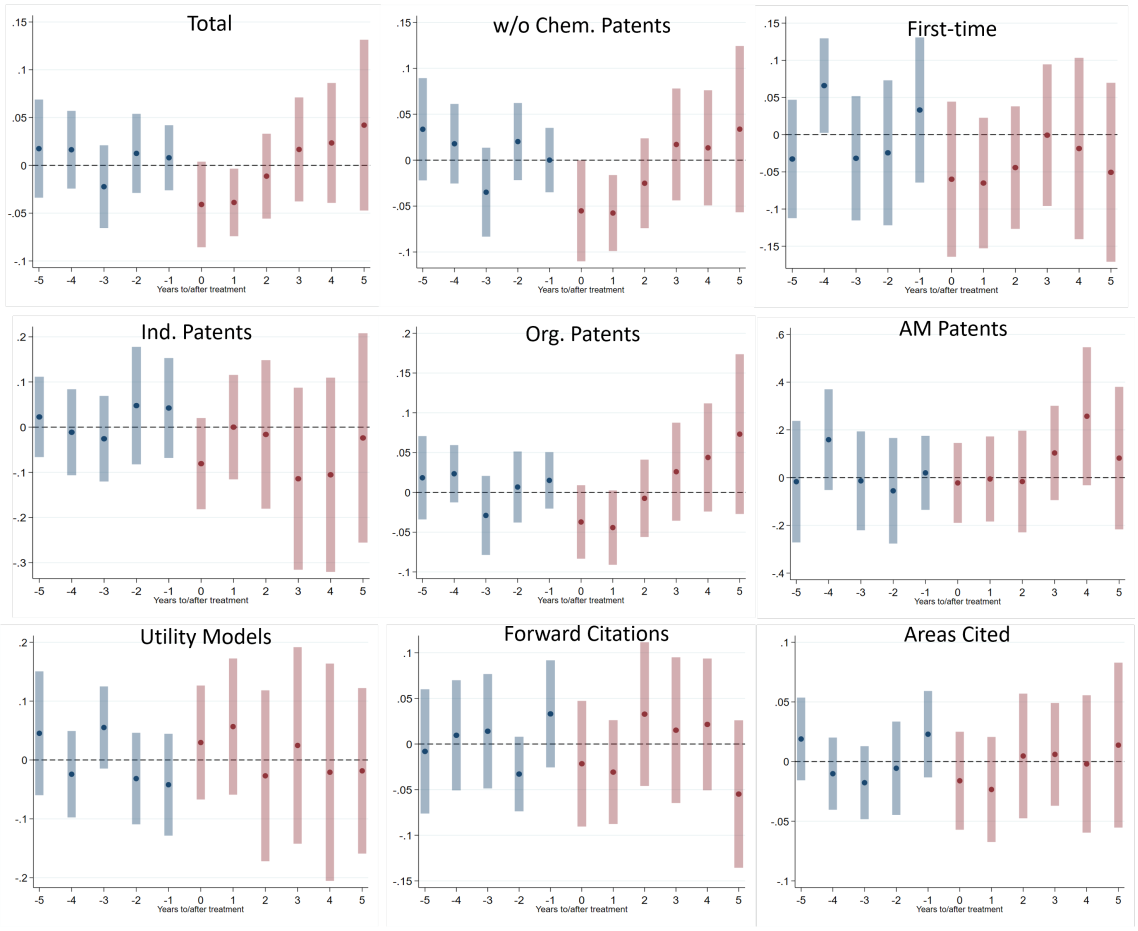


Figure 3.5: Event study estimates for all dependent variables (FUA-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables, focusing on university-integrated FabLabs. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included.

Table 3.2: Impact of FabLabs on dependent variables (FUA-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	-0.00523	-0.0125	-0.0465	-0.0211	0.00498	0.0462	0.00838	0.00602	-0.00490
	(0.0204)	(0.0208)	(0.0401)	(0.0693)	(0.0247)	(0.0805)	(0.0578)	(0.0282)	(0.0193)
G2009	0.0634**	0.00752	0.0147	-0.415***	0.134***	-0.0689	0.00422	-0.0562*	0.0133
	(0.0210)	(0.0214)	(0.0307)	(0.0371)	(0.0203)	(0.112)	(0.0400)	(0.0249)	(0.0127)
G2010	-0.00721	0.00381	-0.106	-0.0802	0.0290	0.138*	0.0660	0.0854**	0.0118
	(0.0530)	(0.0259)	(0.146)	(0.0794)	(0.0565)	(0.0601)	(0.0909)	(0.0329)	(0.0229)
G2011	-0.0121	-0.00871	0.0564	0.157	-0.00376	0.0800	0.0196	0.117**	-0.0292
	(0.0383)	(0.0374)	(0.0491)	(0.0854)	(0.0379)	(0.200)	(0.0994)	(0.0396)	(0.0364)
G2012	-0.0279	-0.0130	-0.160*	0.357	-0.0319	0.153	-0.207	0.00266	0.0130
	(0.0368)	(0.0362)	(0.0738)	(0.267)	(0.0926)	(0.202)	(0.154)	(0.0717)	(0.0485)
G2013	0.0140	-0.00211	-0.130	0.103	0.0250	-0.0616	-0.0375	-0.107	-0.0564
	(0.0368)	(0.0452)	(0.138)	(0.1000)	(0.0498)	(0.206)	(0.137)	(0.0744)	(0.0352)
G2014	0.0260	-0.00432	-0.0857	-0.305*	0.0409	0.211	0.0914	0.0276	0.0834
	(0.0563)	(0.0600)	(0.111)	(0.134)	(0.0611)	(0.173)	(0.199)	(0.0866)	(0.0744)
G2015	-0.000418	-0.00659	0.189**	-0.162*	0.00914	-0.0695	-0.136	-0.0273	0.00201
	(0.0720)	(0.0729)	(0.0643)	(0.0780)	(0.0758)	(0.236)	(0.0889)	(0.0462)	(0.0180)
G2016	-0.00247	0.0111	-0.0674	-0.256*	-0.00169	-0.0257	0.243*	-0.0214	-0.0158
	(0.0459)	(0.0475)	(0.0730)	(0.128)	(0.0474)	(0.167)	(0.102)	(0.0636)	(0.0375)
G2017	-0.0846*	-0.0796	-0.155	-0.0270	-0.0827*	-0.0822	-0.143	0.0340	-0.0144
	(0.0417)	(0.0465)	(0.130)	(0.164)	(0.0372)	(0.158)	(0.134)	(0.0333)	(0.0273)
G2018	-0.193***	-0.308***	0.100	0.180	-0.274***	-0.0103	0.00988	-0.219	-0.102
	(0.0606)	(0.0677)	(0.139)	(0.185)	(0.0523)	(0.192)	(0.125)	(0.150)	(0.165)
N	1,632	1,632	1,632	1,632	1,632	1,632	1,632	1,632	1,632
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables. The unit of analysis is the FUA-year, with not-yet-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

To capture more localized effects, I analyzed the effect of FabLabs on the NUTS3 level. Figure 3.6 presents event studies for all dependent variables. The pre-treatment effects are non-significant across variables except for first-time inventors for which a slightly positive and significant trend in period *t-4* suggests careful causal interpretation. Compared to the FUA-level analysis, the ATTs are larger, suggesting that FabLabs' localized effects may be masked in broader FUA-level analyses (Table 3.3). Consistent with the FUA-level findings, FabLab establishment in NUTS3 regions has an overall negative but statistically insignificant impact on total patent applications, non-chemical patent applications, and first-time inventors. The positive yet insignificant effects on

AM-related patents and forward citations persist with larger ATT estimates than at the FUA level. However, organizational patents and utility model applications, which exhibit small but positive ATTs at the FUA level, show a reversed effect at the more granular NUTS3 level. Conversely, individual patents and cited technological areas, negatively correlated with FabLab introduction at the FUA level, display positive but insignificant ATTs at the NUTS3 level.

These differences may be due to the different aggregation levels. At the FUA level, the aggregated data and broader regional trends may obscure localized effects, for example in patents of individual applicants. Furthermore, the NUTS3 analysis captures more local heterogeneity in innovation dynamics, since NUTS3 regions include both urban and rural areas, while FUAs focus exclusively on urban regions. The greater variation at the NUTS3 level may reflect differing in regional innovation ecosystems, including absorptive capacity, industrial specialization, or local policy environments.

The event studies reveal a lagged, significantly negative effect on non-chemical patents four years after FabLab establishment (see Figure 3.6), whereas AM-related patents exhibit an increasing positive long-term trend. Forward citations and cited areas also increase over time, with the latter showing even a statistically significant increase. This observation suggests that the impact of FabLabs on NUTS3 regions builds up over time, indicating a positive influence on the patent quality indicators and the number of AM-related patents.

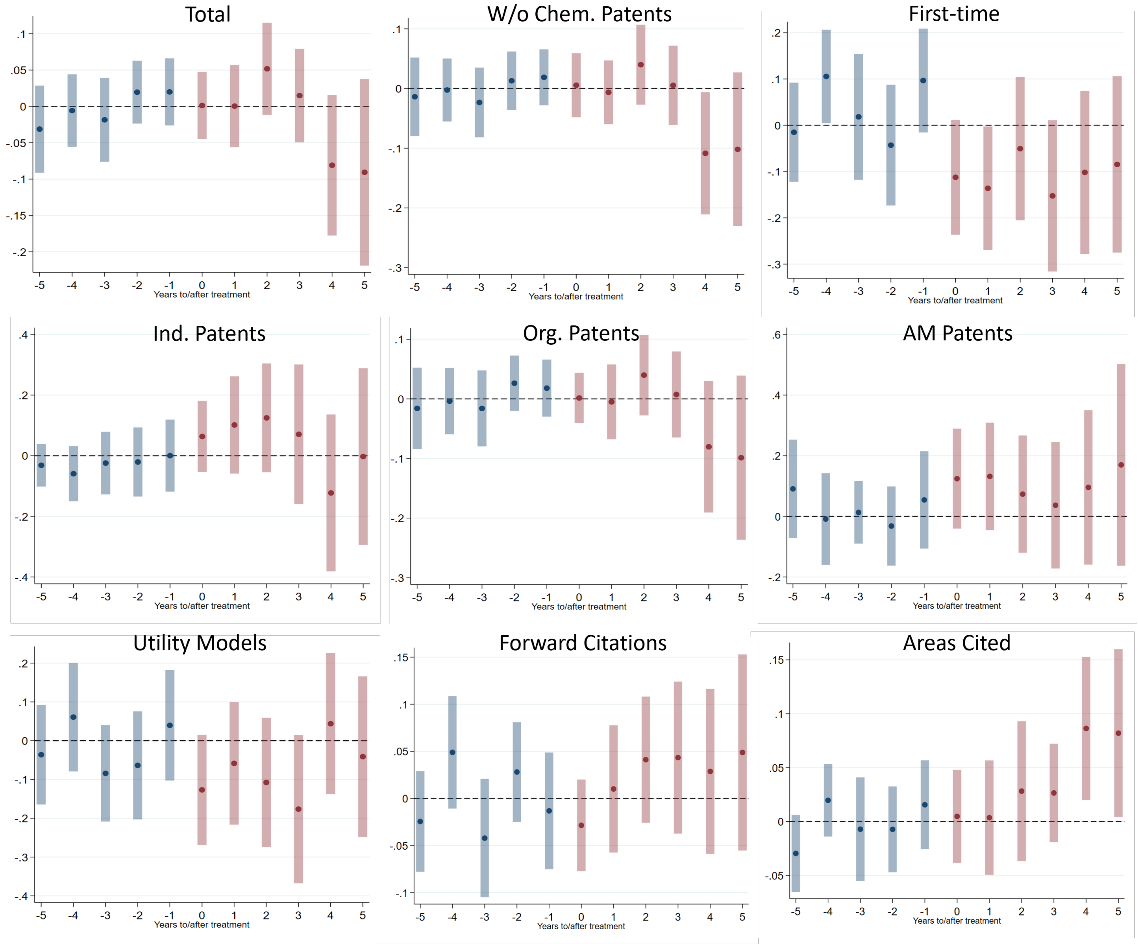


Figure 3.6: Event study estimates for all dependent variables (NUTS3-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included.

As at the FUA level, the first FabLab established in 2009 has the most notable impact at the NUTS3 level. However, unlike at the FUA level, the first FabLab significantly reduces the number of total patent applications, non-chemical patent applications, individual patents, organizational patents, utility models, forward citations, and areas cited. In contrast, following the first FabLab first-time inventors and AM-related patents increase significantly, while these coefficients were insignificant at the FUA-level. These findings underscore the importance of analyzing at the NUTS3 level to capture localized effects. Cohort 2010 also reveals positive and statistically significant effects on these variables. In contrast, later cohorts display negative effects on first-time inventors and positive but insignificant effects on AM-related patents. Of particular interest is the observation that across cohorts FabLabs increase patent quality indicators especially in those of interdisciplinarity. This suggests that, while the overall quantity of patents may not undergo a substantial increase, the quality, appears to be enhanced post-FabLabs.

In general, the results suggest that the impact of FabLabs varies over time. Earlier cohorts (e.g., 2009) experienced declines in the number of total patents, non-chemical patents, individual patent applications, quality measures, and utility models at the NUTS3 level. However, starting with the 2014 cohort, these coefficients transitioned to positive values, indicating an increase in the number of patents, utility models, and quality measures. From the 2016 cohort onward, the number of patents increased post-FabLab, while utility models declined.

On the one hand, this pattern may reflect the benefits of the growing FabLab network in fostering knowledge exchange, supported by the FabFoundation. Another potential explanation is regional differences in the integration of the FabLab within the local community and innovation ecosystem. The positive and partly significant effects on the organizational patents indicate that firms increasingly use the FabLabs for ideation and prototyping, which is consistent with the online survey results that firms and startups frequently use FabLabs.

The predominantly negative yet insignificant effect on first-time inventors may be indicative of the observation that users primarily work on personal projects and that FabLabs foster a culture of free knowledge sharing. Although the FabCharter does not prevent commercialization, these values may discourage FabLab users from patenting their ideas (Gertler, 2003). However, this hypothesis appears contradicted by the positive significant effects on both first-time inventors and AM-related patents especially in the first two cohorts. A more plausible explanation may be that the positive effects need more time to develop. In addition, not all creations are patentable, and not every invention is actually patented (Franke & Lüthje, 2020; Clancy, 2023). Furthermore, regional differences in absorptive capacity and institutional support may also affect first-time inventor trends (Asheim & Gertler, 2009; Carayannis & Campbel, 2009; Granstrand & Holgersson, 2020).

FabLab exhibit mixed and localized effects on regional innovative activities at both the FUA and NUTS3 levels. While both levels show generally negative, statistically insignificant effects on total patents, non-chemical patents, and first-time inventors, the effects are larger at the NUTS3 level. Notably, the more pronounced increase in AM-related patents and forward citations at the NUTS3 level suggests a stronger localized benefit on these inventive activities.

Differences between the two levels of analysis underscore the importance to consider different aggregation levels. The shift from negative to positive effects on patents by individual applicants and interdisciplinary patents at the NUTS3 level, suggests that the granular and regional analysis can uncover localized innovation dynamics which are not evident at the broader FUA level. These findings are consistent with the literature stating that users are more likely to innovate if knowledge and tools are locally close by (Svensson & Hartmann, 2018; Maravilhas & Martins, 2019; Audretsch & Belitski, 2020).

The decline in utility model applications across cohorts and in both NUTS3 and FUA regions suggests a shift in the nature of innovation, possibly reflecting the broader decrease in utility model use in Germany rather than a FabLab-specific effect. Overall, the results suggest a complex impact of FabLabs over time, with varying effects influenced by cohort-specific factors or innovation ecosystems. To explore these varying effects in greater detail, the next section presents an analysis of heterogeneous impacts.

Table 3.3: Impact of FabLabs on dependent variables (NUTS3-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	-0.0111	-0.0189	-0.0851	0.0188	-0.0143	0.143	-0.0936	0.0154	0.0254
	(0.0296)	(0.0305)	(0.0540)	(0.0830)	(0.0340)	(0.0810)	(0.0576)	(0.0304)	(0.0219)
G2009	-0.459***	-0.476***	0.196***	-1.045***	-0.461***	0.413***	-0.265***	-0.326***	-0.253***
	(0.0227)	(0.0226)	(0.0294)	(0.0467)	(0.0226)	(0.0364)	(0.0613)	(0.0185)	(0.0156)
G2010	0.0588*	0.0303	0.139**	-0.204*	0.132***	0.871**	-0.0470	0.0986	-0.00348
	(0.0254)	(0.0316)	(0.0520)	(0.0946)	(0.0256)	(0.312)	(0.101)	(0.0876)	(0.0271)
G2011	-0.0252	0.0198	-0.144	0.177	-0.0521	0.262	-0.228*	-0.0657	0.0138
	(0.0604)	(0.0676)	(0.141)	(0.297)	(0.0930)	(0.197)	(0.113)	(0.0949)	(0.0338)
G2012	-0.0699	-0.101	-0.129	-0.169***	-0.0509	0.0779	-0.172	0.0141	-0.0463
	(0.0965)	(0.0827)	(0.0741)	(0.0453)	(0.0874)	(0.112)	(0.176)	(0.0505)	(0.0730)
G2013	-0.0978	-0.149*	-0.0880	-0.0580	-0.0897	-0.197	0.0424	0.0944	0.143*
	(0.0780)	(0.0694)	(0.193)	(0.181)	(0.102)	(0.224)	(0.154)	(0.105)	(0.0572)
G2014	0.0251	0.00177	-0.194	-0.00350	0.0113	0.00445	0.0398	0.0320	0.0911*
	(0.0487)	(0.0481)	(0.108)	(0.158)	(0.0557)	(0.125)	(0.140)	(0.0372)	(0.0438)
G2015	0.0355	0.0304	0.0797	0.0331	0.0270	0.352	0.0656	0.0520	0.0538**
	(0.0595)	(0.0547)	(0.207)	(0.180)	(0.0525)	(0.204)	(0.209)	(0.0602)	(0.0185)
G2016	0.116*	0.132*	-0.142	0.144	0.116*	0.00590	-0.0633	0.0801	-0.0833
	(0.0545)	(0.0626)	(0.124)	(0.132)	(0.0585)	(0.212)	(0.101)	(0.0548)	(0.0456)
G2017	0.0836**	0.0889***	0.0554	0.695***	0.0431	0.312	-0.854***	-0.0253	0.0416
	(0.0290)	(0.0269)	(0.142)	(0.162)	(0.0310)	(0.291)	(0.250)	(0.0733)	(0.0605)
G2018	-0.0607	-0.0567	-0.144	0.147	-0.0625	0.335*	-0.0845	-0.0593	0.000358
	(0.107)	(0.113)	(0.0980)	(0.196)	(0.113)	(0.157)	(0.196)	(0.0673)	(0.0428)
N	6,800	6,800	6,800	6,800	6,800	6,800	6,800	6,800	6,800
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables. The unit of analysis is the NUTS3-year, with never and not-yet-treated observations serving as the control group. Population and GDP are included as control variables., and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.6.2 Heterogeneous Effects

To assess the impact of FabLabs on local inventive activity, this analysis examines heterogeneity across key dimensions. Specifically, it evaluates the effects of university-integrated FabLabs, FabLabs' influence in regions with below- and above-average patenting activity, and differential impacts across urban and rural NUTS3 regions and within varying community workshop ecosystems.

Heterogeneous Effects of University-Integrated FabLabs

Universities play a central role in innovation ecosystems by fostering knowledge creation and diffusion (Jaffe, 1989; Carayannis & Campbel, 2009; Audretsch, Hülsbeck, & Lehmann, 2012). By providing access to advanced tools and facilitating interdisciplinary collaboration, FabLabs can amplify these effects. Therefore, this analysis focuses exclusively on FabLabs established within universities and excludes those established outside universities. As in the main analyses, the control group consists of never and not-yet treated regions, and population and GDP are used as covariates.

University-integrated FabLabs were established in FUAs in 2009, 2012, 2013, 2014, and 2018 and are associated with an an overall significant 22% decline in first-time applicants (see Table C.3). Three out of five cohorts exhibit negative coefficients for first-time inventors, two of which are statistically significant. While most of the other overall ATTs are negative and insignificant, AM-related patents and patents by individual applicants exhibit small but statistically insignificant positive effects.

The event study analysis shows significant pre-trends for AM-related patents, utility models, and the number of areas cited per patent three to four years before treatment (Figure C.4). These pre-trends indicate potential limitations in statistical power due to the small number of treated units. Post-treatment, first-time inventors experience a sustained decline, while patents by individual applicants exhibit short-term positive effects. Non-chemical patents and organizational patents initially decline but subsequently recover to positive, albeit statistically insignificant, coefficients.

At NUTS3 level, the impact of university-integrated FabLabs differs from the FUA-level trends (see Figure C.5). While non-chemical patenting exhibited a positive trend at the FUA level, they decline significantly five years post-treatment at the NUTS3 level. Similarly, utility model applications experience a significant medium-term decline at the NUTS3 level, which partially recovers over time. The sharp initial drop in first-time inventors post-FabLabs may reflect prospective first-time inventors postpone patent applications to develop their ideas further using the newly available FabLab resources.

The overall negative effect in regions with university-integrated FabLabs may partly be due to a general decline in university patenting following Germany's 2002 policy reform that transferred intellectual property rights from individual professors to universities (see Czarnitzki et al. (2017) for details). Another explanation for the negative impact on first-time inventors and patenting at large could be that university-integrated FabLabs are predominantly used by students, who use the facilities for skills acquisition and academic projects rather than commercial innovation. A descriptive analysis of FabLab users in Germany could provide further insights, but is beyond the scope of this study.

Despite these declines, university-integrated FabLabs appear to stimulate innovation in complex technological areas. AM-related patents exhibit stronger effects compared to the full NUTS3 sample, and interdisciplinary patents—measured by the number of areas cited—show a significant increase over time, albeit with larger standard errors. The 16% decline in utility model applications may reflect stronger institutional support during the patenting process provided by the university. This reinforces the importance of interactions with innovation ecosystems in shaping innovations (Asheim & Gertler, 2009; Peschl & Fundneider, 2012).

Compared to the full NUTS3 analysis (see Table 3.3), university-integrated FabLabs have a stronger negative effect on first-time inventors, patents by individual applicants, organizational patents, utility models, and forward citations. However, AM-related patents and interdisciplinary patents see stronger positive effects in the university subsample, suggesting that FabLabs embedded in universities may act as catalysts for innovation in complex technological fields.

Cohort-level results further support this interpretation. The first German university-integrated FabLab (cohort 2009) exhibits significant increases in first-time inventors and AM-related patents, suggesting that the early FabLabs enabled tinkerers to become inventors. However, subsequent cohorts do not demonstrate these positive effects on first-time inventors, but rather exhibit overall negative coefficients. Nonetheless, these coefficients are mitigated compared to the full sample, suggesting that university-integrated FabLabs exert a moderating influence on the overall negative trend. Furthermore, university-integrated FabLabs increase the interdisciplinarity of patents across cohorts.

The impact of university-integrated FabLabs on inventive activity is twofold. While they are associated with a decline in total patenting, first-time inventors, and utility models, they also foster AM-related patents and interdisciplinary innovation. These results suggest that university FabLabs primarily foster innovation in complex technological domains, while their integration into university ecosystems shifts the nature of inventive output away from first-step innovations.

Heterogeneous Effects of Regions with Patenting Activity Above or Below Average

Following the analysis of university-integrated FabLabs, this section examines whether the regional patenting activity influences Highly innovative regions—those with more advanced innovation ecosystems may affect results differently than less innovative ones (Audretsch, Hülsbeck, & Lehmann, 2012). To address this, I analyze FabLabs' effects separately for regions with patenting activity below and above the median prior to the establishment of the first FabLab in 2009.

At the FUA level, Table C.5 shows that FabLabs in regions with below-median patenting activity yield larger overall ATTs but with greater standard errors. The signs mostly align with the full sample, except for organizational patents and the forward citations, which turn negative in the below-median sample. First-time inventors show more pronounced negative effects compared to the full sample. In below-median patenting regions, FabLabs seem to benefit individuals who are familiar with patenting rather than attracting new inventors.

Event studies show no significant pre-trends (see Figure C.6), but reveal a persistent negative trend for first-time inventors, that is larger than in the full sample. Organizational patents decline significantly in the short-term but recover over time, while AM-related patents and utility models show positive yet insignificant trends accompanied by large standard errors.

In above-median regions, FabLabs have predominantly negative and insignificant overall effects (Table C.7), except for organizational patents which exhibit positive effects. The event studies indicate significant pre-trends in AM-related patents and forward citations, suggesting that the observed effects may partly be due to underlying dynamics rather than FabLabs' establishment (Figure C.8). First-time inventors are less affected, and patents by individual applicants decline.

At the NUTS3 level, the results are more differentiated. In below-median regions (see Table C.6), FabLabs significantly reduce the number of first-time inventors, whereas in above-median regions (Table C.8), the overall effect is positive but not statistically significant. Similar to the FUA analysis, individual applicants benefit more from FabLabs in NUTS3 regions with below-median patenting activity, especially in later cohorts, while organizational patents show significantly greater increases in above-median regions. AM-related patents are significantly and positively affected by FabLabs in above-median regions, with this effect persisting across cohorts. Utility models decline in both samples, but the decrease is more pronounced and partly significant in below-median regions. Patent quality, as measured by forward citations, improves more in above-median regions, with the ATT for forward citations being twice as large as in the full sample and consistently positive across cohorts.

The event studies for below-median regions indicate a positive anticipation effect for the first-time inventors, which moderates the significant post-treatment decline (see Figure C.7). One possible explanation is that potential first-time inventors accelerate patenting before engaging with FabLabs. This may be due to concerns about intellectual property security in an open-innovation environment. The subsequent decline in first-time inventors may then reflect a shift towards experimentation and skill acquisition within FabLabs, delaying patenting activity. This aligns with the idea that FabLabs facilitate learning and prototyping before innovations reach a patentable stage. For all other variables, no significant pre-trends exist. Non-chemical and organizational patents decline from the fourth year post-FabLab onward, while patents by individual applicants and cited areas increase in the long term. In the above-median NUTS3 sample, first-time inventors experience a significant positive trend (Figure C.9). AM-related patents also increase, although large standard errors result in less precise estimates. Utility models exhibit negative trends which appear to recover over time.

These findings highlight how pre-existing local innovation environments shape FabLabs' impact. In regions with a high concentration of first-time inventors before FabLabs were established — indicating a more supportive ecosystem for new inventors — FabLabs are more effective in enabling first-time inventors. Conversely, in regions with fewer inventors, FabLabs benefit those already familiar with patenting, negatively affecting new inventors. Regions with stronger AM-related patenting see an additional boost to this technology and patent quality indicators post-FabLab establishment. These results highlight FabLabs' role as enabling spaces for experimentation and skill-building, with their impact varying depending on regional innovation dynamics. To further refine these insights, the next section examines how FabLabs influence inventive activity in rural versus urban settings.

Heterogeneous Effects of Rural and Urban NUTS3 regions

Building on the analysis of patenting intensity, I now examine whether FabLabs exert a differential impact in rural versus urban regions. This differentiation is possible only at the NUTS3 level, as FUA regions are, by definition, urban.

A comparison of urban and rural areas reveals that FabLabs have a stronger and more positive impact in urban settings. The number of first-time inventors increases 4% in urban areas, albeit insignificantly, while decline significantly in rural areas by 17% (see Tables C.9 and C.10). This pattern reverses for patents by individual applicants, with rural areas being positively and urban areas negatively affected, though neither coefficient is statistically significant. Organizational patents statistically insignificantly increase in urban areas, while rural areas experience a decline, reflecting the concentration of firms in urban regions (Argente et al., 2020). These trends are consistent with earlier findings that FabLabs benefit individual applicants in below-median patenting regions more, while fostering first-time inventors in highly active regions.

FabLabs positively influence AM-related patents in both settings, though the effect is stronger in urban areas, which experience a significant increase of 31%, compared to a statistically insignificant 5% increase in rural areas (see Tables C.9 and C.10). This suggests that FabLabs are more effective in fostering AM-related innovation in urban regions, where supporting infrastructure and expertise already exist. Utility models decline in both urban and rural regions, but for rural areas the effect is more pronounced. This suggests that FabLabs in rural settings are less effective in stimulating utility model applications, possibly due to differences in technological specialization or commercialization pathways.

Patent quality, as measured by forward citations, also exhibits an urban-rural divide. In urban areas the number of forward citations increase significantly, while the effect is negative and insignificant in rural areas. Both urban and rural areas show positive but non-significant effects on the cited areas, with the effect being larger in magnitude for urban regions, suggesting a slightly stronger influence of FabLabs on the breadth of citations in urban regions.

In the long run, rural areas experience a persistent decline in non-chemical and organizational patents, starting four years after the establishment of FabLabs (see Figure C.11). First-time inventors follow a negative trend that persists over time. Individual applicants benefit especially in the first three years, although the effect is less clear in later years due to large standard errors. The impact on AM-related patents and cited areas are weaker than in the full sample.

Urban areas, by contrast, display more consistent positive long-term effects (Figure C.10). Despite a negative pre-treatment trend, non-chemical and organizational patents increase post-treatment, although statistically insignificant. First-time inventors follow a wave-like pattern post-treatment, with almost significant increases every three years. AM-related patents increase significantly over time post-treatment, though large standard errors suggest small sample size limitations. Utility models and patent quality indicators also trend positively in urban areas.

A more granular cohort analysis within urban areas reveals that early FabLab cohorts negatively affect non-chemical patents, but later cohorts experience significant positive effects (Table C.9). A similar pattern is observed for patents by individual applicants. First-time inventors in urban regions exhibit significantly positive impacts across cohorts, reinforcing the notion that FabLabs encourage tinkerers to become inventors. Organizational patents experience mixed effects, but generally benefit from FabLabs in urban areas. AM-related patents consistently increase across cohorts, with most increases being statistically significant. Utility models increase in cohorts from 2012 to 2014. However, subsequent cohorts suggest a shift in focus towards patent applications, particularly within AM technology. Patent quality overall improves, as evidenced by the increases in forward citations and the breadth of cited areas.

In summary, FabLabs have a stronger and more positive impact in urban regions, particularly for first-time inventors, AM-related patents, and patent quality indicators. In contrast, rural regions show more limited and often negative effects. These findings align with Cuntz and Peuckert (2023), who show that urban Hackerspaces foster more digital startups in Germany than rural Hackerspaces. The present heterogeneity analysis also helps explain in part why previous FUA-level analyses yielded more positive results than those at the NUTS3 level, as urban regions mainly drive FUA-level results.

One possible explanation why urban areas benefit more from FabLabs is be the presence of a more developed innovation ecosystem. This offers enhanced access to resources, commercialization opportunities, and knowledge spillovers. In addition, existing community workshops may have fostered a culture of knowledge sharing and user innovation, further enhancing FabLabs' effectiveness. Prior results already indicated that pre-existing innovation ecosystem influence FabLabs' impact. The next section examines this further by analyzing how the presence of community workshops prior to FabLab establishment influences their impact on local inventive activities.

Heterogeneous effects of the community workshop ecosystem

The previous analysis found a greater impact of FabLabs in urban areas. However, in rural areas, other existing community workshops may mask the true effect of FabLabs. Therefore, the following analysis distinguishes between urban and rural areas at the NUTS3 level. Broader FUA-level results are also discussed. The analysis will follow the empirical strategy outlined in section 3.5.2.

The heterogeneity analysis at the urban NUTS3 level reveals distinct patterns, especially for first-time inventors and AM related patents. However, the total number of patents, non-chemical patents, and the patents by organizational applicants exhibit significantly negative pre-trends one period before treatment (see Tables C.16-C.19).[27] These negative anticipation effects could reflect innovators delaying patenting to refine their invention in the new FabLab.

Table C.16 demonstrates that in regions with both a Makerspace and a Hackerspace, FabLabs increase the total number of patents in the long term. The initial significant drop in applications indicates an adaptation period for both users and FabLabs. However, the subsequent positive and significant effect supports the hypothesis that pre-trends reflect delayed patenting. Further, it suggests that FabLabs address previously unmet demand for specialized innovation infrastructure. The pre-existing community workshops likely reflect a supportive ecosystem conducive to user-driven innovation.

In contrast, regions with only a Makerspace or Hackerspace experience a persistent negative long-term trend after FabLab entry. In Makerspace-only regions, FabLabs' immediate and medium-term effects are negative, but show a slightly positive long-term trend. This pattern suggests competition for users, especially if the new FabLab and the incumbent Makerspace offer similar tools, making it difficult for the FabLab to differentiate itself. When FabLabs are the first community workshop, initial effects are negative but turn positive over time. This suggests that FabLabs face challenges integrating into the local community and building a diverse user base in regions with limited experience in user-driven innovation or collaborative knowledge sharing. For non-chemical and organizational patents, the results mirror those for the total number of applications (see Tables C.16 and C.18). At the broader FUA level, the significant positive effects in urban NUTS3 regions with both Hackerspaces and Makerspaces are moderated to insignificant negative levels (see Table C.11). This attenuation of effects may be due to averaging over the broader FUA-level, potentially masking local effects.

For rural areas there is suggestive evidence, that FabLabs positively influence non-chemical patents applications in regions where either a Makerspaces or both Makerspace and Hackerspace existed prior to the FabLab. In regions with both Hacker- and Makerspace, the effects are inconsistent but suggest a potential long-term positive effect. When FabLabs are the first community workshop, short-term effects are

[27] A formal test confirms that the pre-treatment coefficients are significantly different from zero, warranting cautious causal interpretation of these results.

slightly positive, but long-term effects turn negative. However, large standard errors in rural settings may mask the actual effects. These findings underline FabLabs' potential in enabling innovation, particularly in rural areas, when embedded within an ecosystem conducive to open and user-driven innovation (see Figure C.13).

Table C.17 examines first-time inventor applications at the urban NUTS3 level and all pre-trends are statistically insignificant. The most substantial gains from FabLabs occur in regions with both a Makerspace and a Hackerspace, where FabLabs create synergies that boost first-time inventors at t+2 and in the long-term. These synergies likely stem from an established user-driven innovation culture, enhanced commercialization opportunities, and resources that support an additional community workshop. Moreover, the results suggest that potential first-time inventors benefit specifically from the tools and machines offered by the newly established FabLab.

Conversely, regions with only a pre-existing Makerspace experience a negative impact, possibly due to competition for resources and users between the FabLab and the incumbent Makerspace, especially if tools offered overlap. When FabLabs are the first community workshop, first-time inventor applications increase significantly at t+2. Despite a decline at t+3, the coefficients return to positive levels thereafter. This temporary drop may reflect the time required for prototype development, skill acquisition, and patent filings. These findings are reflected at the broader FUA-level (see Table C.12).

In rural areas, FabLabs most consistently benefit first-time inventors when they are the first community workshop (see Figure C.14). While regions with both a Hackerspace and a Makerspace see increases, regions with only one of these workshops exhibit no measurable impact. Positive anticipation effects one period before the FabLab establishment suggest either limited statistical power or local expectations before the FabLab opens. Despite the overall negative effect of FabLabs on first-time inventors in rural areas (see section 3.6.2, the present findings indicate that FabLabs can enable tinkerers to become inventors in rural areas when they serve as the first community workshop in previously underserved regions.

Analyzing individual patenting, regions with either a Makerspace or Hackerspace prior to the FabLab (Table C.17) exhibit an initial significant decline that recovers by t+3 but turns negative again in the long-term. A similar pattern emerges when only a Makerspace precedes the FabLab. These findings might point to competition effects during the early post-treatment phase. In contrast, a positive effect is observed when both an Makerspace and an Hackerspace precede the FabLab. However, this benefit diminishes in the long-term. Where FabLabs are the first community workshop, initial effects are insignificant but turn positive in the long term, likely reflecting FabLabs' integration challenges. These results align with the broader FUA-level findings (see Table C.12).

Overall, the presence of prior community workshops often results in initial negative effects on patent applications, followed by partial recovery over time. In contrast, FabLabs established as the first community workshop in a region exhibit initial positive effects that diminish in the short-run but rebound in the long run, indicating a potential lag in the innovation response.

In rural areas, FabLabs have positive long-term effects, especially if a Makerspace existed prior to the FabLab (see Figure C.15). This suggests a complementarity between these community workshops likely attributable to available local resources, an engaged pool of potential inventors, and a user innovation culture that facilitates FabLabs' integration. When FabLab are the first community workshop, the outcomes align with urban regions, exhibiting an initial increase that subsequently declines but recovers over time.

In urban regions, the effect of FabLabs on AM-related patents shows no pre-trends, though the presence of large standard errors indicate limited statistical power (Table C.18). The presence of a single prior workshop (either an Makerspace or Hackerspace) generates an immediate positive effect, which fades or turns negative by t+3, particularly in regions with only a Makerspace. But in the long run, the trends turn positive, suggesting that FabLabs initially fulfill unmet demand for digital fabrication tools but face integration and adaptation challenges in the subsequent years. The more pronounced negative effect in regions with only a Makerspace preceding the FabLab could additionally reflect short-term competition effects for users and resources.

In contrast, FabLabs in regions with both a Makerspace and a Hackerspace, or in regions without prior workshops, exhibit persistent positive trends. Established community workshops likely contribute to an ecosystem that facilitates FabLabs' integration within the local community.In regions without prior community workshops, FabLabs serve as catalysts for AM innovation. The positive effects persist at the FUA level, underscoring FabLabs' potential to foster innovation in complex technologies (see Table C.13).

In rural areas, FabLabs also increase AM-related patent applications, especially where both Maker- and Hackerspaces existed prior to the FabLab (see Figure C.17). These findings emphasize the importance of established innovation ecosystems for the development of complex technologies in rural areas and highlight FabLabs' potential to cater to previously unmet demand for AM, thereby enhance the support for complex technologies. When FabLabs are the first community workshop, the effects are positive but small and insignificant, further indicating that AM technologies require a robust ecosystem.

FabLabs weakly increase utility model applications in urban areas, particularly in regions with prior community workshops (see Table C.19). The presence of several community workshops suggests an established innovation ecosystem that facilitates commercialization opportunities through firms or stronger startup support mechanisms (i.e., incubators and accelerators), which in turn drive utility model filings. However, when FabLabs are the first community workshop, the overall negative trend persists (see Figure C.2), indicating that regions without pre-existing community workshops, and thus a weaker innovation ecosystem, may exhibit less strategical patenting behaviour, opting instead for more conventional patenting practices. At the FUA-level, the overall trend is negative, although a positive short-term is observed when FabLabs serve as the first community workshop (see Table C.14). In rural areas, FabLabs' impact on utility models is predominantly negative, suggesting that FabLabs are more effective in boosting utility model applications when integrated into regions with complementary infrastructure and resources (see Figure C.18).

The quality measures are both predominantly positively affected by FabLabs in the different urban specifications. Forward citations increase significantly in areas where only a Makerspace preceded the FabLab (see Table C.20). This suggests that FabLabs complement Makerspaces by providing advanced tools and fostering knowledge exchange. In contrast, in regions with both a Makerspace and a Hackerspace FabLabs have no measurable impact. In these settings, the incumbent workshops may already offer digital fabrication tools alongside established robust knowledge sharing mechanisms, limiting FabLabs' ability to make an immediate contribution to the number of forward citations. In regions without any community workshops FabLabs appear to act as catalysts for impactful patents by providing an infrastructure that enables local inventors to develop their ideas. In rural areas, there is suggestive evidence that forward citations increase post-FabLabs across all specifications (see Figure C.19). This indicates that in rural areas, FabLabs, with their standardized tools and knowledge-sharing culture, are more effective in fostering impactful innovations than other types of community workshops.

The cited areas exhibit significant pre-trends in the urban specification, which warrants caution in interpreting causality. The results provide suggestive evidence that interdisciplinarity increases across specification, with the most pronounced impact observed in areas where both Makerspace and Hackerspace preceded the FabLab (see Table C.20, Panel (b)).At the broader FUA-level, the predominantly positive impacts of FabLabs persist but are very small (see Table C.15, columns (5)-(8)). In rural areas, while no significant pre-trends are observed, the results are less consistent than in urban settings, likely due to limited statistical power. Early periods exhibit mixed and sometimes negative trends across specifications, but a positive long-term trend emerges (see Figure C.20).

The heterogeneity analysis at the urban NUTS3 level reveals that FabLabs' impacts vary significantly depending on the characteristics of the pre-existing innovation ecosystem. Urban regions with both a Makerspace and a Hackerspace benefit the most, experiencing long-term increases in patenting activity, first-time applications, and AM innovation. These positive effects are likely driven by synergy effects within the established community workshop ecosystem, including funding resources, user engagement, commercialization opportunities, and FabLabs' ability to address previously unmet demand for digital fabrication tools. Moreover, such established ecosystems enable FabLabs to contribute to an existing culture of user-driven innovation and resource-sharing, fostering both patent quantity and quality.

In urban regions without prior community workshops, FabLabs generate long-term positive impacts, particularly on first-time inventors, AM-related patents and forward citations. However, they face initial adaptation challenges, especially in utility model applications. In contrast, urban regions with only a Makerspace, FabLabs initially compete for users and resources, leading to crowding out effects. Over time, however, their local integration fosters particularly AM-related patents, first-time inventors, and patents by individual applicants.

FabLabs in rural areas have the most significant impact when they are the first community workshop, providing an innovation infrastructure and culture that enables innovation, particularly for first-time inventors. However, in these regions FabLabs also face prolonged integration times, likely due to limited user bases and weaker pre-existing innovation ecosystems, which delay FabLabs' ability to foster innovation. In rural regions with any type of incumbent community workshop, FabLab predominantly yield positive impacts, although larger standard errors indicate limited statistical power.

In summary, urban settings demonstrate stronger and faster impacts, while rural settings reveal the transformative potential of FabLabs in filling gaps in innovation infrastructure, though impacts are delayed and less consistent.

3.6.3 Robustness Checks

To ensure that the results are not influenced by the decline in the number of patents, first-time inventors, or utility models in 2008, I conducted a robustness check by excluding the years 2003-2009 from the analysis. Tables C.21 and C.22 present the results for FUA and NUTS3 regions, respectively.

For FUA regions, the overall ATTs are insignificant across variables, with minimal differences from the main specification in Table 3.2. The ATTs for first-time inventors and AM-related patents persist, while those for utility models and quality measures decrease in size. Notably, the overall ATT for individual patents turns positive in the robustness analysis, though the effect size remains very small. At the cohort level, the effects remain unchanged.

In NUTS3 regions, the overall ATTs are also insignificant (see Table C.22) and similar in sign to the main specification in Table 3.3. The restricted analysis shows larger ATTs for first-time inventors, patents by individual applicants, and quality measures. The remaining ATTs are smaller in magnitude. As with FUA regions, the cohort-level effects are unchanged.

The robustness check, excluding 2003-2009, reveals that the overall ATTs remain insignificant for both FUA and NUTS3 regions, with minimal changes compared to the main analysis. Notable variations include a small positive shift in individual patents at the FUA level and larger ATTs for first-time inventors and quality measures in NUTS3 regions.

3.7 Conclusion

In times where innovation is becoming increasingly complex and knowledge creation more intricate, community workshops like FabLabs—designed to provide access to digital fabrication tools, foster interdisciplinary knowledge exchange, and democratize innovation—offer a potential mechanism for reaching the knowledge frontier. By enabling user-driven innovation, FabLabs contribute to the broader discourse on how open-access resources shape inventive activity.

This study provides a comprehensive assessment of FabLabs' impact on regional innovation in Germany, measured by patent and utility model applications, patent quality indicators, and first-time inventors. While previous research emphasizes the importance of local access to knowledge and tools in fostering innovation (e.g., Audretsch & Belitski, 2020; Furman, Nagler, & Watzinger, 2021), this study reveals a complex and heterogeneous effect shaped by regional preconditions.

Although FabLabs do not significantly increase overall patenting activity at either the FUA or NUTS3 level, FabLabs demonstrate stronger localized effects, particularly in fostering AM-related patents and forward citations, suggesting a positive influence on technologically complex and high-impact innovation. Especially university-integrated FabLabs contribute to AM-related patents.

Regional preconditions play a critical role in shaping FabLabs' effectiveness. In already highly inventive regions, FabLabs amplify inventive activity and foster especially AM-related technologies, and first-time inventors which likely also encompass accidental inventors–users who did not initially intend to invent. This contrasts with Bao (2025), who finds that U.S. Makerspaces primarily attract entrepreneurs rather than non-inventors. In less inventive regions, FabLabs primarily benefit those already familiar with patenting, likely due to their access to advanced technology, while also fostering patent interdisciplinarity. However, over time, FabLabs might also contribute to building innovative capacity in these regions. This may be an avenue for future research.

FabLabs exert a more substantial and consistent impact in urban areas, particularly in terms of first-time inventors, AM-related patents, and patent quality indicators. Urban regions, with their developed innovation ecosystems, exhibit sustained patenting growth over time, while rural regions experience weaker and more inconsistent effects. This suggests that the effectiveness of FabLabs depends on the presence of complementary resources, networks, and commercialization infrastructure, which are more prevalent in urban environments.

Further heterogeneity analysis underscores the role of pre-existing community workshops in shaping FabLabs' impact. Urban regions with both Makerspaces and Hackerspaces benefit the most, experiencing long-term increases in patenting, AM-related patenting, and first-time inventors. This suggests that FabLabs' positive effects stem primarily from access to advanced tools, as the culture of knowledge sharing and user-driven innovation is already established within the existing community workshop ecosystem. In rural areas, FabLabs demonstrate the greatest potential when they serve as the first community workshop, filling gaps in innovation infrastructure and fostering inventive activity. However, weaker pre-existing innovation ecosystems, dispersed population, and smaller user bases prolong adaptation periods and may thus delay measurable innovation gains.

Overall, the findings indicate that FabLabs are not a one-size-fits-all solution for regional innovation but rather interact dynamically with local innovation ecosystems. In urban settings, they amplify existing inventive activity, particularly in complex technological domains, and enable new (accidental) inventors, with access to tools likely being the key mechanism. In rural areas, FabLabs can serve as catalysts for innovation, albeit with delayed and less consistent impacts. In these regions both access to tools and the culture of knowledge diffusion may play a role. These results emphasize the importance of tailoring FabLab policies and support mechanisms to regional conditions, ensuring that they complement rather than compete with existing innovation infrastructures.

One potential avenue for maximizing FabLabs' impact is explicitly integrating them with local incubators or accelerators to facilitate the transition from idea generation to commercialization. This could help innovations mature into fully developed products or businesses. The TU Ilmenau exemplifies this by recently linking its university-integrated FabLab with the university-owned incubator, Ilmkubator (Technische Universität Ilmenau, 2024).

Despite providing novel insights into the localized effects of FabLabs, this study faces several limitations. First, the inability to track individuals over time limits the ability to estimate the direct impact of FabLab usage on an individual's inventive output. As a result, this study cannot fully disentangle the relative contributions of different FabLab features—such as standardized digital fabrication tools, the knowledge sharing culture, or the global FabLab network—to fostering innovation. Future research could explore how FabLabs affect the probability of becoming an inventor over time.

Second, despite controlling for various regional characteristics, the potential for selection bias remains. Unobservable regional factors may influence the decision to establish a FabLab, making it difficult to fully rule out endogeneity concerns. Future studies could address this by employing quasi-experimental approaches with richer micro-level data.

Third, limited statistical power, particularly in rural analyses, result in large standard errors, weakening causal inference. Expanding the dataset to include the broader

DACH region (Germany, Austria, and Switzerland) could help address this issue. Additionally, future research could explore FabLabs' impact on inventive outcomes beyond patents, such as startup creation, software development, and product innovation. Especially, since prior studies have focused on Hackerspaces and Makerspaces in fostering entrepreneurship but not FabLabs (van Holm, 2017; Cuntz & Peuckert, 2023; Bao, 2025). While FabLabs' role in fostering AM-related startups has been examined (Schwierzy, 2025), broader investigations into FabLabs role in entrepreneurial ecosystems may be insightful.

By shedding light on the heterogeneous impacts of FabLabs across different regional contexts, this study contributes to a more nuanced understanding of their role in fostering innovation. FabLabs democratize access to tools and knowledge, but their impact is contingent on pre-existing innovation ecosystems, user composition, and local institutional frameworks. These insights underscore the importance of tailoring policy interventions to regional contexts, ensuring that FabLabs can effectively complement and strengthen open innovation ecosystems.

4

Value Estimates for U.S. Patent Grants

4.1 Introduction

There is by now a vast literature demonstrating that patents play an important role in creating incentives for invention and innovation (Moser, 2005; Hall & Helmers, 2024). Therefore, it is not surprising that private and public actors seek patent protection for many of their inventions. Given the high propensity to patent, especially in industrialized countries, patent data have become an important tool for detecting and monitoring technical developments. While patents do not cover all sectors of the economy and all innovation activities equally well[1], and while the incentives for patenting

*This chapter is based on joint work with Jonathan Federle and Dietmar Harhoff. For this chapter, I used ChatGPT and DeepL-Write to refine grammar and wording.

[1]In some cases, innovators may choose to rely on secrecy or on fast market introduction of products. See Levin et al. (1987) and Cohen et al. (2002) for discussions of the trade-offs between different forms of appropriation.

have been subject to important changes over time, patent systems are nonetheless the most comprehensive systematic collection of information on technical advances. That makes them well-suited for addressing a broad range of research questions.

But patents also continue to pose challenges to researchers: they are typically complex documents and difficult to assess based on their content alone. Their value is idiosyncratic and depends crucially on the technical contribution made above and beyond prior art, but also on the nature of complementary assets that the patent holder commands (Teece, 1986). Patent values are therefore highly heterogeneous, and their distribution is highly skewed, meaning that a small fraction of patents accounts for a large share of the total value represented by all patents in a firm's portfolio, in a sector, or in the overall economy (Scherer, 1965; Harhoff, Scherer, & Vopel, 1998).

Given the heterogeneity of patents' strategic and economic importance, having reliable patent value estimates is crucial for many analytical tasks. Economists have followed various approaches to assessing patent values, ranging from surveys of managers (Harhoff et al., 1999) or inventors (Giuri et al., 2007; Torrisi et al., 2016), over studies of licensing and income generated thereof (Bray & Lee, 2000; Ziedonis, 2007; Abrams et al., 2019; Akcigit et al., 2021) to structural models illuminating the distributional properties of the renewal value of patents (Pakes & Schankerman, 1984; Pakes, 1985; Lanjouw, 1998). Each of these approaches offers unique advantages, but also faces problems: surveys are costly, suffer from substantial non-response, and are naturally constrained in the number of patents covered; the curation of valuation or impact data from firm-level licensing cases is time-consuming and results may be specific to the licensor's policies; structural models of renewal decisions only reveal overall distributional properties, but not value estimates at the patent level.

Any (reasonably) comprehensive, inexpensive, and up-to-date source of patent value measures would therefore be highly advantageous. Stock market data have obvious appeal in this context, as the data-generating process is quasi-automatic and comprehensive for patents held by firms listed on public financial markets. The relationship between medium- to long-term returns in the stock market and the quantity and average quality of a firm's patent stock has already been established in pioneering work by Hall, Jaffe, and Trajtenberg (2005). However, this approach does not readily lend itself to obtaining value estimates for individual patents.

Obviously, using stock market data to estimate the value of patents has its disadvantages as well, as only publicly listed patent owners are covered. Therefore, a valuation of patents held by universities, small business owners, or other non-listed entities is not directly feasible. Moreover, the value of patent applications ultimately not granted cannot be ascertained directly. However, results for patents issued to private entities may allow researchers to compute synthetic (as-if) values for similar patents of non-listed actors or for applications that were rejected (Kline et al., 2019; Hsu et al., 2021).

The most prominent study using stock market reactions to compute value estimates at the patent level is the seminal work of Kogan et al. (henceforth KPSS 2017). KPSS focus on the stock market reaction of publicly listed U.S. firms to patent grants by the USPTO, using a three-day window starting on Tuesdays when patent grants are announced. Within these event windows, they decompose the stock market return into a patent-related and an idiosyncratic (unrelated) component. After adjusting for prior market expectations, they compute the value of individual patents as the average patent-related capital gain per granted patent in a given week to the focal firm. KPSS value estimates have been used in several hundred studies on topics such as the financial implications of innovation on firm value, the impact of age and size on the value of patents, the value of "green" patents, etc. Some authors use control variables (of different sorts) to compensate for possible biases. Undoubtedly, Kogan et al. (2017) represent a groundbreaking contribution, but only a few studies explicitly consider possible weaknesses of KPSS estimates and guard against them empirically.[2] We expand on the seminal work of KPSS, arguing that the original estimates are subject to at least two[3] important biases. We propose a fundamental generalization of their framework to avoid these biases.

First, we change a basic assumption regarding the signal-to-noise ratio (SNR). This variable denotes the proportion of the patent-related return variance to the overall variance of returns during patent-grant days. In the original KPSS framework, the SNR is considered constant over firms and over time. We introduce two degrees of freedom in the SNR by allowing both the patent-related and the patent-unrelated return components to vary at the firm-day level. Most importantly, we allow the SNR

[2]Higham, de Rassenfosse, and Jaffe (2021) comment on the surprisingly low correlation of the KPSS estimates with a large set of other variables that have been shown to have robust statistical association with patent value. W. Wang (2023) estimates the signal-to-noise ratio separately for different types of patents, but restricts attention to singleton patent issues.

[3]We defer the discussion of a third issue to a later version of this paper, but acknowledge its presence and implications here. Kogan et al. (2017) assume that the patent as it will or will not be granted is publicly known, and that the patent issue only resolves uncertainty regarding the grant decision. This assumption may be justified for a significant portion of the patents covered in their data, in particular when the application has been disclosed before. However, some patents have been kept secret. In an extension of this paper, we identify all patents that are likely still unknown when granted by the USPTO and provide a correction of our estimates to account for this fact. Our revised estimates indicate that none of the results presented here will change qualitatively.

to vary depending on the number of patents granted. Our revised SNR yields patent values which exhibit substantially higher correlations with external patent-value correlates.

Second, we introduce another degree of freedom for patents that are granted simultaneously with other patents to the same firm on the same day. Naturally, the heterogeneity of patent values that are granted on the same day to the same firm is masked in daily returns. KPSS address this issue by averaging the observed patent-related market value gains over the number of simultaneously granted patents, thus possibly introducing measurement error in about 85% of the patents in their sample. We demonstrate below, that averaging is internally inconsistent with important aspects of their data. The values of singleton grants are systematically related to patent characteristics, and the aggregate value of patents issued simultaneously is similarly related to the aggregates of these patent characteristics. Thus, the empirical evidence contradicts an averaging approach. Lastly, the patent economics literature has long argued that patent values follow a highly skewed distribution, which is most likely altered in an averaging operation.

To address the averaging problem, we estimate hedonic patent value regressions, which allow us to assign weights to patents within the daily[4] bundle of patent grants while keeping the estimated total over simultaneously granted patents constant. The new value estimates show greatly increased correlations with known patent value correlates. Based on our computations we also provide a completely revised dataset for nearly 3 million patent grants at the USPTO.

[4]Since the USPTO announces patent grants on every Tuesday, we focus on daily data.

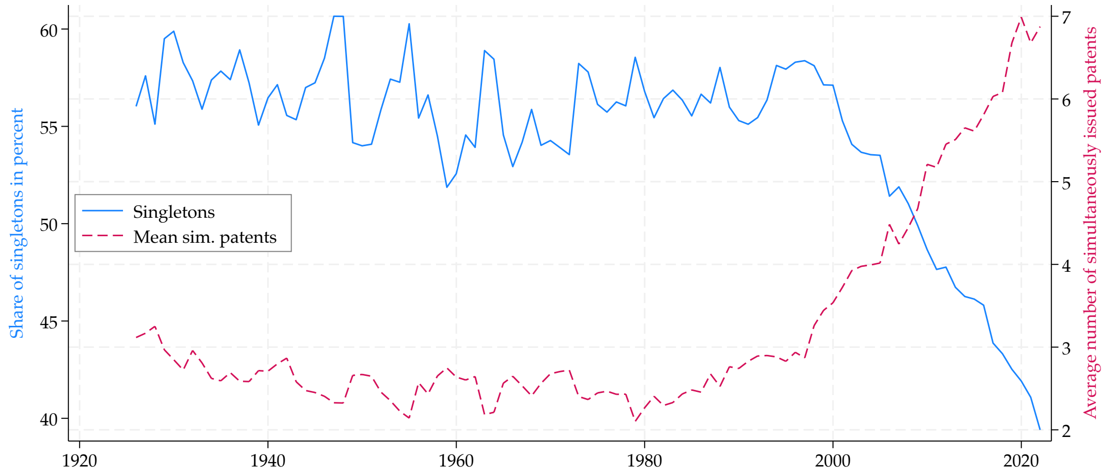


Figure 4.1: Simultaneously issued patents and singletons over time

Notes: Time in years depicted on x-axis. Blue solid line denoted on left y-axis and outlines the share of patents issued as singletons (no other patent issued on the same firm day) in percent. Red dashed line denotes the annual average of the number of simultaneously issued patents on the same firm day.

There is considerable heterogeneity over time in the number of grants issued per day to any given firm. Figure 4.1 shows the share of singleton patent grant events and the average number of patents issued per day over time. Both variables remain stable from 1925 to the mid-1980s, after which the average number of patents issued per day begins to increase and the share of singleton grant events starts to decline. We argue that this trend is directly related to the sharp increase in patenting activities by U.S. firms in the 1980s.[5] While there may be various explanations for the rise of patenting and litigation, the clear time patterns revealed in Figure 4.1 point to a problem in the currently used KPSS data, going beyond the threat of a time-invariant size bias. KPSS data have been used to study patent litigation, patent portfolio strategies, and the role of patents for VC-financed companies. Many of these phenomena became more prevalent in the 1980s, coinciding with changes in the number of patents issued per day - a central variable driving the signal-to-noise ratio which is held constant in the original framework. Consequently, the bias in KPSS data may affect studies of these phenomena in complex ways unless researchers prudently control for its causes. We argue that, rather than relying on controls and multiple robustness exercises, it may

[5]Hall and Ziedonis (2001) argue that this development initially started in the semiconductor industry and then spilled over into other industries. It was driven by firms' efforts to build patent portfolios that could be used for litigation, which also became more frequent during these years. Moreover, the 1980s also saw growing importance of the venture capital (VC) financing model in the USA. See Hall (2004) for further exploration of the rise of patents in the 1980s.

be preferable to develop a revised set of estimates for the value of patent grants. We propose an approach for doing so and provide such a dataset.

The remainder of this paper is structured as follows. Section 4.2 introduces our theoretical framework and its empirical implementation. Section 4.3 describes the construction and properties of our dataset. We proceed to discuss our revised value estimates and their properties in Section 4.4. Section 4.5 presents a discussion of three important empirical relationships which may be viewed differently when using the original KPSS data or our estimates: we study the nexus between value estimates and citations, the relationship between firm size and other patent-level characteristics and patent value. We briefly conclude in Section 4.6.

4.2 Theory and Empirical Framework

4.2.1 Basic Assumptions

Throughout this section, we describe the empirical framework employed to infer patent values from stock market returns. Our point of departure is similar to the framework of KPSS: The market-adjusted return of firm i on day t, $R_{i,t}$ is given by,

$$R_{i,t} = \nu_{i,t} + \varepsilon_{i,t} \,. \tag{4.1}$$

Here, $\nu_{i,t}$ is the aggregate value of all patents of firm i which have been granted in day t expressed as a fraction of the firm's prior day market capitalization and discounted by the market's pre-grant assessment of the grant being successful.Variable $\varepsilon_{i,t}$ denotes the idiosyncratic component of the firm's return, i.e., the component of returns unrelated to patent grants.

It follows that, in absolute terms, the aggregate value of all patents granted on a firm day is given by

$$\lambda_{i,t} = (1-\bar{p})^{-1} * E[\nu_{i,t}|R_{i,t}] * M_{i,t-1} \,, \tag{4.2}$$

where $M_{i,t-1}$ is the pre-grant market capitalization of firm i and $\bar{p} = 0.56$ is the unconditional probability of a successful patent application throughout the entire sample.[6] Adjusting for the success probability assumes that the market has already priced the

[6] W. Wang (2023) relaxes the assumption of a homogeneous grant rate, but applies this generalization only to singleton patent issues.

patent value prior to the patent grant — discounted, however, by the chance of the patent not being approved after all. In this regard, we mirror the assumptions of KPSS by assuming constant grant probabilities and complete public knowledge of patent values before the grant.

Backing out the conditional patent values, $E[v_{i,t}|R_{i,t}]$, required to compute (4.2), is contingent on distributional assumptions of both $v_{i,t}$ and $\varepsilon_{i,t}$. Since patent grants are assumed to have strictly positive value[7], KPSS impose a positively truncated normal distribution on them, i.e.,

$$v_{i,t} \sim N^{+}(0, \sigma^2_{v,i,t}) . \tag{4.3}$$

The idiosyncratic return component follows a random walk and is assumed to be normally distributed such that

$$\varepsilon_{i,t} \sim N(0, \sigma^2_{\varepsilon,i,t}) . \tag{4.4}$$

Given these distributional features, it follows that

$$E[v_{i,t}|R_{i,t}] = \delta_{i,t}R_{i,t} + \sqrt{\delta_{i,t}}\sigma_{\varepsilon,i,t}\frac{\phi\left(-\sqrt{\delta_{i,t}}\frac{R_{i,t}}{\sigma_{\varepsilon,i,t}}\right)}{1-\Phi\left(-\sqrt{\delta_{i,t}}\frac{R_{i,t}}{\sigma_{\varepsilon,i,t}}\right)} , \tag{4.5}$$

with ϕ and Φ denoting the standard normal pdf and cdf, respectively. $\delta_{i,t}$ is the signal-to-noise ratio (SNR) of observed returns and given by

$$\delta_{i,t} = \frac{\sigma^2_{v,i,t}}{\sigma^2_{v,i,t} + \sigma^2_{\varepsilon,i,t}} . \tag{4.6}$$

4.2.2 Signal-to-Noise Ratio

Our first major deviation of the KPSS framework comes into play when estimating the SNR.[8] KPSS assume this ratio to be constant across the entire sample. The underlying assumption is that both $\sigma^2_{v,i,t}$ and $\sigma^2_{\varepsilon,i,t}$ vary over firms and time, but only in constant proportions to each other. Yet, the idiosyncratic return of firms is — by definition

[7] This assumption is probably justified for the vast majority of cases, assuming that the grant does not bring "bad news", e.g., in the form of major restrictions of the claims of the patent, relative to the claims that the market has expected.

[8] To estimate the SNR, KPSS employ the following equation: $\log(R^2_{i,d}) = \gamma I_{i,d} + cZ_{i,d} + \mu_{i,d} + u_{i,d}$, where $R_{i,d}$ denotes the three-day return of firm i starting on day d, and $I_{i,d}$ is an indicator variable equal to one if at least one patent was issued to firm i on day d. The control variable, $Z_{i,d}$, comprises day-of-week dummies, and $\mu_{i,d}$ denotes interacted firm-year fixed effects. The sample is restricted to firms that have been granted at least one patent over the sample period. The SNR is then recovered as $\hat{\delta} = (1 - e^{-\gamma})$. For details, see Kogan et al. (2017, p. 679)

— independent of patent-related information. This is particularly problematic in the context of multiple patents being granted simultaneously, as we would expect the variance of the aggregate value of all patents granted, $\sigma^2_{v,i,t}$, to increase in the number of patents granted. However, since the variance of the idiosyncratic return component, $\sigma^2_{\varepsilon,i,t}$, is independent of patent grants, it cannot systematically vary with the number of granted patents. Thus, a constant SNR mechanically imposes the value variance of one patent to equal that of multiple simultaneously granted patents. Since the variance determines not only the dispersion but also the mean of the positively truncated distribution defining the value of patents in equation (4.3), failure to account for a varying SNR invariably leads to biased patent valuations.

To illustrate the inherent issue, consider the following scenario: A representative example of a patent-unrelated risk captured by the idiosyncratic return variance, $\sigma^2_{\varepsilon,i,t}$, would be the headquarters of firm i to catch fire on day t. Imposing $\sigma^2_{\varepsilon,i,t}$ to vary in constant proportions with $\sigma^2_{v,i,t}$ and acknowledging that the $\sigma^2_{v,i,t}$ should vary with the number of granted patents, effectively implies the probability of patent-unrelated risks like the headquarters catching fire to increase in constant proportions with the number of patents granted by the USPTO on a given day. Since this is arguably implausible, a constant SNR implies that patent values are mechanically biased depending on the number of other patents simultaneously issued.

Instead of imposing a constant SNR across all firm-day observations, we back out both the patent-unrelated return variance and the patent-related variance and let them vary at the firm level. Leveraging the fact that the idiosyncratic return component is — by definition — independent of patent-related returns, we can compute the measurement error $\sigma^2_{\varepsilon,i,t}$ as the realized variance of returns on non-grant days of firm i in the year preceding patent grant day t.[9]

We now proceed with the definition of the patent-related variance of returns. Recall from equation (4.1) that, on each firm-day, the overall return of a firm is given by the portion of patent-related and patent-unrelated returns. Thus, the total variance of

[9] In contrast, KPSS back out the variance of the measurement error using their estimate of the SNR through $\sigma^2_{\varepsilon,i,t} = 3\sigma_{i,t}(1+3d_{i,t}\hat{\delta})^{-1}$. Here, $\sigma_{i,t}$ is the volatility of firm i in year t and $d_{i,t}$ is the fraction of trading days on which patents have been granted. Their approach of estimating the measurement error is not feasible when allowing the SNR to differ across the number of patents granted. We provide robustness tests reproducing the original KPSS estimator using our measurement error and show that the resulting patent values are almost identical to the original specification with the Pearson correlation exceeding 0.99.

returns is characterized by the following formula:

$$\sigma^2_{i,t} = Var(v_{i,t}) + \underbrace{Var(\epsilon_{i,t})}_{=\sigma^2_{\varepsilon,i,t}} + \underbrace{Cov(v_{i,t}, \varepsilon_{i,t})}_{=0} \tag{4.7}$$

In other words, the total return variance of a firm is simply the sum of the variances of its patent-related returns, its idiosyncratic returns, and their covariance. However, the covariation between patent-related returns and the idiosyncratic return—by definition—must be equal to zero. Since the distribution of patent values follows a positively truncated normal distribution, its variance corresponds to:

$$Var(v_{i,t}) = \sigma^2_{v,i,t}(1 - 2/\pi). \tag{4.8}$$

Equipped with these inputs, we can back out the observed patent value variance for any day on which a patent has been granted using Equation (4.7). Thus, we define the observed value variance of patents granted to firm t on day i as the difference between the overall variance of the return of firm i on day t and the variance of the patent-unrelated return which we defined above as the variance of the returns of firm i on non-grant days in the year preceding day t.

While this non-parametric estimate has intuitive properties and yields, on average, an unbiased estimate of the patent value variance, it also has limitations. Most importantly, it can take on negative values if the return variance on a given day is unusually low compared to the reference period (the year prior to the patent grant).

To address this issue, we estimate how the variance of patent values depends on the number of granted patents. To this end, we introduce the following linear model:

$$\sigma^2_{v,i,t} = \alpha_i + \beta \frac{\#Patents}{M_{i,t-1}} + u_{i,t} \tag{4.9}$$

Thus, we relate the observed value variances of patents granted to firm i on day t to a set of firm fixed effects and the ratio of the number of patents granted to the pre-grant market value of the firm.

In relating the variance of patents not only to the number of patents but to the number of patents relative to the firm's market capitalization, we seek to account for single patents having a larger impact on the values of small firms than on big corporations — at least in *relative* terms. Upon estimating this model, we find that the estimate of β is highly significantly positive (p-value < 0.001).

In the next step, we predict the patent value variances, based on our estimated model above such that:

$$\hat{\sigma}^2_{v,i,t} = \hat{\beta} \frac{\#Patents}{M_{i,t-1}}. \tag{4.10}$$

In this way, we arrive at our estimate for the firm-day specific variance of the patent-related return variance.

Having both estimates for the patent-related variance ($\hat{\sigma}^2_{v,i,t}$) and our non-parametric estimate for the variance of returns unrelated to patents ($\hat{\varepsilon}^2_{v,i,t}$), we can now proceed to plug these values into the formula for the SNR as specified in (4.6). In contrast to the original specification, our newly derived SNR now varies at the firm-day level and depends on the number of patents simultaneously granted. Equipped with these inputs, we proceed to compute the aggregate value of all patents granted to firm i on day t, as given by equation (4.5).

4.2.3 Simultaneous Patent Grants

At this juncture, KPSS average the values of all simultaneously granted patents in a given firm week. In this way, the value estimates forfeit any heterogeneity across simultaneously issued patents. Given that both the number of patents granted and the value of patents correlate with firm size, the averaging procedure severely disturbs the patent distribution not only at the sample level but also across firms and time. Moreover, while the mean value may be a reasonable estimator if no additional information about patent value is available, such an assumption is internally inconsistent with the KPSS estimates (and would be inconsistent with our estimates using an amended computation of the SNR). We will show below that the value estimates for singleton patents vary systematically with characteristics of the patents that are known at the grant event (or even earlier), and that this also holds for aggregate relationships. This is not a novel insight, and a large literature has studied such correlations for the purpose of predicting patent value, either for aggregate portfolios or single patents (Harhoff, Scherer, & Vopel, 1998; Hall, Jaffe, & Trajtenberg, 2005; Higham, de Rassenfosse, & Jaffe, 2021).

We seek to overcome the problem of computing patent values for individual patents by regressing patent value estimates on a set of patent quality characteristics publicly known at the time of the grant. We then use these estimates to distribute realized capital gains across simultaneously issued patents according to their estimated value. In doing so, we maintain the total value estimate for the bundle of patents issued. Our reweighing of patent value estimates is therefore mean-preserving (over patents simultaneously granted to the firm). Formally, let $J_{i,t}$ denote the set of patents granted to firm i on day t. Furthermore, let Q denote a set of patent-related quality characteristics. To back out the general relationship between each characteristic and patent values around grants, we estimate the following linear regression

$$\lambda_{i,t} = \alpha_i + \beta_t + \sum_{q \in Q} \left(\psi_q \sum_{j \in J_{i,t}} q_j \right) + u_{i,t}\,, \tag{4.11}$$

where α_i and β_t denote a set of firm and time fixed-effects, respectively, and $u_{i,t}$ denotes the error term. In this way, ψ_q, yield estimates of how much each patent quality characteristic $q \in Q$ contributes to realized capital gains. Based on these estimates, we define the predicted value of each patent j as[10]

$$\hat{\eta}_j = \sum_{q \in Q} \hat{\psi}_q q_j \,. \tag{4.12}$$

Lastly, we use $\hat{\eta}_j$ to disaggregate the firm-day patent values from equation (4.2) to the patent level. Thus, we define

$$\xi_j = \frac{\hat{\eta}_j}{\sum_{p \in J_{i,t}} \hat{\eta}_p} \lambda_{i,t} \,, \tag{4.13}$$

where j_i refers to the issuing firm and j_t refers to the grant day of patent j. In this way, we step back from KPSS' original methodology of distributing capital gains in equal proportions to simultaneously granted patents but instead distribute them in a qualitatively meaningful way using a set of publicly known quality characteristics.

4.3 Data

We base our analysis on patent value data from KPSS, utilizing the most recent dataset available through their public GitHub repository.[11] This dataset, an extension of the original data used in their 2017 QJE paper, includes 3,160,453 patents issued by the USPTO from January 5, 1926, to December 29, 2022. The KPSS dataset associates each patent with a unique permanent firm identifier (permco) provided by the Center for Research in Security Prices (CRSP). Leveraging this identifier, we merge the patent data with daily stock market data for 8,432 U.S. firms between 1926 and 2022 from the Center for Research in Security Prices (CRSP).[12] This joint coverage of stock and patent data yields a final dataset of 3,155,528 patents. Hence, we can compute patent value estimates along the lines of KPSS with our SNR amendment for almost the full KPSS 2022 sample (99.6%).

[10] Note that $\hat{\eta}_j$ does not include fixed effects and therefore has no qualitative interpretation in absolute but only in relative terms.
[11] See https://github.com/KPSS2017
[12] We utilize variables from CRSP such as daily closing price, stock return, value-weighted market index, and shares outstanding. From the closing price and shares outstanding, we compute each firm's daily market capitalization.

Table 4.1: Summary statistics of patent characteristics

	N	Mean	Median	SD	min	max
Panel A: characteristics of all patents						
#patents	2,703,605	23.99	10	38.68	1.00	637
GDP (patent family)	2,703,605	20.25	17.55	10.78	5.24	76.13
claims (grant)	2,703,605	16.42	16.00	11.15	1.00	868
PL refs (wgt. 3yr cit.)	2,703,605	304.97	55.00	1,167.25	0.00	32,715
NPL refs	2,703,605	5.50	0.00	23.69	0.00	3,230
areas cited	2,703,605	3.62	3.00	3.05	0.00	35
Panel B: patent characteristics in weeks with one grant per firm (singletons)						
#patents	354,670	1.00	1.00	0.00	1.00	1.00
GDP (patent family)	354,670	16.89	14.96	10.19	5.24	76.13
claims (grant)	354,670	16.84	15.00	13.31	1.00	683
PL refs (wgt. 3yr cit.)	354,670	278.10	27.00	1185.00	0.00	32,318
NPL refs	354,670	7.33	0.00	28.06	0.00	2,372
areas cited	354,670	3.91	3.00	3.47	0.00	35
Panel C: patent characteristics in weeks with multiple grants						
#patents	2,348,935	27.47	13.00	40.38	2.00	637
GDP (patent family)	2,348,935	20.76	17.87	10.77	5.24	74.51
claims (grant)	2,348,935	16.36	16.00	10.78	1.00	868
PL refs (wgt. 3yr cit.)	2,348,935	309.02	60.00	1,164.49	0.00	32,715
NPL refs	2,348,935	5.22	1.00	22.95	0.00	3,230
areas cited	2,348,935	3.58	3.00	2.98	0.00	35
Panel D: aggregate characteristics for firm week events						
#patents	678,804	3.99	1.00	8.94	1.00	637
GDP (patent family)	678,804	80.66	23.56	223.75	5.41	14,752.77
claims (grant)	678,804	65.42	24.00	153.06	2.00	10,850
PL refs (wgt. 3yr cit.)	678,804	1,214.58	73.00	4,977.08	0.00	226,974
NPL refs	678,804	21.91	1.00	91.32	0.00	5,979
areas cited	678,804	14.42	5.00	33.25	0.00	2,239

Notes: The table shows summary statistics of patent characteristics. #patents are the number of patents that are issued per day. *GDP (patent family)* is the cumulative GDP covered by all patent filings based on the priority filing of the focal US patent. *claims (grant)* are the number of claims in the grant document. *PL refs* are the number of U.S. patent literature references weighted by three-year citations, and *NPL refs* are the number of non-patent literature (NPL) references. *areas cited* are the number of different technical areas cited by the focal patent.

Moreover, we obtain a broad range of patent characteristics as part of our hedonic value estimation from PATSTAT.[13] These characteristics include the cumulative GDP covered by all patent filings based on the priority filing of the focal US patent (GDP-weighted size of the international patent family)[14], U.S. patent literature references weighted by three-year citations, claims in the grant document, non-patent literature (NPL) references, and the technical areas[15] cited in the grant document. Since some of the value correlates are only available for patents granted after 1970, our hedonic analysis is currently restricted to 2,703,605 patents issued between 1970 and 2022.

The summary statistics for patent characteristics of these 2,703,605 patents are presented in Table 4.1, Panel A. On average, each patent is issued together with 24 other patents. The patent families associated with the issued patents cover, on average, a GDP of US$ (1982) 20.3 trillion. The issued patents typically contain 16 claims and 305 weighted patent literature references (weighted by 3-year citations). On average, the issued patents have 5.50 non-patent literature references. Additionally, they cite an average of 3.62 different technological areas in their backward citations.

We also characterize the singleton observations, defined as single patent grants on a given day and the non-singleton observations separately (Table 4.1, Panel B and C). A comparison of the panels points to some selectivity: singleton patent families cover countries with less GDP (16.89 vs. 20.76 US$ trillion), have fewer patent references (278.10 vs. 309.02), but a larger number of non-patent literature references, i.e., are closer to science (7.33 vs. 5.22). Moreover, singleton patents cite prior art from more technical areas more frequently than non-singletons (3.91 vs. 3.58). These differences are statistically highly significant.

Finally, in Panel D, we provide descriptive statistics at the firm-day level, comprising 678,804 firm-day observations (Table 4.1, Panel D). On average, firms receive 4 patents per day, with a maximum of 637 patents granted in a single firm-day. We will use these data in our hedonic estimates of aggregated patent bundles issued to firms.

[13] PATSTAT database, version Autumn 2023

[14] We use the DOCDB patent family definition provided in our PATSTAT data. The GDP data are taken from the Penn World Tables and complemented with information from Jordà, Schularick, and Taylor (2017) and Funke, Schularick, and Trebesch (2023).

[15] We use the technical area classification developed by Schmoch (2008). See https://www.wipo.int/export/sites/www/ipstats/en/docs/wipo_ipc_technology.pdf.

4.4 Results

In this section, we establish and contrast the results from estimating patent values using the KPSS methodology and our revised model. The first subsection deals with the effects and impact of modifying the SNR, and the second subsection outlines solutions to the averaging problem.

4.4.1 Fixing the Signal-to-Noise Ratio

The natural starting point revealing the dynamic properties of the SNR is to let its constituents, the patent-unrelated return variance and the patent-related variance, vary at the firm-day level (see Section 4.2.2).

The latter is obtained in estimating regression 4.9, where we regress the observed patent-related variance of returns on the number of patents granted relative to the firms' market capitalization. The results are depicted in Table 4.2. As is apparent from the table, the variance of the patent-related return components strongly varies with the number of patents granted relative to the grantee's pre-grant market value. The association is highly significant (p-value < 0.01).

Table 4.2: Estimating the patent-related variance

	$\sigma^2_{v,i,t}$
$\frac{\#patents}{MV_{i,t-1}}$	0.00351***
	(0.000629)
FXT	YES
Adj. R^2	0.062
N	41,699,393

Notes: Table shows results of estimating regression (4.9). It shows how the variance of the patent-related return component depends on the relative number of patents granted to firms. Both the regressor and the dependent variable are winsorized at the 0.1% and 99.9% levels. FXT: firm-year fixed effects. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.

The value of the coefficient extends beyond mere descriptives as it allows us to compute revised SNRs based on the number of patents granted relative to the firms' market values.

To illustrate the impact of assuming a constant SNR, consider Table 4.3 which shows the average SNR in the KPSS and our revised specification. The former assumes a constant SNR, whereas the latter implies a SNR that varies at the firm-day level. As becomes apparent from the table, a fixed SNR masks substantial heterogeneity in the number of patents simultaneously granted. However, it is important to note that a direct comparison between the KPSS SNR and our revised version is inherently limited, as KPSS analyze three-day returns, whereas our specification focuses on grant-day returns.

Table 4.3: SNR and number of simultaneously granted patents

	Signal-to-noise ratio	
#patents	KPSS	FHK
1	0.0145	0.0649
5	0.0145	0.08
10	0.0145	0.0873
20	0.0145	0.1331
50	0.0145	0.1442

Notes: The table shows the number of simultaneously granted patents against the estimated SNR for the original specification of KPSS and our revised specification.

Taking into account the differing SNRs also affects the patent value properties of the sample at large, as shown in Table 4.4. Since the distribution of the number of simultaneously granted patents carries a substantial right tail, the KPSS specification yields substantially higher values for the majority of patents in the overall sample. Our explanation is that the KPSS estimate overstates patent values for the most frequent observations in the sample, i.e., for event windows with only one or few patent grants. In fact, a static SNR yields mean patent values almost four times higher and median patent values that are about 50% higher than those resulting from our revised SNR. This may help to resolve a puzzle that KPSS themselves note in their paper (see p. 682) where they state that in comparison to estimates by Giuri et al. (2007) their average value of patent estimates “seems a bit high”.

Table 4.4: Patent value statistics by SNR specification

SNR	Mean	Median	SD	SD / Mean	Skewness	Maximum
KPSS	10.81	3.25	30.01	2.78	14.96	2,683.24
FHK	2.95	2.14	3.03	1.02	2.42	65.24

Notes: N=3,158,981. All values in 1982 $US million. Expressed in 2025 $US million, the mean of KPSS's value estimates is 41.94, compared to 11.45 for the FHK estimates. The maximum value in the KPSS estimates is 10,410.97, whereas in the FHK estimates, it is 253.13.

Higham, de Rassenfosse, and Jaffe (2021) observe that KPSS values exhibit low correlations with observable patent characteristics. However, as shown in Table 4.5, allowing the SNR to vary dynamically by firm and day significantly improves these correlations. On average, accounting for differing SNRs increases correlation values at the aggregate level by 50% compared to the original KPSS values.

Table 4.5: Correlation between capital gains and cumulative value correlates

Patent characteristic	KPSS	FHK
#patents	0.277	0.617
GDP (patent family)	0.263	0.541
PL references (wgt. by 3yr cites)	0.399	0.571
claims (grant)	0.332	0.648
NPL references	0.407	0.552
US cites 3yrs	0.357	0.589
US cites 10yrs	0.332	0.577

Notes: The table shows the correlation between value estimates (either the KPSS original or our SNR-corrected FHK estimates) and a number of patent characteristics as summarized in Table 4.1, Panel D. Patent characteristics are aggregated over simultaneously issued patents. The patent characteristics comprise the cumulative GDP covered by the international patent family (*GDP (patent family)*), the number of U.S. patent literature references weighted by three-year citations (*PL refs*), the number of claims in the grant document (*claims (grant)*), the number of non-patent literature (NPL) references (*NPL refs*), the number of different technical areas cited by the focal patent (*areas cited*), and the number of USPTO citations within 3 years and 10 years of filing (*US cites 3yrs* and *US cites 10yrs* respectively). N=678,743.

In addition, we regress the value estimates from KPSS and our approach on aggregated value correlates to highlight differences between the static and dynamic SNR. Table 4.6 reports the R^2 and within-R^2 of these regressions. To distinguish the explanatory power of the number of simultaneously granted patents from other correlates, we first regress the estimates on the number of patents alone and then include the additional correlates. The extremely low within-R^2 in column (3) suggests that in the KPSS specification, nearly all variation is explained by firm-year fixed effects. In contrast, in our dynamic SNR-based specification, the within-R^2 remains substantial even after including firm-year fixed effects.[16]

Table 4.6: Regression analyses between capital gains and cumulative value correlates

	KPSS			FHK		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	22.31***	18.72***	40.66***	4.929***	1.460***	5.970***
	(0.156)	(0.156)	(0.0472)	(0.0215)	(0.0194)	(0.00858)
Adj. R^2	0.0799	0.241	0.966	0.345	0.560	0.957
Within R^2	0.0799	0.241	0.000759	0.345	0.560	0.461
N	855,618	855,618	833,641	855,614	855,614	833,637
#patents	YES	YES	YES	YES	YES	YES
Value correlates	NO	YES	YES	NO	YES	YES
FXT	NO	NO	YES	NO	NO	YES

*Notes:*The dependent variables are the original KPSS capital gains and our SNR-corrected FHK estimates. Standard errors in parentheses. Patent characteristics are aggregated over simultaneously issued patents. The patent characteristics comprise the cumulative GDP covered by the international patent family, the number of U.S. patent literature references weighted by three-year citations, the number of claims in the grant document, the number of non-patent literature (NPL) references, the number of different technical areas cited by the focal patent, and the number of USPTO citations within 3 years and 10 years of filing. We also include the quadratic transformations of these patent characteristics. FXT: firm-year fixed effects. *** p<0.01, ** p<0.05, * p<0.

4.4.2 Fixing the Averaging Problem

To address the averaging issue, we consider different empirical approaches. All of these can be considered hedonic estimates as we model the value of the patent as a function of observable characteristics. We limit our analysis to information known at the time of the patent grant at the USPTO. These characteristics have already been described in section 4.3. We compute four different hedonic value estimates in order to study their properties. Our exercise is motivated by the observation that the original KPSS estimates appear to be only weakly correlated with known patent value proxies. Averaging for a large share of the sample (85%) may well contribute to these

[16]The results are similar if we winsorize the dependent variables at the 0.1% and 99.0% levels (see Table D.1).

low correlations. Naturally, a reweighing using hedonic regressions will strengthen correlations with the variables used as hedonic characteristics and predictors. However, as we show below, the correlation with ex post citation measures - which are not included in the hedonic exercise - also increases. The following regressions use estimates of the total market value gain to firms in a given three-day event window. We apply our SNR correction as described in section 4.2.2.

Our four estimators are as follows (we relegate regression results to the appendix). The first estimation, *HED-SNG*, focuses on singleton observations and models patent values as a function of the five characteristics, firm fixed effects, as well as technical area and grant year fixed effects. The regression results of this specification are reported in Table D.2, based on 354,633 observations. As expected, we find positive and statistically highly significant coefficients for most of the variables. The coefficients for the number of areas cited in the patents' backward references and for the NPL references are negative. Since we are mostly interested in obtaining a reasonable prediction of patent value, we refrain from an in-depth interpretation of the coefficients. The fit of the regression is not particularly good in the specification without fixed effects, but improves with the inclusion of year and technical area dummies. Most of the variation is explained by firm fixed effects. We use the results from column (3) for obtaining predictions for all patents and use these to reweigh the averaged patent value estimates such that the firm-day estimate of the total gain in market value is preserved. Note that by definition, the estimated values for singleton grants are maintained here (as in the following models), the rescaling only affects patents granted simultaneously.

Our second model, *HED-AGG*, follows the aggregation logic outlined in equation (4.11) and in the rescaling equation (4.13). We report the results of the regressions in Table D.3. Given the necessity of an additive specification, we use quadratic transformations of our aggregate hedonic characteristics to allow for deviations from a strictly linear relationship. In the absence of aggregation bias, we can use the coefficient estimates to predict pseudo-values at the patent level. We use these as weights to improve upon the averaged values. Since using the coefficients to predict at the individual patent level results in negative predictions for some cases, we impose a value of 0.05 (50,000$US) in these cases. Our third model, *HED-TOP*, is based on the observation that the value of patent portfolios is strongly determined by their most valuable patents (Scherer & Harhoff, 2000). To estimate the relationship at the aggregate level, we encode all patent characteristics as order statistics. Specifically, we define our regressors as the number of cases, within a given firm-week event, in which an individual patent's characteristics fall into the top five percent and the upper quartile (but not the top five percent) of the distribution. The distribution is computed within technical area and over five-year intervals. Similarly to the previous case, where we used aggregates of the value correlates, predicting values for individual patents results in a few instances of negative predictions, which we handle as previously described.[17] The regression results are presented in Table D.4.

Finally, we use external value information for 3,874 patents, for which we have direct survey estimates from the PatVal project described in Torrisi et al. (2016). Using a log-linear specification, we derive coefficients, then compute predictions for our full sample and use these to reweigh the contributions of these patents to the total market value gain observed within a particular event window (*HED-PATVAL*). The regression results are presented in (Table D.5).

The value estimates resulting from the four models are summarized in Table 4.7.[18] The rescaling applied does not alter the mean of the distributions. Hence, we preserve the property that, across all estimates, the mean of our value estimates is, on average, 87% lower than the mean of the KPSS values. The median and skewness values vary considerably across the four models. In particular, the *HED-AGG* and *HED-PATVAL* estimators exhibit considerably greater skewness and higher maxima compared to both the original KPSS estimators and our SNR-adjusted estimators with averaging (*FHK*).

Table 4.7: Descriptive statistics for value estimates

	Mean	Median	SD	Skewness	p99	Max.
KPSS	11.89	3.66	32.15	14.11	128.51	2,683.24
FHK	2.61	1.96	2.58	2.16	12.01	53.78
HED-SNG	2.61	1.93	2.64	2.25	12.46	53.7
HED-AGG	2.61	0.89	5.02	7.63	22.73	317.57
HED-TOP	2.61	1.76	2.85	2.33	13.27	53.78
HED-PATVAL	2.61	1.52	3.55	5.06	16.84	221.64

Notes: N= 2,703,605. All value estimates in 1982 $US million. *HED-SNG* are the values predicted using the log-specification with singletons only. *HED-AGG* are the prediction results based on the linear specification for aggregated patent grants. *HED-TOP* are the predicted values of the specifications for aggregates using upper percentile dummies. *HED-PATVAL* are the prediction results using the PatVal survey data. In 2025 $US million, the mean of the hedonic specifications is 10.13. The maximum values for *HED-AGG* and *HED-PATVAL* are 1,232.17 and 859.96, respectively. In contrast, the KPSS mean in 2025 $US million is 42.25, with a maximum value of 10,410.97.

[17]We replace negative values with 0.05 in 50.42%cases of the predicted estimates in the *HED-AGG* specification, and for 2% in the *HED-TOP* specification.

[18]Note that we consistently utilize specifications without firm fixed effects.

4.5 Applications

4.5.1 Citations

As a first application we explore the relationship between citations and patent value estimates. Computing average values for simultaneously granted patents risks introducing measurement error – high value patents granted at the same time as low-value ones would be underestimated, while the value of lesser patents would be overstated. We are concerned that averaging introduces biases, resulting in an artificially low correlation between value correlates and the KPSS value estimates. In other words, relying on value estimates with these biases is likely to distort estimation results in multivariate settings.

Table 4.8: Citation analysis using averaging estimators (KPSS and FHK)

	KPSS			FHK		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(1+cites)	0.213***	0.168***	0.003***	0.113***	0.087***	0.001***
	(0.009)	(0.009)	(0.001)	(0.005)	(0.005)	(0.000)
N	2,703,450	2,703,450	2,685,324	2,703,446	2,703,446	2,685,320
R^2	0.130	0.235	0.960	0.147	0.249	0.959
Firm size	NO	YES	YES	NO	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
FE	CXT	CXT	CXT, FXT	CXT	CXT	CXT, FXT

Notes: The dependent variable is log(patent value estimate). See KPSS (2017), section III, Table II for their original estimates. Standard errors in parentheses, clustered by grant year. Firm size is approximated using the market value of the firm in the week preceding the grant event. CXT: tech classification x grant year. FXT: firm fixed effects x grant year. *** p<0.01, ** p<0.05, * p<0.1.

The exercise undertaken to address this concern essentially replicates an analysis conducted by KPSS (section III). KPSS presented a graph in which they plot a citation measure against their value estimates, identifying an almost linear relationship (see their Fig. II). But the data underlying this analysis is at the aggregate level, where observations are pooled to the level of percentiles. The patent-level regression analysis described in KPSS's Table II appears to confirm our concerns. Initially, the relationship between forward citations and patent value estimates is (without additional controls) strong and statistically significant, yielding an elasticity of 0.174. Once more controls are successively introduced – such as size, volatility, and fixed effects for firm years and technology years – the coefficient diminishes and is no longer statistically significant in the final specification. We now demonstrate that our value estimates exhibit a different pattern.

Table 4.9: Citation analysis using hedonic estimators (*HED-PATVAL* and *HED-AGG*)

	HED-PATVAL			HED-AGG		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(1+cites)	0.239***	0.214***	0.152***	0.266***	0.249***	0.196***
	(0.007)	(0.006)	(0.005)	(0.008)	(0.008)	(0.012)
N	2,703,446	2,703,446	2,685,320	2,703,446	2,703,446	2,685,320
R^2	0.187	0.261	0.845	0.146	0.168	0.574
Firm size	NO	YES	YES	NO	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
FE	CXT	CXT	CXT, FXT	CXT	CXT	CXT, FXT

Notes: The dependent variable is log(patent value estimate). See KPSS (2017), section III, Table II for their original estimates. Standard errors in parentheses, clustered by grant year. Firm size is approximated using the market value of the firm in the week preceding the grant event. CXT: tech classification x grant year. FXT: firm fixed effects x grant year. *** p<0.01, ** p<0.05, * p<0.

We first perform the same exercise as KPSS using averaged values based on a constant SNR (denoted KPSS in Table 4.8). This approach replicates the pattern previously described. When we use averaged values based on the dynamic SNR estimate, the results are qualitatively similar. In contrast, using hedonic estimates (*HED-PATVAL* and *HED-AGG*, see Table 4.9) yields significantly different results. Even with the most aggressive introduction of fixed effects, the citation coefficient decreases but remains relatively large (0.152 in column (3) and 0.196 in column (6)). The HED-SNG coefficient estimates behave similar to the averaging estimators, confirming our view that these may be affected by selection bias and that the rescaling based on singletons introduces only minor changes in the patent-level value estimates. The *HED-TOP* coefficient estimates are almost identical to the *HED-PATVAL* and *HED-AGG* estimates (see Table D.6).

The exercise allows us to conclude that averaging has the expected effect of artificially reducing the partial correlation between citations and patent value measures. Our result also contradicts the conclusion by KPSS that citations are better considered a measure of scientific value than of monetary value.

Although we do not currently endorse a specific variant of our hedonic estimators, the properties illustrated in Table 4.7 suggest that these estimators are preferable to the original KPSS estimates.

4.5.2 Patent Term Extension, Standard Essential Patents, and Firm Size

As a second application, we investigate the relationship between additional patent- and firm-level characteristics and our value estimates, as well as those proposed by KPSS. As we noted in the previous application, the use of average values for simultaneously granted patents in KPSS may introduce measurement errors. Specifically, we assess and compare the correlation between estimated patent values and key factors, including patent term extensions, standard essentiality, and firm size.

To analyze the relationship between patent value and patent term extensions, standard essentiality (SEPs), and firm size we regress the estimated patent values on indicator variables. These indicators are expected to be positively correlated with patent value estimates. We include an indicator variable that equals one if the patent belongs to a patent family that contains at least one SEP.[19] Information whether a patent's term has been extended by the USPTO is obtained from PATSTAT. To account for firm size, we include a categorical variable capturing the number of employees per firm-year, based on data from COMPUSTAT.

[19]This classification is based on Brachtendorf, Gaessler, and Harhoff (2023)

Table 4.10: Patent term extension, SEPs, firm size and value estimates

	KPSS	HED-AGG	HED-PATVAL
	(1)	(2)	(3)
patent term ext.	-0.013	0.227***	0.168***
	(0.0466)	(0.0625)	(0.0395)
SEP	0.074**	0.218***	0.461***
	(0.035)	(0.040)	(0.023)
Missing	1.186***	0.625***	0.673***
	(0.100)	(0.072)	(0.070)
251–500 Employees	0.207***	0.176***	0.144***
	(0.033)	(0.023)	(0.021)
501–2.500 Employees	0.492***	0.373***	0.346***
	(0.039)	(0.027)	(0.026)
2.501–10.000 Employees	0.833***	0.585***	0.560***
	(0.050)	(0.042)	(0.039)
10.001–50.000 Employees	1.145***	0.749***	0.768***
	(0.082)	(0.057)	(0.056)
50.001–250.000 Employees	0.984***	0.596***	0.670***
	(0.102)	(0.067)	(0.066)
>250.000 Employees	1.235***	0.849***	0.954***
	(0.127)	(0.087)	(0.084)
Constant	-0.165*	-0.979***	-0.529***
	(0.083)	(0.056)	(0.055)
N	2,702,158	2,702,154	2,702,154
R^2	0.866	0.504	0.753
Year FE	YES	YES	YES
Firm FE	YES	YES	YES
CXT	YES	YES	YES

*Notes:*The dependent variable is ln(patent value estimate). The base category for firm size are firms with 250 or less employees (see Table D.7). Standard errors in parentheses, clustered by grant year. CXT: tech classification x grant year. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.10 reports the regression results, controlling for year, firm fixed effects, and year-technology-class fixed effects. In both our *HED-AGG* and *HED-PATVAL* specifications, patents with extended terms are highly significantly and positively correlated with patent value. In contrast, the correlation between KPSS estimates and patent term extension is negative and statistically insignificant. Similarly, the relationship between SEPs and patent value varies across specifications. SEPs are significantly positively correlated with patent value in our specifications, whereas the KPSS estimates yield substantially smaller coefficients. The results also underscore the positive correlation between firm size and patent value. In all three specifications, the coefficients increase almost linearly with the number of employees per firm-year. However, the coefficients in the KPSS estimates are much larger compared to our *HED-AGG* and *HED-PATVAL*.

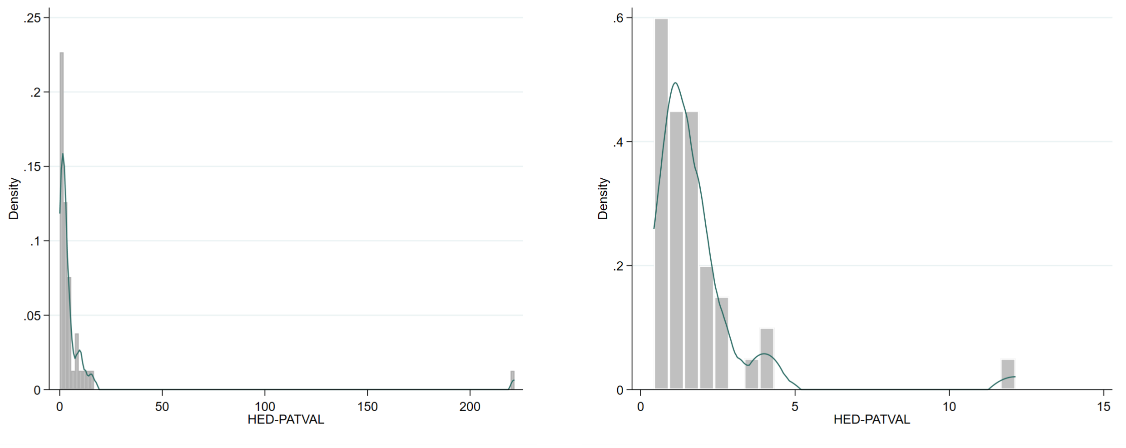
Overall, this analysis confirms the hypothesis that the averaging reduces the partial correlation between patent-level characteristics and patent value estimates.

4.5.3 Patent Value Examples

We further illustrate the variation in value estimates among simultaneously granted patents with two examples.

First, on May 3, 2022, Apple was granted 43 patents, including the patent with the highest value in our *HED-PATVAL* specification, estimated at $US 221.64 million. While KPSS assign the same value of $US 77.58 million to each of these patents, our *HED-PATVAL* estimates exhibit significantly greater variation, as shown in Figure 4.2 (all value estimates are in 1982 $US). Second, on April 5, 1994, IBM received 41 patents. Again, the histogram in Figure 4.2 demonstrates that our *HED-PATVAL* specification captures meaningful variation across patents, whereas KPSS assigns a uniform value of $US 3.37 million to all.

These examples highlight the ability of our improved method to better capture heterogeneity in patent value among simultaneously granted patents.



(a) Apple **(b)** IBM

Figure 4.2: Distribution of patent value estimates

Notes: Panel (a) depicts the distribution of patent values, estimated using the *HED-PATVAL* specification, for 43 patents simultaneously granted to Apple on May 3, 2022. In contrast, KPSS assign each of these 46 patents a uniform value of $77.58 million (deflated to 1982). Panel (b) shows the corresponding distribution for 41 patents granted to IBM on April 5, 1994, where KPSS similarly attribute a uniform value of $3.37 million (deflated to 1982) per patent.

4.6 Conclusion

In their pioneering work, KPSS introduced a novel method for computing patent value estimates based on short-term stock market reactions to patent grant events. Their approach has sparked renewed interest in employing patent value estimates across various analytical studies. Computing the estimates can be highly standardized, and the resulting data can cover a broad range of sectors and time periods.

In our analysis, we identify two sources of bias in the KPSS value estimates that may impose limitations in subsequent applications. We propose major amendments to their approach to address these issues.

We first argue that assuming a constant SNR introduces a non-trivial bias. Allowing both the patent-related and patent-unrelated return components to vary at the firm-day level, and permitting the SNR to adjust based on the number of simultaneously granted patents, introduces two degrees of freedom that lead to significant changes in value estimates. These adjustments make the estimates more consistent with prior literature and internally coherent. Comparing our dynamic SNR to KPSS's static SNR in relation to the number of simultaneously issued patents further reveals substantial heterogeneity, which the static SNR in KPSS masks. Compared to KPSS estimates, our

FHK estimates of patent-level capital gains to firms are considerably lower. The mean of the KPSS patent value distribution is nearly five times higher than ours, while the median is more than 50% higher.

We address the limitations of the averaging approach by using hedonic regressions, which yield weights that allow us to redistribute value among individual patents issued simultaneously to a given firm. This intervention improves the correlation between our modified values and established value correlates. Examining the relationship between forward citations and patent value estimates, we also show that our hedonic adjustment significantly affects the citation coefficient in a log-linear value regression, even though forward citations to the focal patent are not included in the hedonic computations. In addition, the hedonic patent values exhibit a stronger correlation with patent value indicators, such as standard essentiality and patent term extensions, compared to KPSS. Furthermore, we find suggestive evidence that the value of patents granted to smaller firms is overestimated in KPSS. Considering the central role of firm size in various empirical studies within innovation research, industrial organization, and related fields, we expect this bias to spill over to other variables, such as capital investment, R&D expenditures, and IT-related activities.

The contribution of this paper is first one of clarification – despite of the widespread use of the KPSS value data, there have been only very few investigations seeking to understand which assumptions drive the main results of KPSS. We think that our work will enhance the understanding of the approach chosen by KPSS and that it will offer significant improvements. Moreover, we hope that the paper will open the discussion on how stock market reactions can be better exploited for the purpose of economic research on innovation and competition. Towards that objective, we plan to extend our analysis provided here to specific relationships such as the nexus between patent value and R&D inputs, the impact of science and knowledge externalities on patent value, and the influence of competition on stock market assessments of patent value. Finally, the value estimates provided in this paper may have applications that we as authors may not anticipate.

A

Appendix to Chapter 1

Megaprojects, Digital Platforms, and Productivity – Evidence from the Human Brain Project

A.1 Figures

A.1.1 Figures – Geographical Diversity

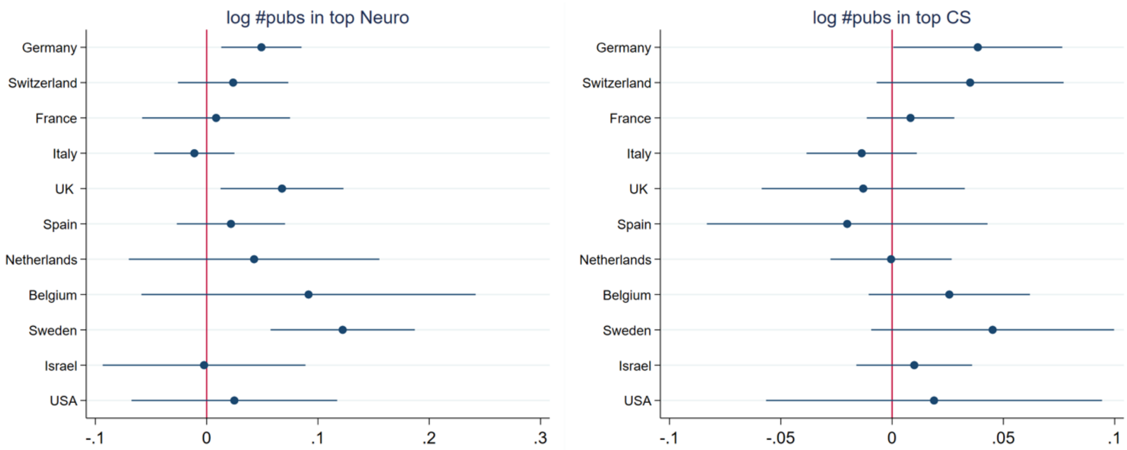


Figure A.1: Regression results incl. country-specific interaction terms - top journal publications (2000-2023)

Notes: This figure depicts the coefficients based on equation $y_{it} = \beta HBP_{it} * country_{it} + \delta_i + \delta_t + \varepsilon_{it}$. For each individual, we assigned the country indicator based on the country with which they were affiliated at the time of their initial participation in HBP. The order of countries reflects the overall engagement of these countries. The baseline outcomes for both regressions are statistically significant at the 1% level. For the number of publications in top neuroscientific journals it equals 0.1211041 and for those in top CS journals it amounts to 0.0406739. Confidence intervals are reported at the 95% level.

A.1.2 Figures – Subsamples (2000-2023)

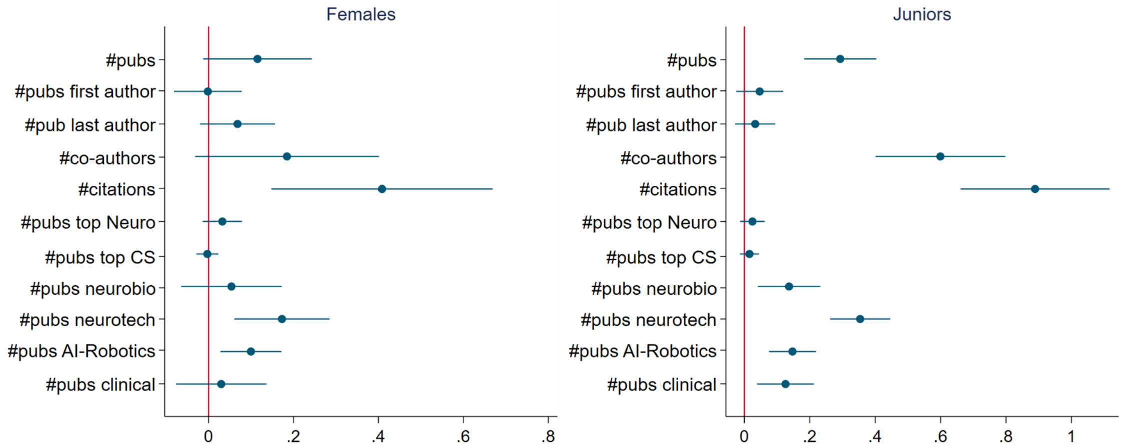


Figure A.2: Regression results for female and junior subsamples

Notes: This figure depicts the coefficients based on equation 1.1 for the female and junior subsample for the years 2000-2023. The #co-authors encompass the unique number of co-authors per author-year. Confidence intervals are reported at the 95% level.

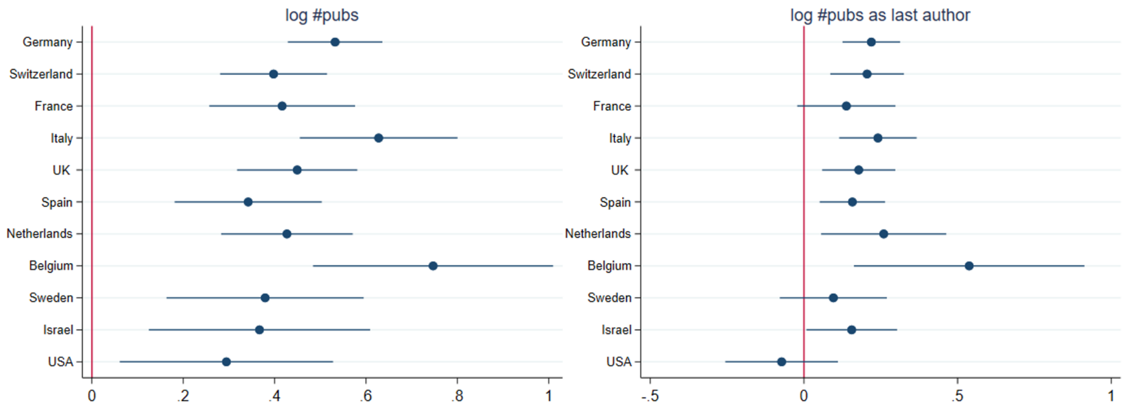


Figure A.3: Regression results incl. country-specific interaction terms - #pubs and #pubs as last author (2000-2023)

Notes: This figure depicts the coefficients based on equation $y_{it} = \beta HBP_{it} * country_{it} + \delta_i + \delta_t + \varepsilon_{it}$. For each individual, we assigned the country indicator based on the country with which they were affiliated at the time of their initial participation in HBP. The order of countries reflects the overall engagement of these countries. For the number of publications (#pubs), the base outcome is highly significant at the 1% level and equals 0.3829277. Similarly, for the number of publications as the last author (#pubs as last author), the base outcome is also statistically significant at the 1% level, amounting to 0.1506259. Confidence intervals are reported at the 95% level.

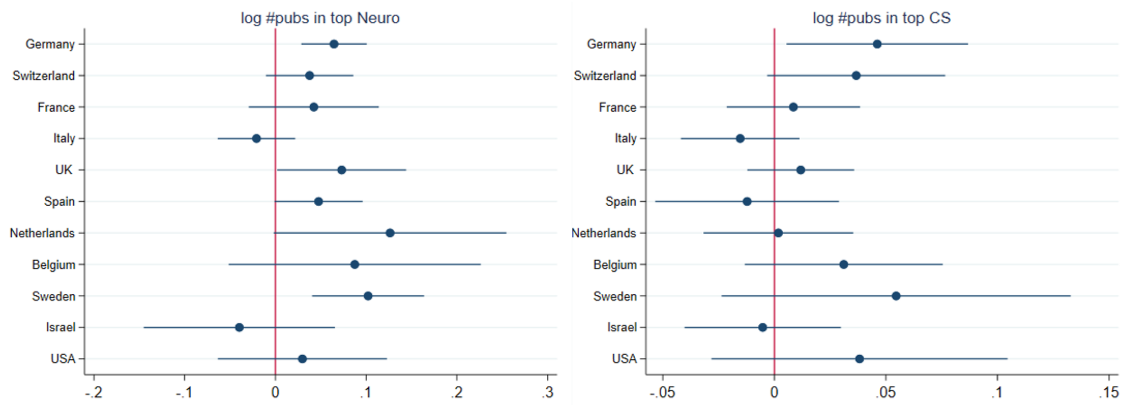
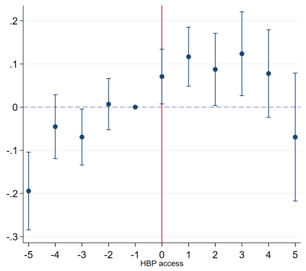


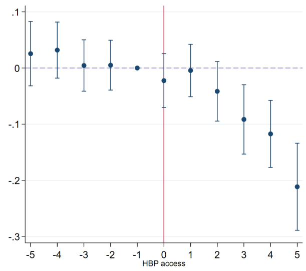
Figure A.4: Regression results incl. country-specific interaction terms - top journal publications - 2000-2023

Notes: This figure depicts the coefficients based on equation $y_{it} = \beta HBP_{it} * country_{it} + \delta_i + \delta_t + \varepsilon_{it}$. For each individual, we assigned the country indicator based on the country with which they were affiliated at the time of their initial participation in HBP. The order of countries reflects the overall engagement of these countries. Confidence intervals are reported at the 95% level.

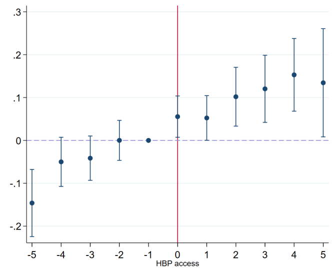
A.1.3 Figures – Event Studies



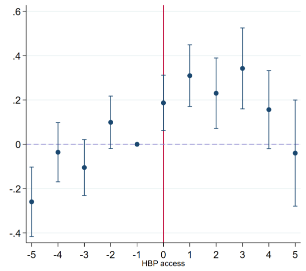
(a) #pubs



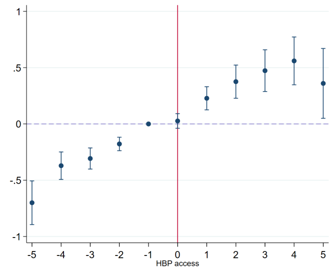
(b) 1st author



(c) last author



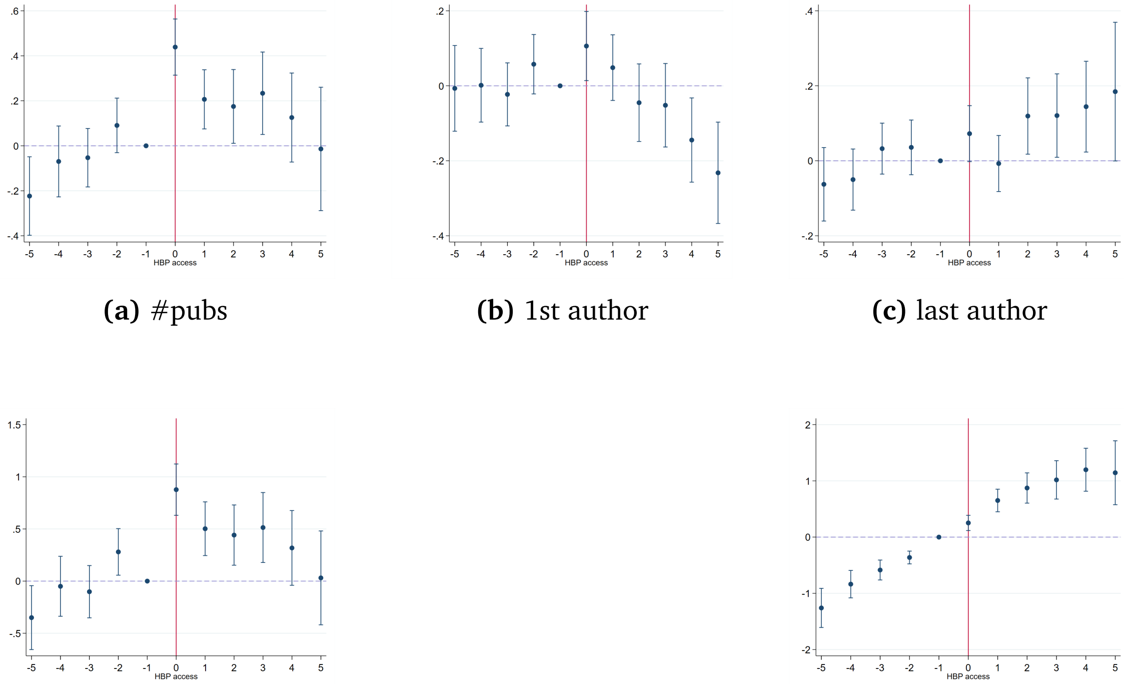
(d) #co-authors



(e) #citations

Figure A.5: Event studies for publications and collaborations

Notes: This figure presents event study estimates based on equation (1.2) for the full sample. All variables are in logarithms. The specification includes year and author fixed effects, with standard errors clustered at the author level. Confidence intervals are reported at the 95% level.

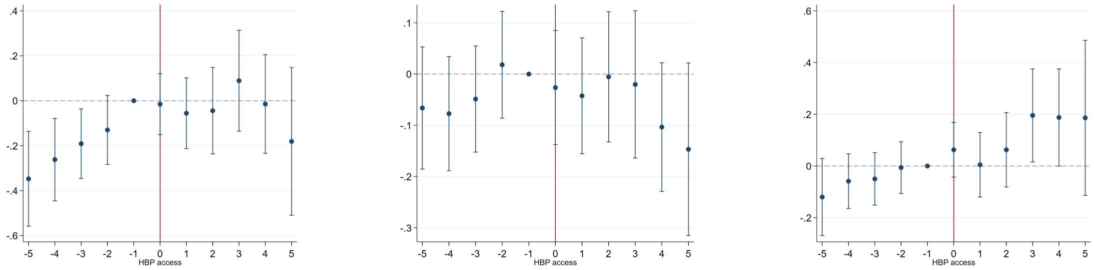


(a) #pubs **(b)** 1st author **(c)** last author

(d) #co-authors **(e)** #citations

Figure A.6: Event studies for publications and collaborations - junior sample

Notes: This figure presents event study estimates based on equation (1.2) for the junior subsample. All variables are in logarithms. The specification includes year and author fixed effects, with standard errors clustered at the author level. Confidence intervals are reported at the 95% level.

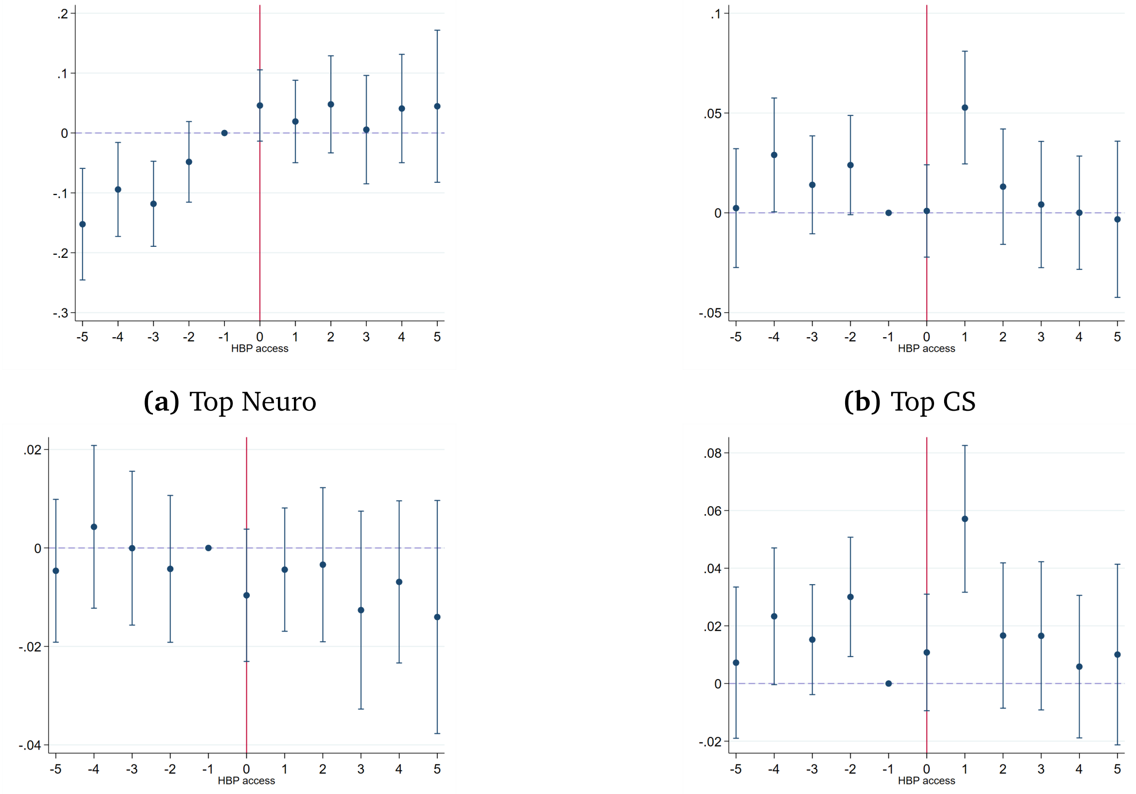


(a) #pubs **(b)** 1st author **(c)** last author

(d) #co-authors **(e)** #citations

Figure A.7: Event studies for publications and collaborations - female sample

Notes: This figure presents event study estimates based on equation (1.2) for the female subsample. All variables are in logarithms. The specification includes year and author fixed effects, with standard errors clustered at the author level. Confidence intervals are reported at the 95% level.



(a) Top Neuro

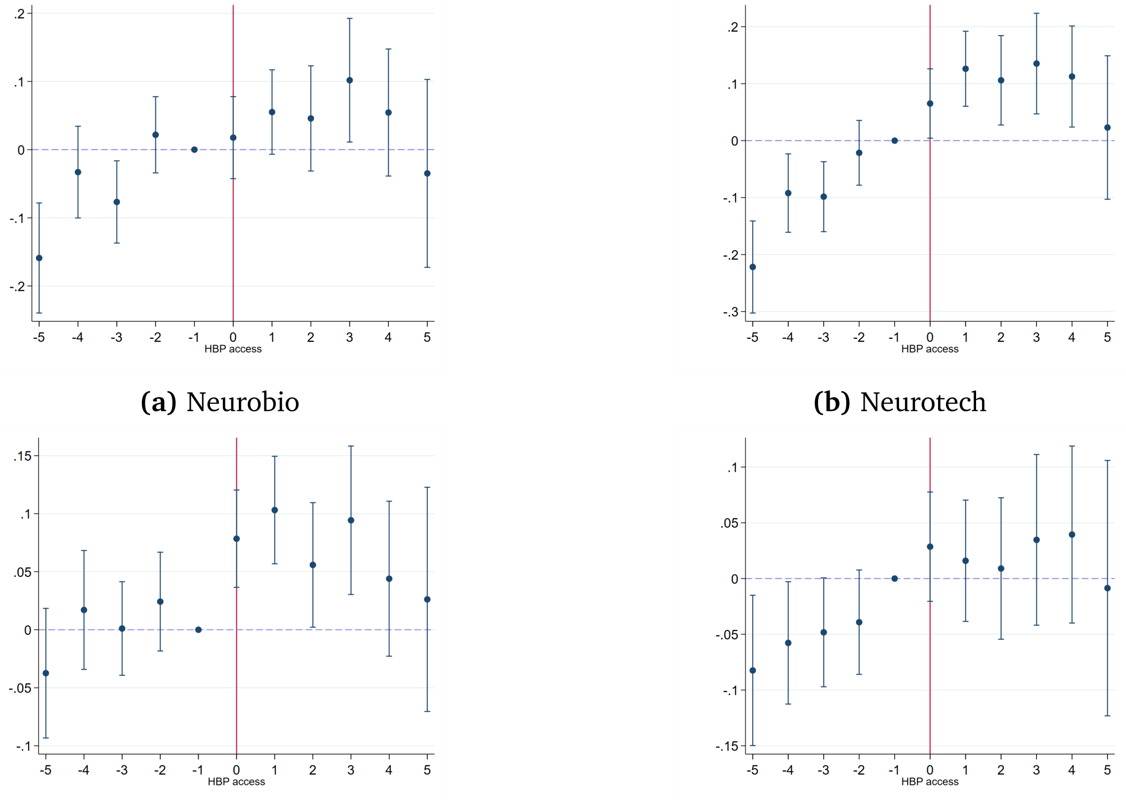
(b) Top CS

(c) CS A*

(d) CS A

Figure A.8: Event studies for journal quality

Notes: This figure presents event study estimates based on equation (1.2) for the full sample. All variables are in logarithms. The specification includes year and author fixed effects, with standard errors clustered at the author level. Confidence intervals are reported at the 95% level.



(a) Neurobio

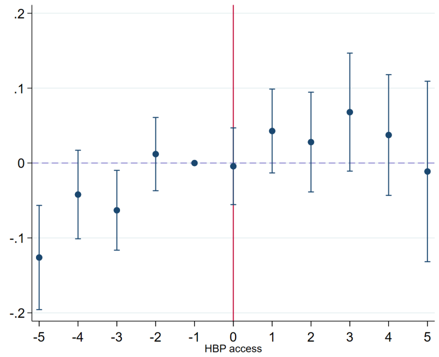
(b) Neurotech

(c) AI-Robotics

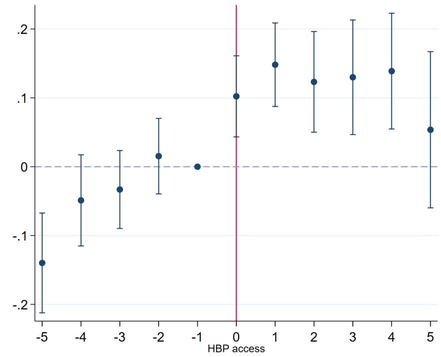
(d) Clinical

Figure A.9: Event studies for topic classification

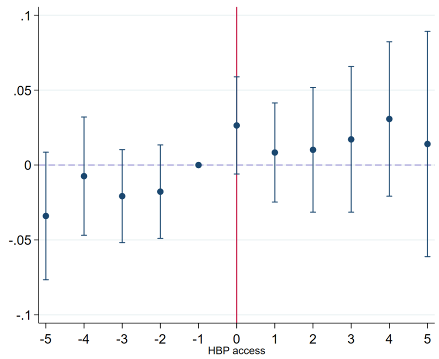
Notes: This figure presents event study estimates based on equation (1.2) for the full sample. All variables are in logarithms. The specification includes year and author fixed effects, with standard errors clustered at the author level. Confidence intervals are reported at the 95% level.



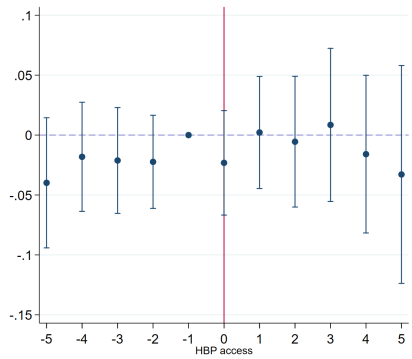
(a) Neurobio



(b) Neurotech



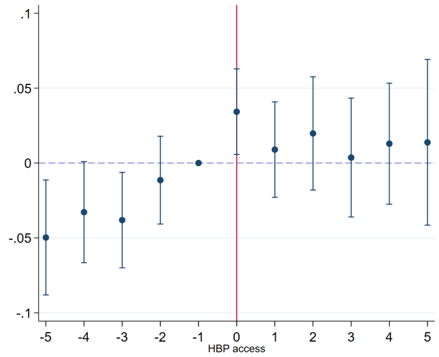
(c) AI-Robotics



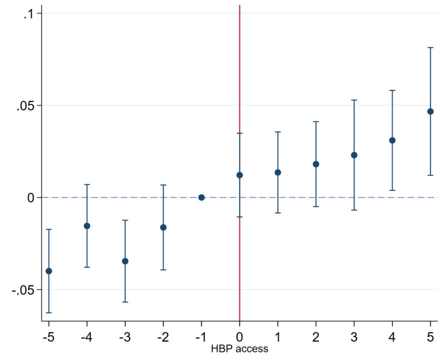
(d) Clinical

Figure A.10: Event studies for topic classification with the highest probability

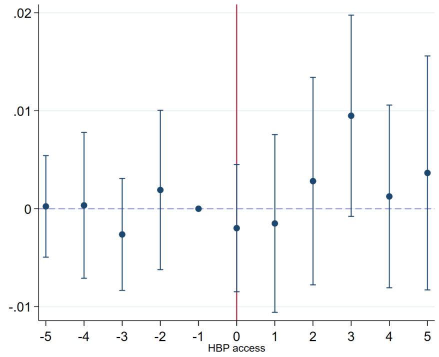
Notes: This figure presents event study estimates based on equation (1.2) for the full sample. All variables are in logarithms. The specification includes year and author fixed effects, with standard errors clustered at the author level. Confidence intervals are reported at the 95% level.



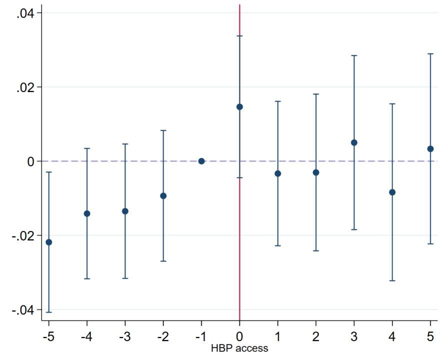
(a) Neurobio



(b) Neurotech



(c) AI-Robotics



(d) Clinical

Figure A.11: Event studies for topic classification in top neuro journals

Notes: This figure presents event study estimates based on equation (1.2) for the full sample. All variables are in logarithms. The specification includes year and author fixed effects, with standard errors clustered at the author level. Confidence intervals are reported at the 95% level.

A.2 Tables

A.2.1 Tables – Descriptive Analysis

Tables – Keywords and Topics

Table A.1: Keywords and topics for classification

Topic	Keywords	
ai_robotics	artificial intelligence & robotics	artificial intelligence; machine learning; neurorobotics
clinical	clinical applications, treatment, & care management	brain atlas; clinical trial; management; medical device; medical information; neuroethics; neurological treatment; new drug; patient care
neurobio	fundamental neuroscience & neurobiology	behavioral neuroscience; cellular neuroscience; clinical neuroscience; cognitive neuroscience; developmental neuroscience; molecular neuroscience; neurogenetics; neurophysiology; sensory neuroscience; theoretical neuroscience
neurotech	neurotechnology, simulation, & computational tools	BrainScaleS; SpiNNaker; brain simulation; cognitive architecture; computational biology; high-performance computing; neuroinformatics; neuromorphic computing; virtual brain

Tables – Subsample 2000-2023

Table A.2: Summary statistics (2000-2023)

	Freq.	%
Panel A: individual level		
Total	639	100
Female	129	20.19
Male	510	79.81
Active researchers per phase		
Phase 0	266	41.63
Phase 1	318	49.77
Phase 2	393	61.50
# affiliated countries		
Phase 0	23	
Phase 1	20	
Phase 2	21	
Panel B: Publication level		
Total publications	50,365	100
Type of publication		
Article	31,830	63.20
Conference paper	11,762	23.35
Review	2,751	5.46
Book chapter	1,257	2.50
Other (e.g., editorial, letter, book)	2,765	5.49
Type of journal		
Journal	37,166	73.79
Conference proceeding	8,913	17.70
Book / book series	4,269	8.48
Trade journals	17	0.03

Table A.3: Summary statistics: dependent variables (2000-2023)

Dep. variable (author-year)	N	Mean	Std. dev.	Min	Max
Panel A: Publications					
#pubs	15,336	3.2841030	6.3220970	0	131
#pubs as first author	15,336	0.3370501	0.7921521	0	15
#pubs as last author	15,336	1.0966350	2.87581	0	61
#distinct co-authors	15,336	17.85022	54.76995	0	1105
#co-authors	15,336	35.1032900	263.562	0	127.31
avg. #co-authors	15,336	4.1401560	7.8586470	0	100
#citations	15,336	198.862400	568.6459	0	118.01
Panel B: Journal Quality					
#pubs in Top Neuro	15,336	0.2192879	0.6953914	0	9
#pubs in Top Computer Science	15,336	0.0983307	0.5392768	0	11
#pubs in Computer Science A*	15,336	0.0372327	0.3523400	0	10
#pubs in Computer Science A	15,336	0.0610981	0.382000	0	8
Panel C: Topic Classification					
#pubs in topic Neurobio	15,336	2.2098330	4.8811690	0	109
#pubs in topic Neurotech	15,336	1.7704090	4.2150270	0	130
#pubs in topic AI-Robotics	15,336	0.6877934	2.4537340	0	67
#pubs in topic Clinical	15,336	1.1628850	3.1470130	0	49
Panel D: Probability of Topic Classification					
#pubs in topic Neurobio	15,336	0.2162334	0.3230883	0	1
#pubs in topic Neurotech	15,336	0.1987906	0.3074422	0	1
#pubs in topic AI-Robotics	15,336	0.0449272	0.1394472	0	1
#pubs in topic Clinical	15,336	0.0827588	0.1897918	0	1
Panel E: Topic Classification with the highest probability					
#pubs in topic Neurobio	15,336	1.3658710	3.4506730	0	44
#pubs in topic Neurotech	15,336	1.0783780	2.9003870	0	93
#pubs in topic AI-Robotics	15,336	0.2915363	1.5629710	0	49
#pubs in topic Clinical	15,336	0.5525561	1.9040510	0	44
Panel F: Topic Classification in Top Neuro Journals					
#pubs in topic Neurobio	15,336	0.2042254	0.6672929	0	8
#pubs in topic Neurotech	15,336	0.0562728	0.2762689	0	4
#pubs in topic AI-Robotics	15,336	0.0055425	0.0751171	0	2
#pubs in topic Clinical	15,336	0.0419927	0.2445302	0	5

A.2.2 Tables – Junior and Female Subsamples

Table A.4: HBP participation and publications - junior sample (author-year level)

	(1)	(2)	(3)	(4)
Panel A: Journal Quality (Top neuroscience/CS outlets)				
	Top Neuro	Top CS	CS A*	CS A
HBP	0.0162	0.0167	-0.00386	0.0116
	(0.0160)	(0.0148)	(0.00449)	(0.00716)
LHS mean	0.1157	0.0643	0.0217	0.0426
Panel B: Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.108***	0.328***	0.127***	0.117***
	(0.0413)	(0.0430)	(0.0294)	(0.0380)
LHS mean	1.3343	1.3340	0.5482	0.6041
Panel C: Probability of Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0238	0.163***	0.0189*	0.0306*
	(0.0181)	(0.0245)	(0.0105)	(0.0171)
LHS mean	0.1985	0.2487	0.0694	0.0824
Panel D: Topic Classification with the highest probability				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0358	0.284***	0.0350*	0.0446
	(0.0348)	(0.0402)	(0.0192)	(0.0296)
LHS mean	0.7662	0.8758	0.2809	0.3005
Panel E: Topic Classification in Top Neuro Journals				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.00988	0.0138*	-0.00271	-0.00476
	(0.0161)	(0.00768)	(0.00368)	(0.0101)
LHS mean	0.1022	0.0426	0.0071	0.0250
Observations	2,955	2,955	2,955	2,955
#authors	197	197	197	197

Notes: This table reports results estimating equation 1.1 "junior-at-the-start" sample (2008-2022). The sub-sample consists of HBP participants who entered with the status junior (under-/graduates, junior faculty). Outcome variables are log number of publications. One is added to all variables before taking the logarithm to include years without publications or publications without citations. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.5: HBP participation and publications - female sample (author-year level)

	(1)	(2)	(3)	(4)
Panel A: Journal Quality (Top neuroscience/CS outlets)				
	Top Neuro	Top CS	CS A*	CS A
HBP	0.0401*	-0.00620	-0.00388	-0.00399
	(0.0238)	(0.0133)	(0.00599)	(0.0122)
LHS mean	0.1618	0.0672	0.0202	0.0470
Panel B: Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0484	0.141***	0.0545**	0.0158
	(0.0530)	(0.0515)	(0.0274)	(0.0437)
LHS mean	1.8848	1.675	0.5674	1.0145
Panel C: Probability of Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0114	0.0512*	0.0114*	-0.0130
	(0.0261)	(0.0283)	(0.00668)	(0.0224)
LHS mean	0.2379	0.2151	0.0353	0.1143
Panel D: Topic Classification with the highest probability				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0492	0.108**	0.0373**	-0.0181
	(0.0464)	(0.0476)	(0.0188)	(0.0350)
LHS mean	1.1938	1.0946	0.3069	0.5069
Panel E: Topic Classification in Top Neuro Journals				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0454*	0.0188	0.000411	-0.0126
	(0.0230)	(0.0144)	(0.00328)	(0.0132)
LHS mean	0.1457	0.0553	0.0062	0.0382
Observations	1,935	1,935	1,935	1,935
#authors	129	129	129	129

Notes: This table reports results estimating equation 1.1 female sample (2008-2022). Outcome variables are log number of publications. One is added to all variables before taking the logarithm to include years without publications or publications without citations. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A.2.3 Tables – Subsample by Training

Table A.6: Topic classifications - field of training, junior subsample (author-year level)

	(1)	(2)	(3)	(4)	(5)
	#pubs	1st author	last author	#co-authors	#citations
Panel A: Publications and collaborations					
HBPxEECS	0.166**	0.0511	-0.0189	0.289**	0.450***
	(0.0774)	(0.0465)	(0.0427)	(0.133)	(0.153)
HBPxBio-EECS	0.349***	0.117*	0.00910	0.800***	0.634***
	(0.0972)	(0.0645)	(0.0716)	(0.182)	(0.231)
HBPxNatural/Life Sc.	0.299***	0.0537	-0.000119	0.695***	0.505***
	(0.0646)	(0.0429)	(0.0339)	(0.118)	(0.119)
Constant	0.307***	0.111***	0.0419**	0.553***	1.007***
	(0.0363)	(0.0234)	(0.0207)	(0.0616)	(0.0710)
Observations	2,955	2,955	2,955	2,955	2,955
#authors	197	197	197	197	197
	Neurobio	Neurotech	AI-Robotics	Clinical	
Panel B: Publication topics					
HBPxEECS	-0.00347	0.241***	0.116**	0.0312	
	(0.0582)	(0.0659)	(0.0533)	(0.0514)	
HBPxBio-EECS	0.152*	0.319***	0.171**	0.115	
	(0.0878)	(0.0705)	(0.0778)	(0.0717)	
HBPxNatural/Life Sc.	0.172***	0.325***	0.0731*	0.106**	
	(0.0497)	(0.0564)	(0.0399)	(0.0473)	
Constant	0.221***	0.188***	0.0692***	0.0912***	
	(0.0285)	(0.0312)	(0.0255)	(0.0265)	
Observations	2,955	2,955	2,955	2,955	
#authors	197	197	197	197	

Notes: This table reports results estimating equation $y_{it} = \beta HBP_{it} * field_i + \delta_i + \delta_t + \varepsilon_{it}$. The highest degree field indicator for each individual was assigned based on the field in which they obtained their highest degree. One is added to all variables before taking the logarithm to include years without publications or publications without citations. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.7: HBP publication quality - field of training (author-year level)

	(1) Top Neuro	(2) Top CS	(3) CS A*	(4) CS A
Panel A: full sample				
HBPxEECS	0.0170	-0.00262	-0.00769	0.00479
	(0.0110)	(0.0157)	(0.00749)	(0.0144)
HBPxBio-EECS	0.0485**	0.0247**	0.00500	0.0198**
	(0.0201)	(0.00989)	(0.00542)	(0.00821)
HBPxNatural/Life Sc.	0.0215*	0.0130*	-0.00134	0.0138**
	(0.0120)	(0.00742)	(0.00313)	(0.00685)
Constant	0.126***	0.0443***	0.00955***	0.0356***
	(0.00856)	(0.00578)	(0.00266)	(0.00521)
Observations	9,585	9,585	9,585	9,585
#authors	639	639	639	639
Panel B: junior scholars (junior faculties, graduate students)				
HBPxEECS	-0.00659	0.0510*	-0.00248	0.0539**
	(0.0148)	(0.0281)	(0.0103)	(0.0263)
HBPxBio-EECS	0.0484**	0.0213	-0.00269	0.0239
	(0.0215)	(0.0203)	(0.00986)	(0.0169)
HBPxNatural/Life Sc.	0.0228	0.00157	-0.00303	0.00421
	(0.0184)	(0.0148)	(0.00626)	(0.0139)
Constant	0.0252***	0.00704	0.00352	0.00352
	(0.00947)	(0.00741)	(0.00457)	(0.00567)
Observations	2,955	2,955	2,955	2,955
#authors	197	197	197	197
Panel C: female scholars				
HBPxEECS	0.0261	-0.0505	-0.0204	-0.0379
	(0.0258)	(0.0451)	(0.0201)	(0.0449)
HBPxBio-EECS	0.0229	0.00161	-0.00615	0.00771
	(0.0340)	(0.00969)	(0.00522)	(0.00890)
HBPxNatural/Life Sc.	0.0368	0.00512	0.00206	0.00302
	(0.0243)	(0.0100)	(0.00477)	(0.00917)
Constant	0.0856***	0.0394***	0.00537**	0.0363***
	(0.0185)	(0.0129)	(0.00231)	(0.0130)
Observations	1,935	1,935	1,935	1,935
#authors	129	129	129	129

Notes: This table reports results estimating equation $y_{it} = \beta HBP_{it} * field_i + \delta_i + \delta_t + \varepsilon_{it}$. The highest degree field indicator for each individual was assigned based on the field in which they obtained their highest degree. One is added to all variables before taking the logarithm to include years without publications or publications without citations. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A.2.4 Tables – Difference-in-Difference Results for 2000-2023

Table A.8: Publications and collaborations (2000-2023)

	(1)	(2)	(3)	(4)	(5)
	#pubs	1st author	last author	#co-authors	#citations
Panel A: full sample					
HBP	0.135***	-0.0108	0.0793***	0.281***	0.387***
	(0.0291)	(0.0176)	(0.0238)	(0.0506)	(0.0618)
LHS mean	3.2841	0.3371	1.0966	17.8502	198.8624
Observations	15,336	15,336	15,336	15,336	15,336
#authors	639	639	639	639	639
Panel B: junior scholars (junior faculties, graduate students)					
HBP	0.293***	0.0468	0.0331	0.599***	0.889***
	(0.0560)	(0.0367)	(0.0311)	(0.101)	(0.115)
LHS mean	1.4919	0.2815	0.2451	8.0044	77.6914
Observations	4,728	4,728	4,728	4,728	4,728
#authors	197	197	197	197	197
Panel C: female scholars					
HBP	0.115*	-0.00180	0.0682	0.185*	0.408***
	(0.0647)	(0.0404)	(0.0447)	(0.109)	(0.132)
LHS mean	2.2888	0.3088	0.7048	10.2448	104.1234
Observations	3,096	3,120	3,120	3,120	3,121
#authors	129	130	130	130	131

Notes: This table reports results estimating equation 1.1 on full sample (2000-2023). Outcome variables are log number of publications. One is added to all variables before taking the logarithm to include years without publications or publications without citations. #co-authors report the number of distinct co-authors per author-year. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.9: HBP participation and publications - full sample (2000-2023)

	(1)	(2)	(3)	(4)
Panel A: Journal Quality (Top neuroscience/CS outlets)				
	Top Neuro	Top CS	CS A*	CS A
HBP	0.0351***	0.00746	-0.00484	0.0116
	(0.0123)	(0.00820)	(0.00401)	(0.00716)
LHS mean	0.2193	0.0983	0.0372	0.0611
Panel B: Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0819***	0.155***	0.0872***	0.0675***
	(0.0262)	(0.0261)	(0.0184)	(0.0239)
LHS mean	2.2098	1.7704	0.6878	1.1629
Panel C: Probability of Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	-0.000727	0.0421***	0.00629	-0.00644
	(0.0101)	(0.0126)	(0.00520)	(0.00856)
LHS mean	0.2162	0.1989	0.0449	0.0828
Panel D: Topic Classification with the highest probability				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0450*	0.136***	0.0153	0.00915
	(0.0232)	(0.0240)	(0.0119)	(0.0189)
LHS mean	1.3659	1.0784	0.2915	0.5526
Panel E: Topic Classification in Top Neuro Journals				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0334***	0.0240***	-2.89e-05	0.00924
	(0.0121)	(0.00693)	(0.00183)	(0.00673)
LHS mean	0.2042	0.0563	0.0055	0.042
Observations	15,336	15,336	15,336	15,336
#authors	639	639	639	639

Notes: This table reports results estimating equation 1.1 on full sample (2000-2023). Outcome variables are log number of publications. One is added to all variables before taking the logarithm to include years without publications or publications without citations. Topic class with the highest probability contains 2,226 publications classified in multiple topics with equal maximum probabilities (2,216 cases with two topics, 10 cases with three topics), for which we count each topic as the max-likelihood topic (the results are qualitatively similar when analyzing disaggregated topic combinations). For those we count each topic equally as one. We also analyzed top CS outlets by topic classes, and none of the estimates are statistically significant nor economically meaningful (i.e., all estimates are very close to zero). Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.10: HBP participation and publications - junior sample (2000-2023)

	(1)	(2)	(3)	(4)
Panel A: Journal Quality (Top neuroscience/CS outlets)				
	Top Neuro	Top CS	CS A*	CS A
HBP	0.0246	0.0156	0.000614	0.0149
	(0.0192)	(0.0149)	(0.00583)	(0.0131)
LHS mean	0.0850	0.0444	0.0159	0.0286
Panel B: Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.137***	0.354***	0.147***	0.126***
	(0.0485)	(0.0466)	(0.0364)	(0.0442)
LHS mean	1.0116	0.9323	0.4092	0.4847
Panel C: Probability of Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0171	0.157***	0.0115	0.0207
	(0.0216)	(0.0250)	(0.0109)	(0.0184)
LHS mean	0.1546	0.1747	0.0473	0.0611
Panel D: Topic Classification with the highest probability				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0241	0.284***	0.0219	0.0273
	(0.0419)	(0.0414)	(0.0234)	(0.0330)
LHS mean	0.5539	0.5598	0.1728	0.2058
Panel E: Topic Classification in Top Neuro Journals				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0187	0.0254***	-0.00181	-0.00550
	(0.0192)	(0.00906)	(0.00284)	(0.0107)
LHS mean	0.0759	0.0279	0.0038	0.0192
Observations	4,728	4,728	4,728	4,728
#authors	197	197	197	197

Notes: This table reports results estimating equation 1.1 "junior-at-the-start" sample (2000-2023). The sub-sample consists of HBP participants who entered with the status junior (under-/graduates, junior faculty). Outcome variables are log number of publications. One is added to all variables before taking the logarithm to include years without publications or publications without citations. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.11: HBP participation and publications - female sample (2000-2023)

	(1)	(2)	(3)	(4)
Panel A: Journal Quality (Top neuroscience/CS outlets)				
	Top Neuro	Top CS	CS A*	CS A
HBP	0.0324	-0.00293	-0.00318	-0.00292
	(0.0235)	(0.0131)	(0.00684)	(0.0116)
LHS mean	0.1357	0.0714	0.0284	0.0429
Panel B: Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0538	0.173***	0.0997***	0.0297
	(0.0600)	(0.0568)	(0.0363)	(0.0538)
LHS mean	1.5497	1.2846	0.4503	0.8282
Panel C: Probability of Topic Classification				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	-0.00629	0.0466	0.0116*	-0.0166
	(0.0265)	(0.0301)	(0.00689)	(0.0218)
LHS mean	0.2046	0.1704	0.0273	0.0892
Panel D: Topic Classification with the highest probability				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0245	0.136***	0.0467	-0.0184
	(0.0508)	(0.0501)	(0.0285)	(0.0391)
LHS mean	0.9306	0.7955	0.1935	0.3727
Panel E: Topic Classification in Top Neuro Journals				
	Neurobio	Neurotech	AI-Robotics	Clinical
HBP	0.0345	0.0224*	0.000276	-0.00997
	(0.0227)	(0.0129)	(0.00334)	(0.0154)
LHS mean	0.1253	0.0398	0.0036	0.0297
Observations	3,096	3,096	3,096	3,096
#authors	129	129	129	129

Notes: This table reports results estimating equation 1.1 female sample (2000-2023). Outcome variables are log number of publications. One is added to all variables before taking the logarithm to include years without publications or publications without citations. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B

Appendix to Chapter 2

Collaborating Neuroscience Online – The Case of the Human Brain Project Forum

B.1 Figures

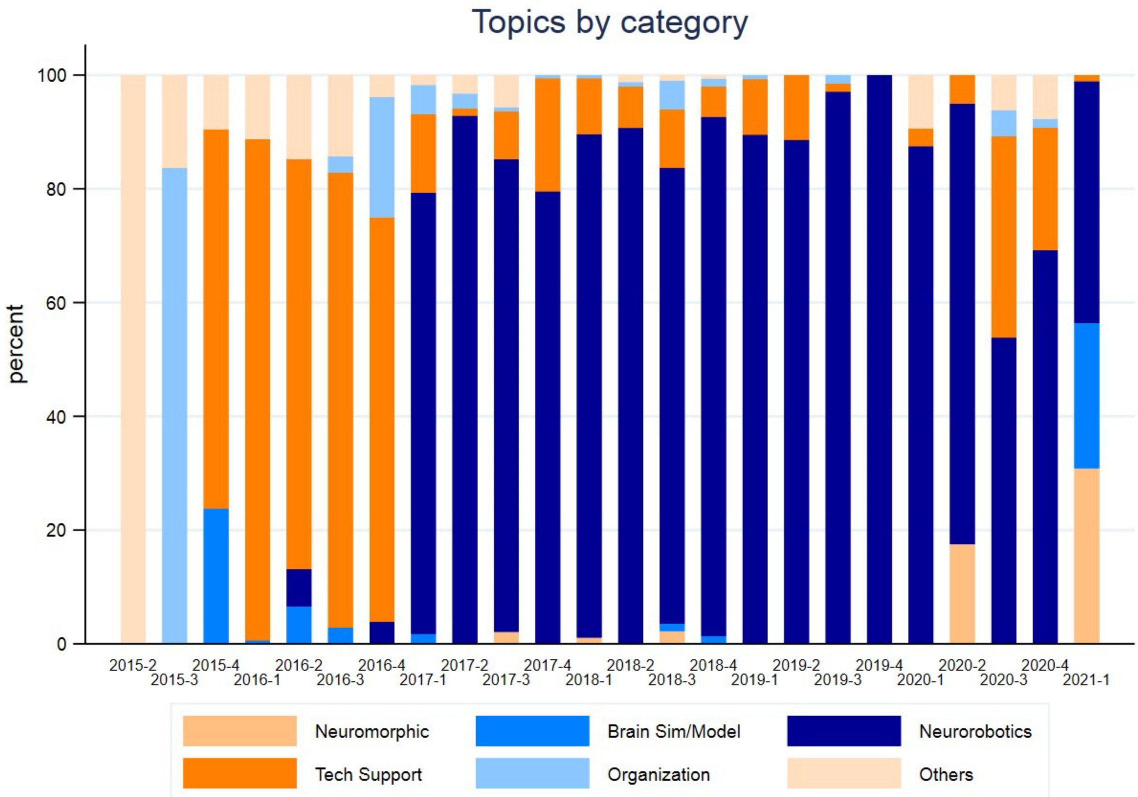


Figure B.1: HBP Forum usage: Topics by content category

Notes: This histogram depicts the proportion of topic categories discussed on the HBP Forum overtime in quarterly units. The major content categories are created based on the official category tags generated on the HBP Forum. We grouped similar tags into major groups in line with the main subproject areas of the HBP. Among the six content categories, Neurorobotics, Neuromorphic, and Brain Simula- tion/Model are categories closely tied with HBP platform-based subproject areas.

Human Brain Project Forum

[CLOSED] Messages on the /husky/odom topic

Neurorobotics

vonarnim

Mar 16

Hi Gabija

So the type of the odom message is here http://docs.ros.org/en/noetic/api/nav_msgs/html/msg/Odometry.html 7 and the Orientation is of TwistWithCovariance type that is defined here http://docs.ros.org/en/noetic/api/geometry_msgs/html/msg/TwistWithCovariance.html 3 which is not a Pose message that you expect (read the message type).

You can see what the message type is by doing

```
rostopic info /husky/odom
```

and then it will show you the right message type

Type: nav_msgs/OdometryPublishers:

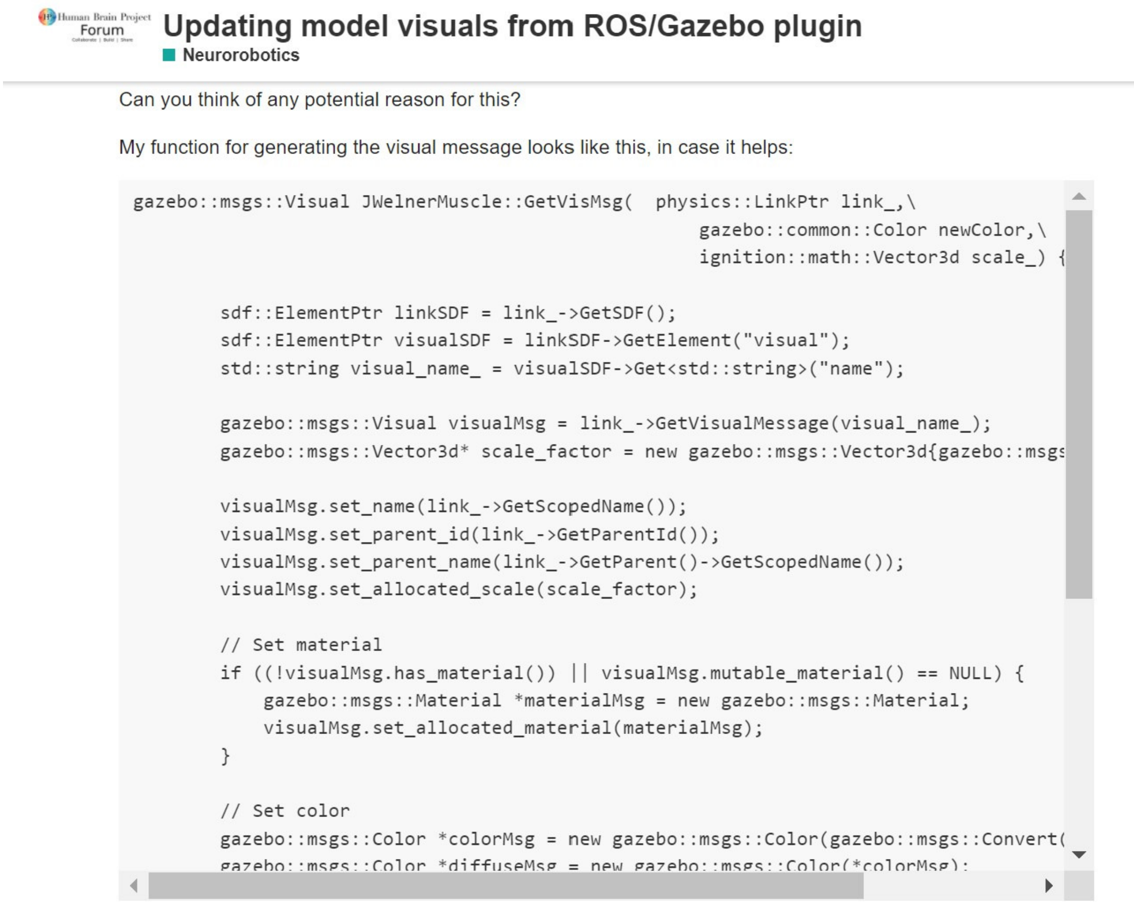
- /gazebo (http://127.0.0.1:43339/ 1)

Subscribers: None

Manos

Figure B.2: Snapshot example of an answer to questions on neurorobotics (with light code)

Notes: This figure provides an example of a question with light code discussed on the HBP Forum.



Human Brain Project Forum

Updating model visuals from ROS/Gazebo plugin

Neurorobotics

Can you think of any potential reason for this?

My function for generating the visual message looks like this, in case it helps:

```
gazebo::msgs::Visual JWelnerMuscle::GetVisMsg(  physics::LinkPtr link_,\
                                                    gazebo::common::Color newColor,\
                                                    ignition::math::Vector3d scale_) {

        sdf::ElementPtr linkSDF = link_->GetSDF();
        sdf::ElementPtr visualSDF = linkSDF->GetElement("visual");
        std::string visual_name_ = visualSDF->Get<std::string>("name");

        gazebo::msgs::Visual visualMsg = link_->GetVisualMessage(visual_name_);
        gazebo::msgs::Vector3d* scale_factor = new gazebo::msgs::Vector3d{gazebo::msgs

        visualMsg.set_name(link_->GetScopedName());
        visualMsg.set_parent_id(link_->GetParentId());
        visualMsg.set_parent_name(link_->GetParent()->GetScopedName());
        visualMsg.set_allocated_scale(scale_factor);

        // Set material
        if ((!visualMsg.has_material()) || visualMsg.mutable_material() == NULL) {
            gazebo::msgs::Material *materialMsg = new gazebo::msgs::Material;
            visualMsg.set_allocated_material(materialMsg);
        }

        // Set color
        gazebo::msgs::Color *colorMsg = new gazebo::msgs::Color(gazebo::msgs::Convert(
        gazebo::msgs::Color *diffuseMsg = new gazebo::msgs::Color(*colorMsg);
```

Figure B.3: Snapshot example of a question on neurorobotics (with a block of code)

Notes: This figure provides an example of a question with a block of code discussed on the HBP Forum.

Human Brain Project
Forum
Collaborate | Build | Share

Discrepancies in connectivity data between paper and NMC portal data

Publications **Markram et al., 2015**

henhok Nov '16

Hello all,

We've been successfully using the connectivity data provided at the NMC portal for constructing neural networks of our own. After going through the paper and looking hard at the data, we however haven't been able to answer some questions regarding differences between the paper and the data at the NMC portal (https://bbp.epfl.ch/nmc-portal/documents/10184/7288948/pathways_anatomy_factsheets_simplified.json 2). Specifically:

- Issue 1: In the paper for the total number of synapses in the microcircuit, you provide the estimate ~37 synapses (36.7 ± 4.2 million synapses). However if you take the sum of total_synapse_count 's, you get a number close to 60 million (59 245 093).
- Issue 2: When plotting the connection_probability variable as a heatmap similarly to Figure 7B in the paper, two spots stand out: L4_BTC:L4_ChC with 40% probability and L4_NGC:L4_ChC with 100% probability. These spots are blank in the paper. Why have they been weeded out from Fig 7B but left in the json data?
- Issue 3: There are also some connections in Figure 7B that don't have a connection_probability set in the json data, eg. L6_NGC:L6_SBC. Why are they missing from the json data?

We'd be really happy if someone took the time to clarify these issues for us.

Cheers,
Henri Hokkanen
PhD student, Univ.of Helsinki

Figure B.4: Snapshot example of an organizational issue discussed (without code)

Notes: This figure provides an example of a question without code discussed on the HBP Forum. More examples on the typical interactions between users within each content categories are available upon request.

C

Appendix to Chapter 3

From Tinkering to Inventing – FabLabs as Catalysts of Innovations

C.1 Figures

C.1.1 Figures – Descriptive Analyses

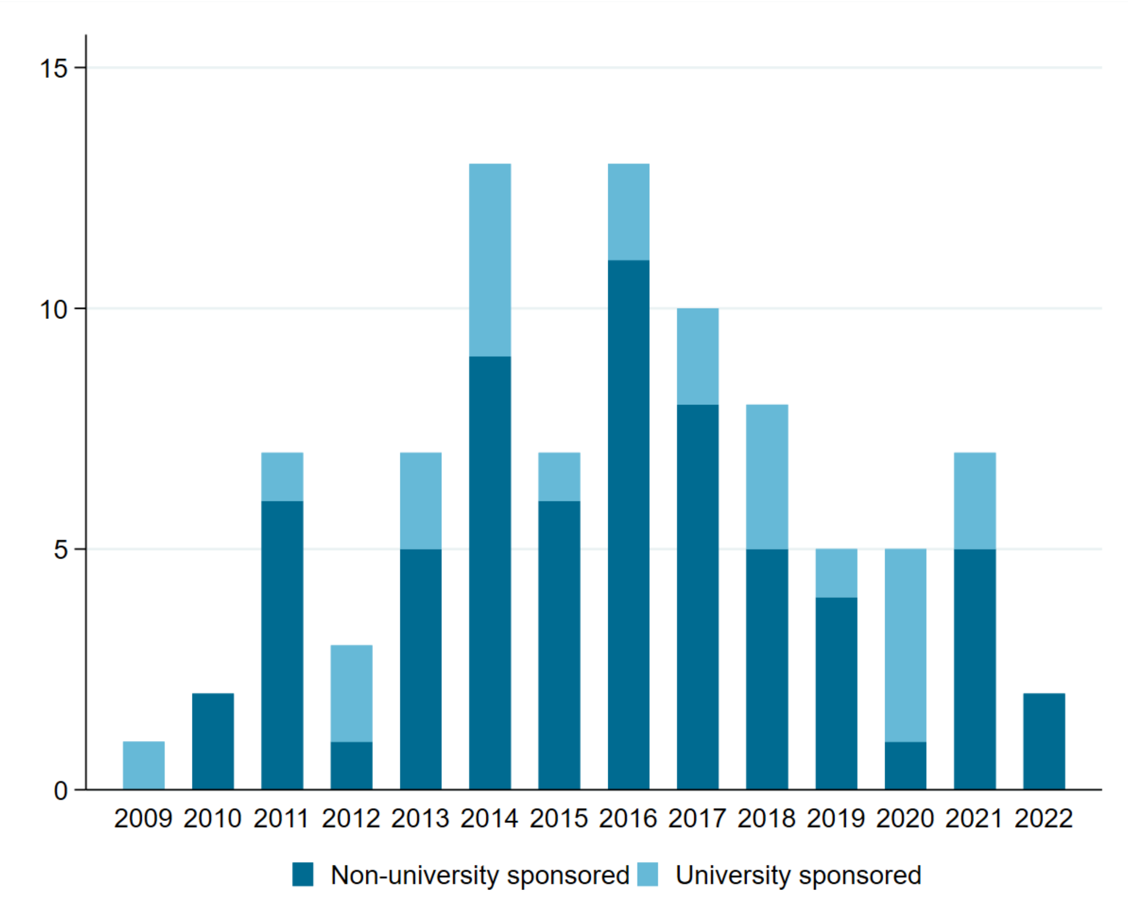


Figure C.1: FabLab establishment by founding year and university integration

Notes: This figure illustrates the development of FabLab establishments over time, distinguishing between those integrated into a university and those that are not.

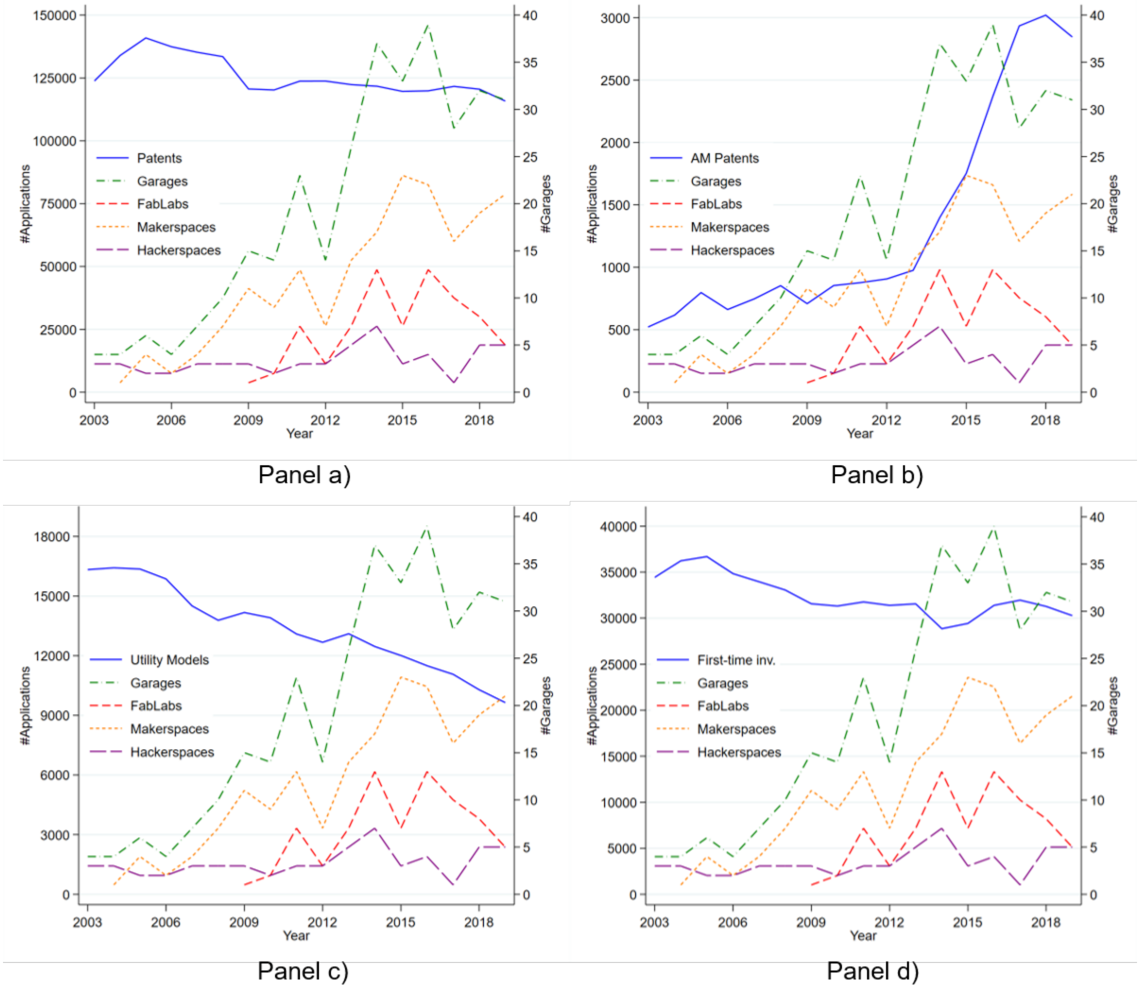


Figure C.2: Trends in patenting activity and community workshops across NUTS3 (2003–2019)

Notes: This figure presents the temporal development of patenting activity and community workshops across all NUTS3 from 2003 to 2019. Panel (a) shows the total number of patent applications, while Panel (b) focuses on patent applications related to AM. Panel (c) displays the number of utility model applications, and Panel (d) illustrates the number of applications filed by first-time inventors.

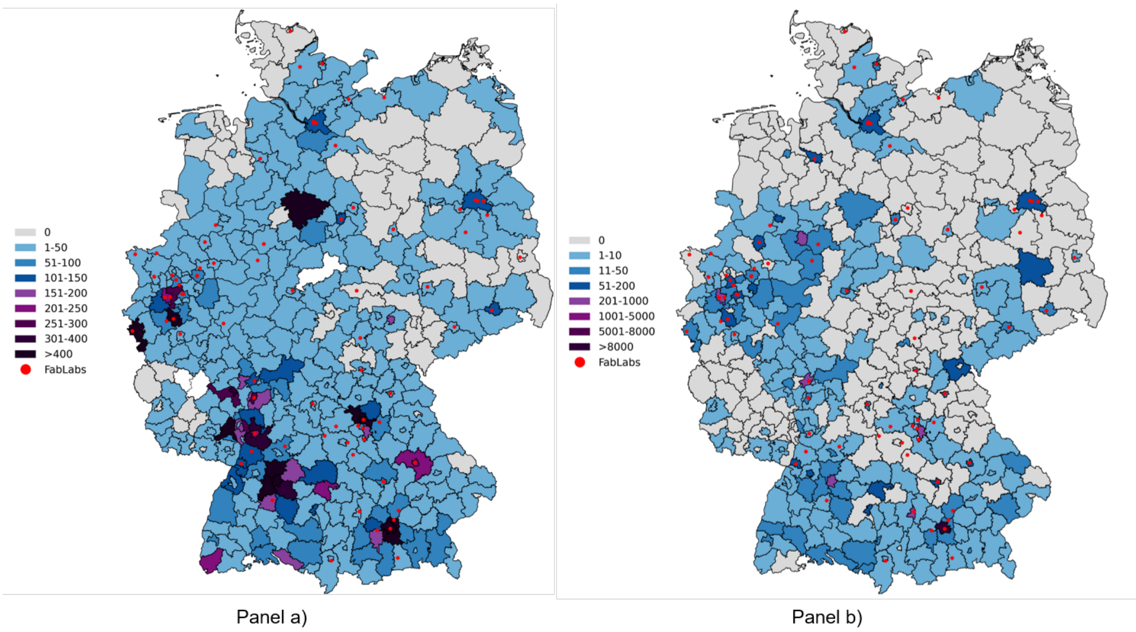


Figure C.3: Regional variation in patenting activity per capita in NUTS3 regions

Notes: This figure illustrates the geographic distribution of patenting activity in NUTS3 regions from 2003 to 2019, normalized by population. Panel (a) displays the number of patent applications per capita, while Panel (b) presents the number of utility model applications per 100 inhabitants.

C.1.2 Figures – Heterogeneous Effects

Figures – University-Integrated FabLabs

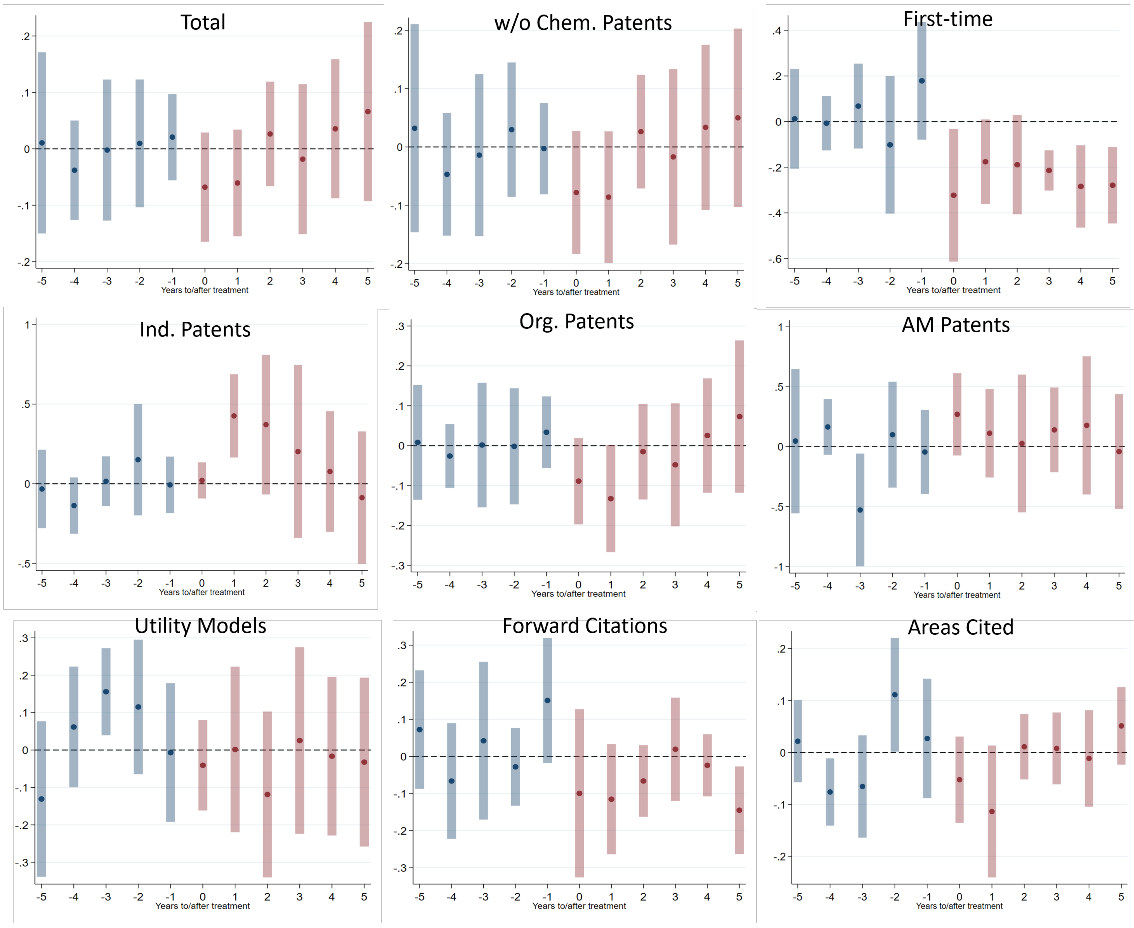


Figure C.4: Event study estimates – university-integrated FabLabs (FUA-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables, focusing on university-integrated FabLabs. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included.

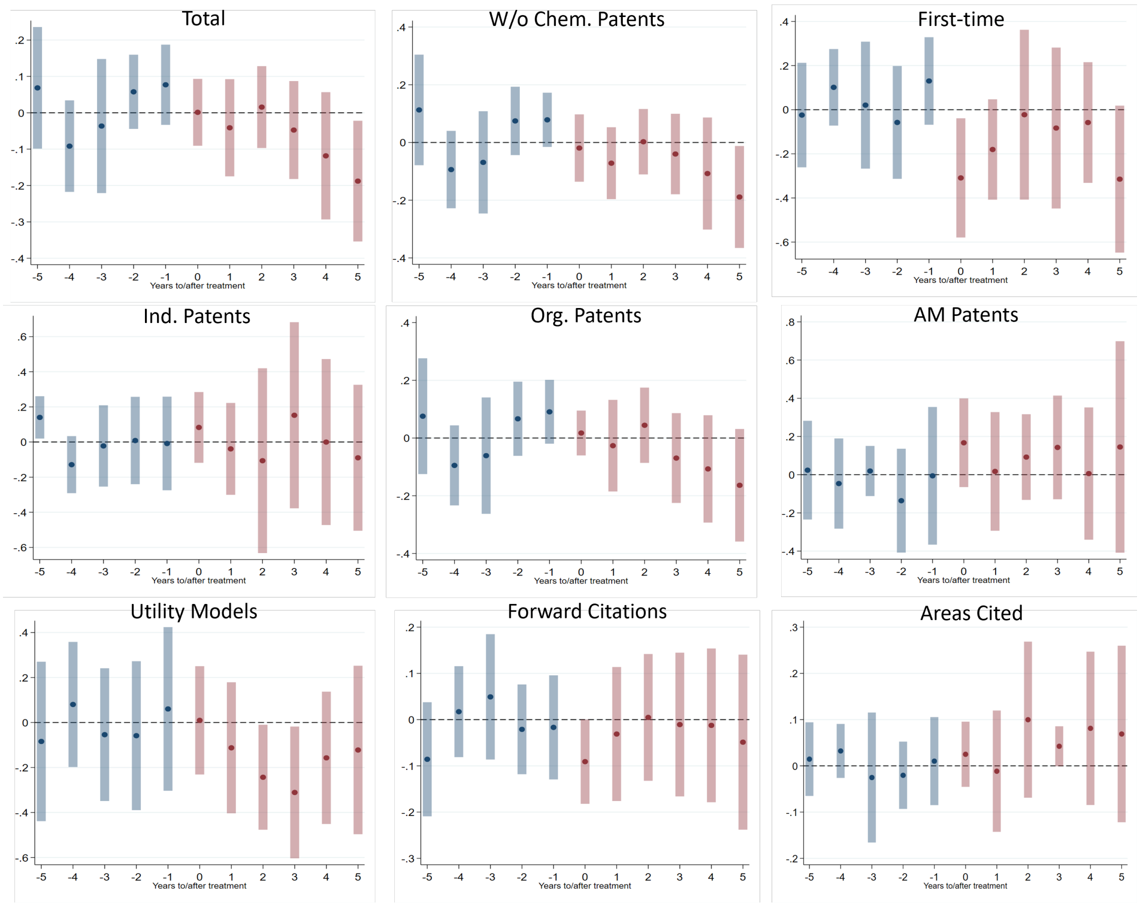


Figure C.5: Event study estimates – university-integrated FabLabs (NUTS3-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables, focusing on university-integrated FabLabs. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included.

Figures – Below-Median Subsample

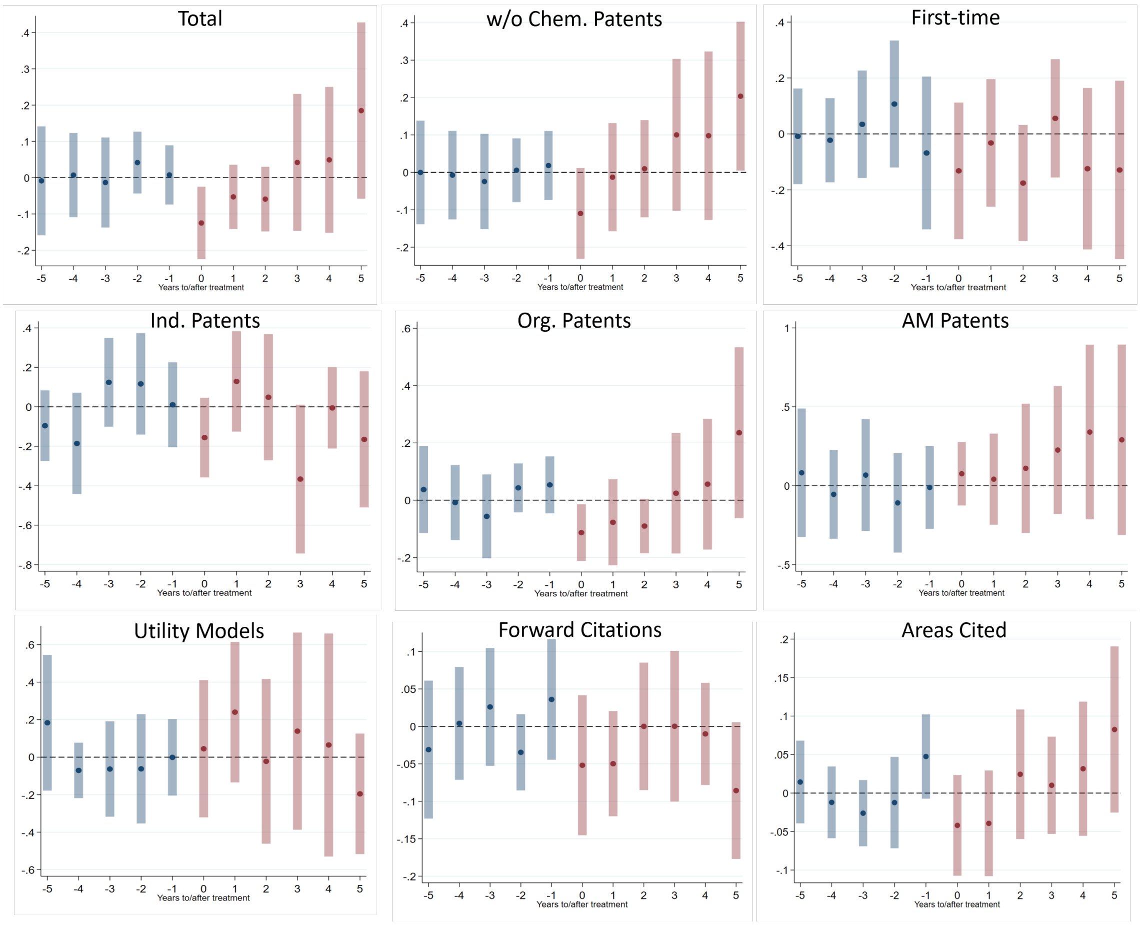


Figure C.6: Event study estimates – below-median subsample (FUA-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables in the below-median subsample. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included.

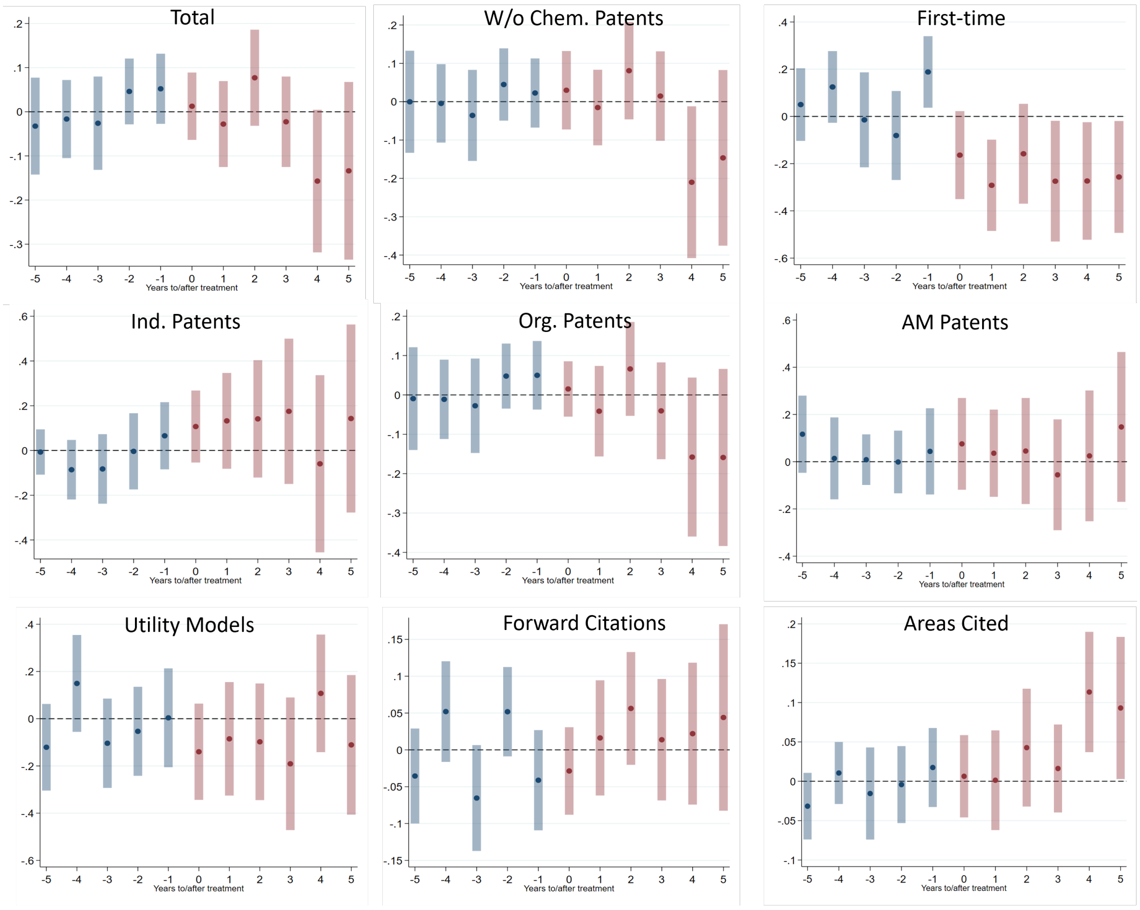


Figure C.7: Event study estimates – below-median subsample (NUTS3-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables in the below-median subsample. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included.

Figures – Above-Median Subsample

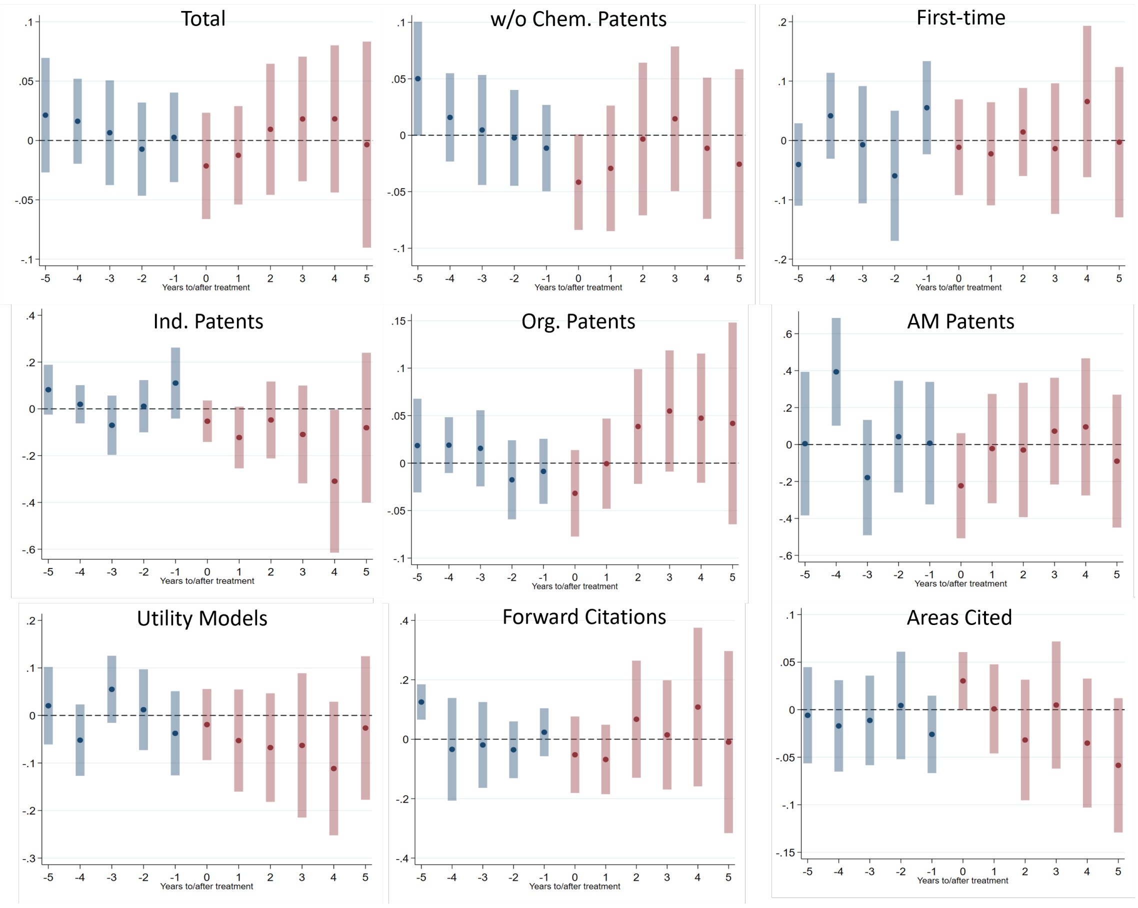


Figure C.8: Event study estimates – above-median subsample (FUA-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables in the above-median subsample. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included.

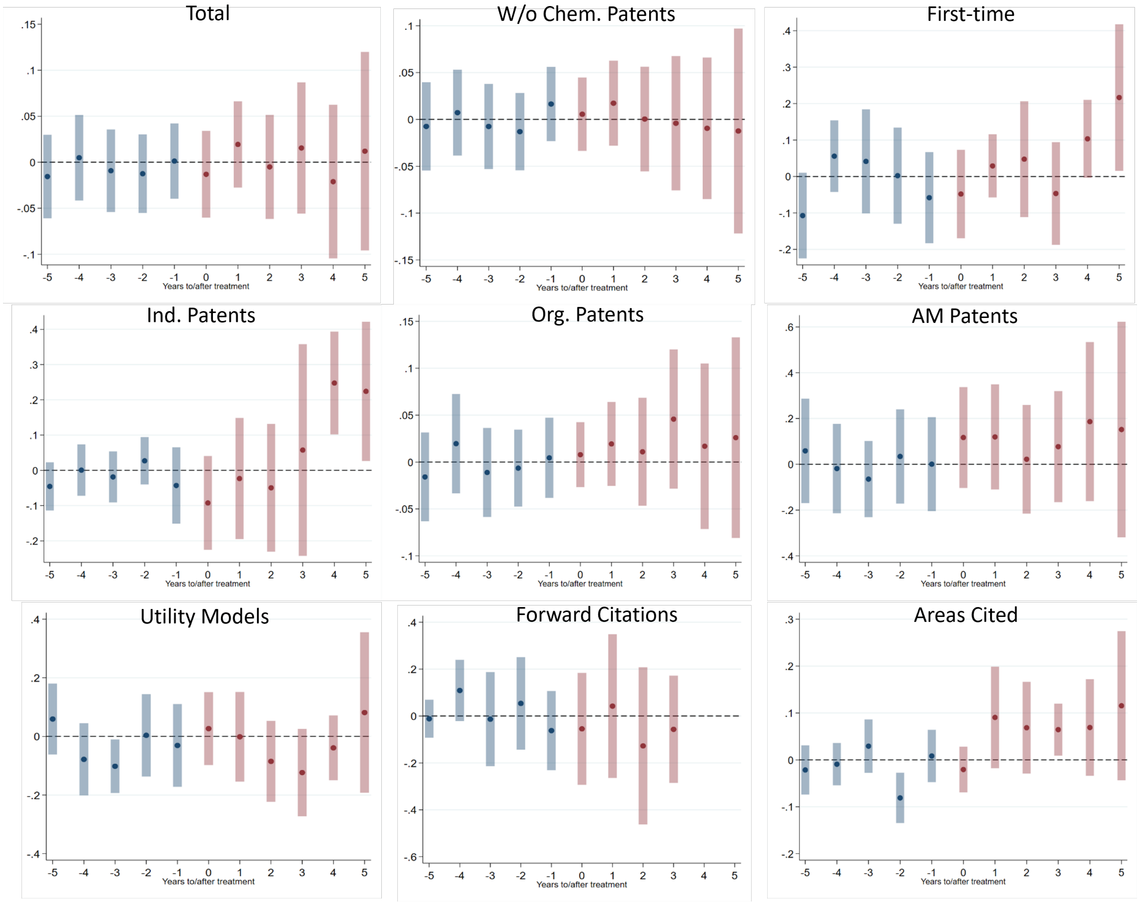


Figure C.9: Event study estimates – above-median subsample (NUTS3-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables in the above-median subsample. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included.

Figures – Urban vs. Rural Effects

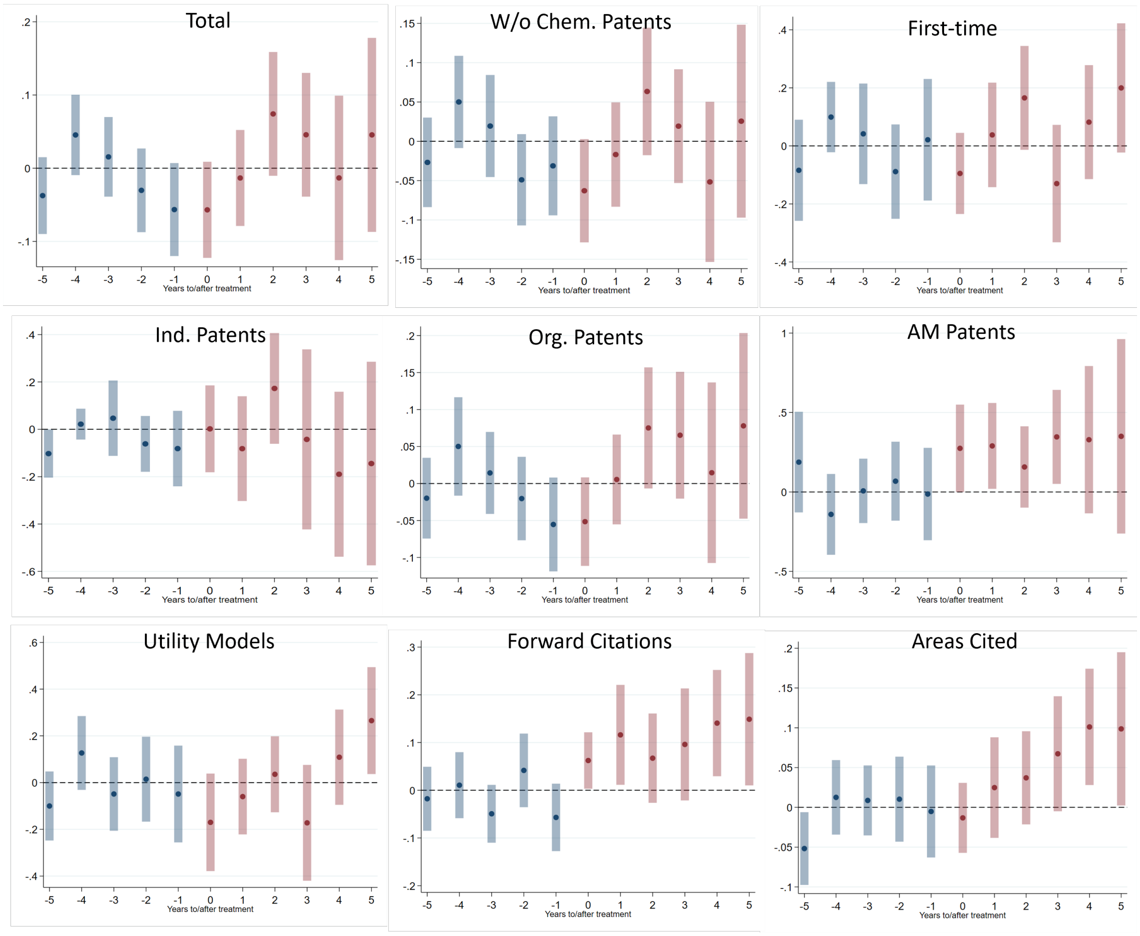


Figure C.10: Event study estimates – urban subsample (NUTS3-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables in the urban subsample. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included.

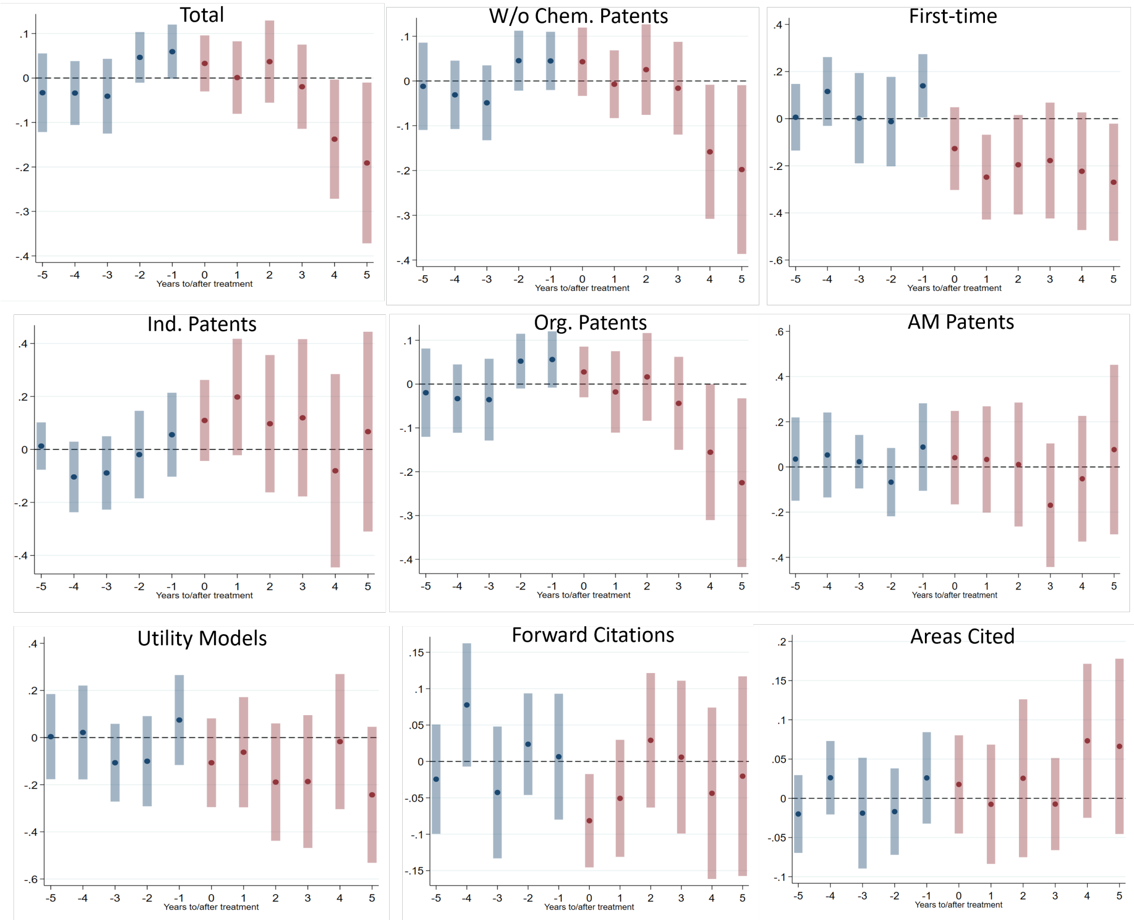


Figure C.11: Event study estimates – rural subsample (NUTS3-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables in the rural subsample. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included.

Figures – Pre-Existing Community Workshops (Rural NUTS3)

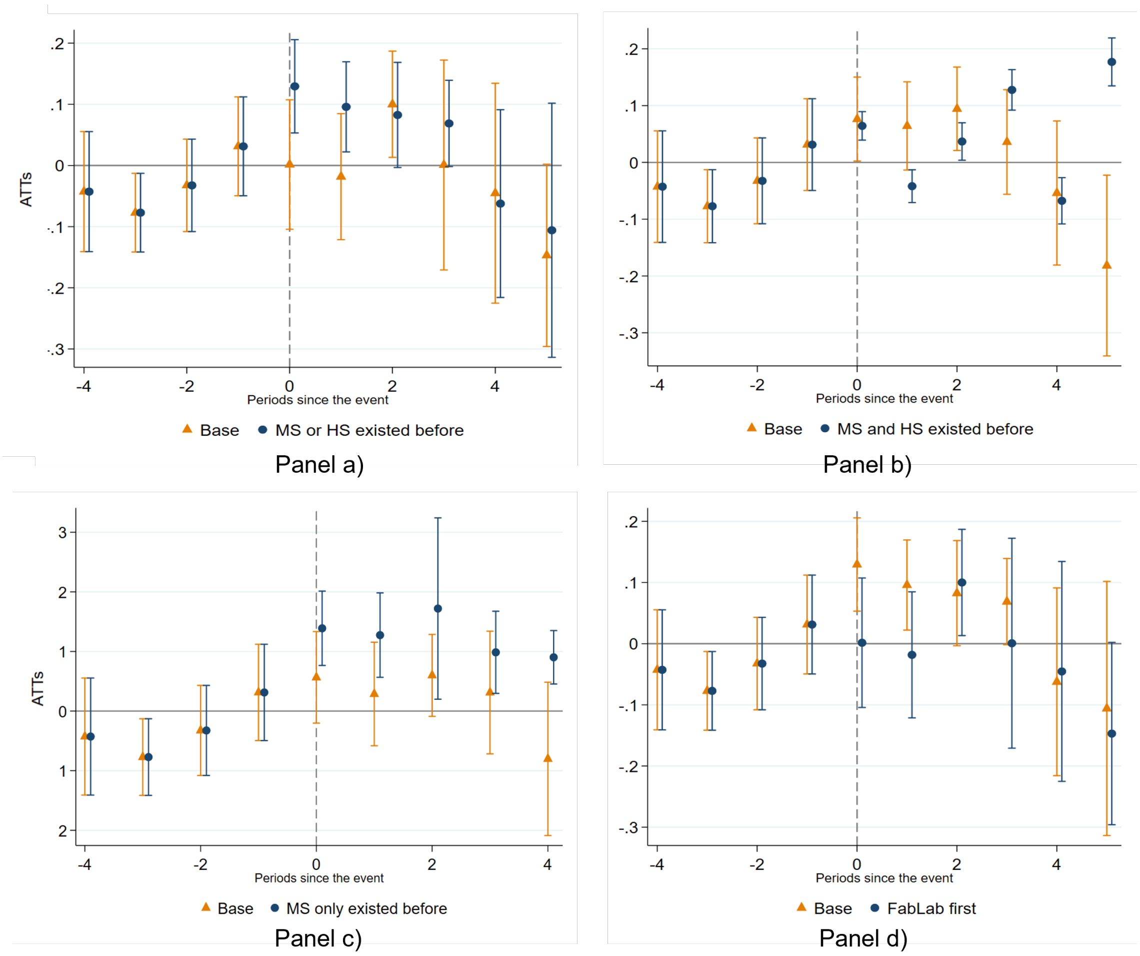


Figure C.12: Event study estimates – total patent applications (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log number of total patent applications, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

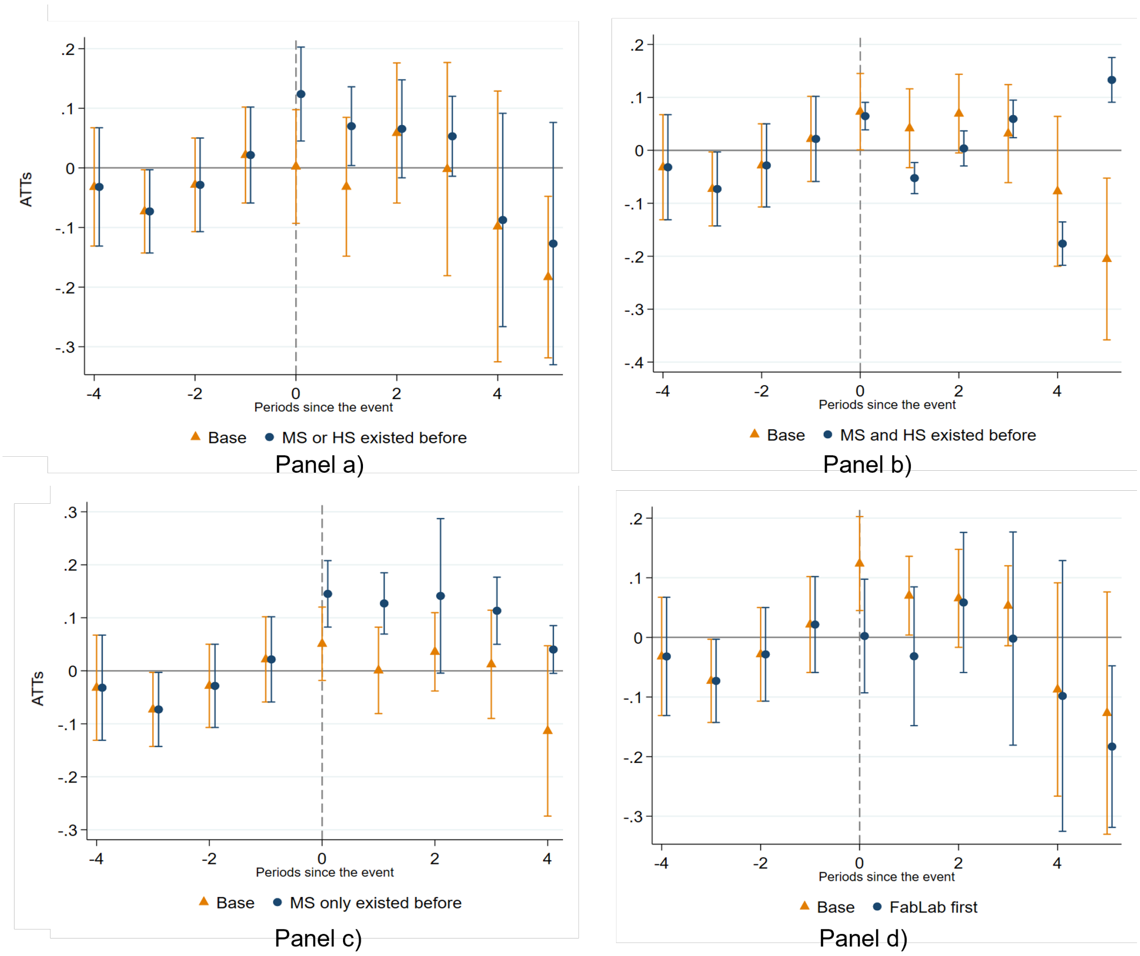


Figure C.13: Event study estimates – non-chemical patent applications (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log number of non-chemical patent applications, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

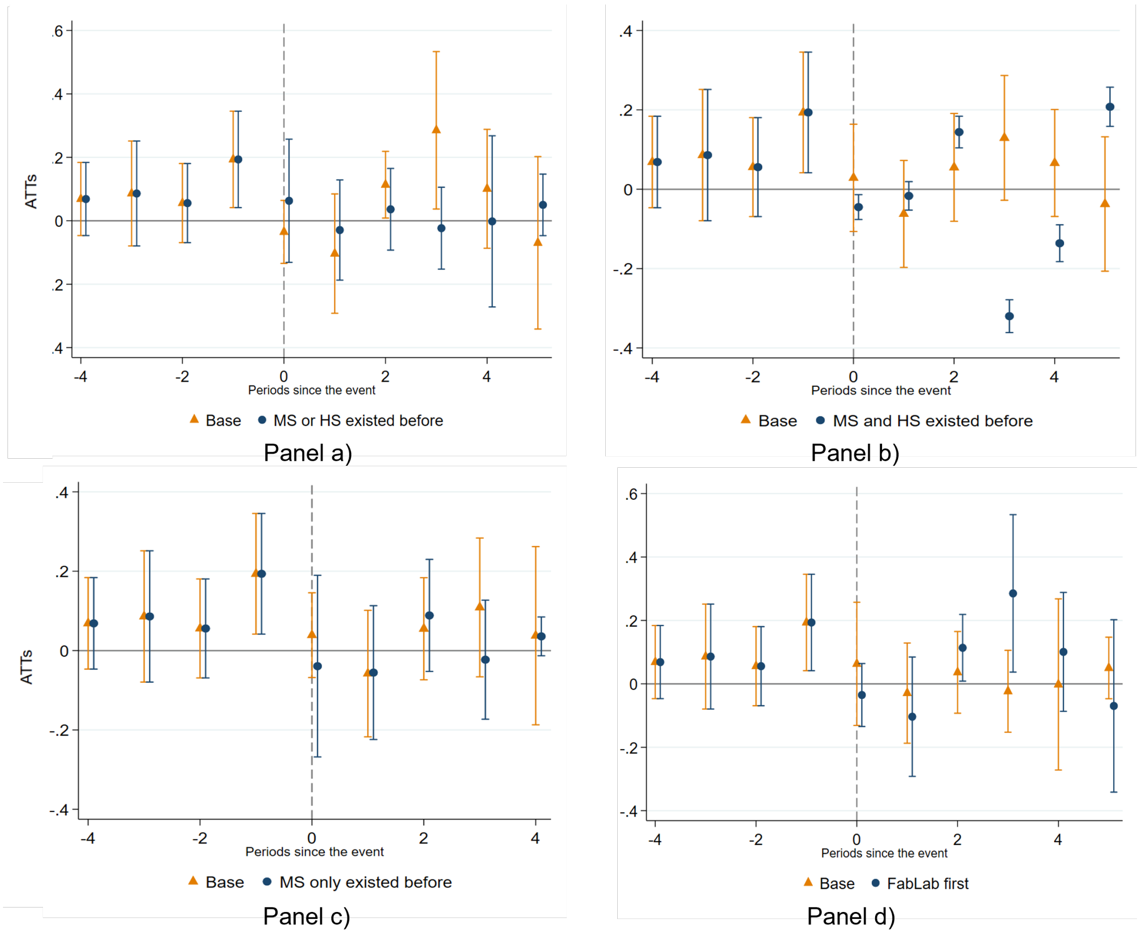


Figure C.14: Event study estimates – applications by first-time inventors (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log number of patent applications by first-time inventors, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

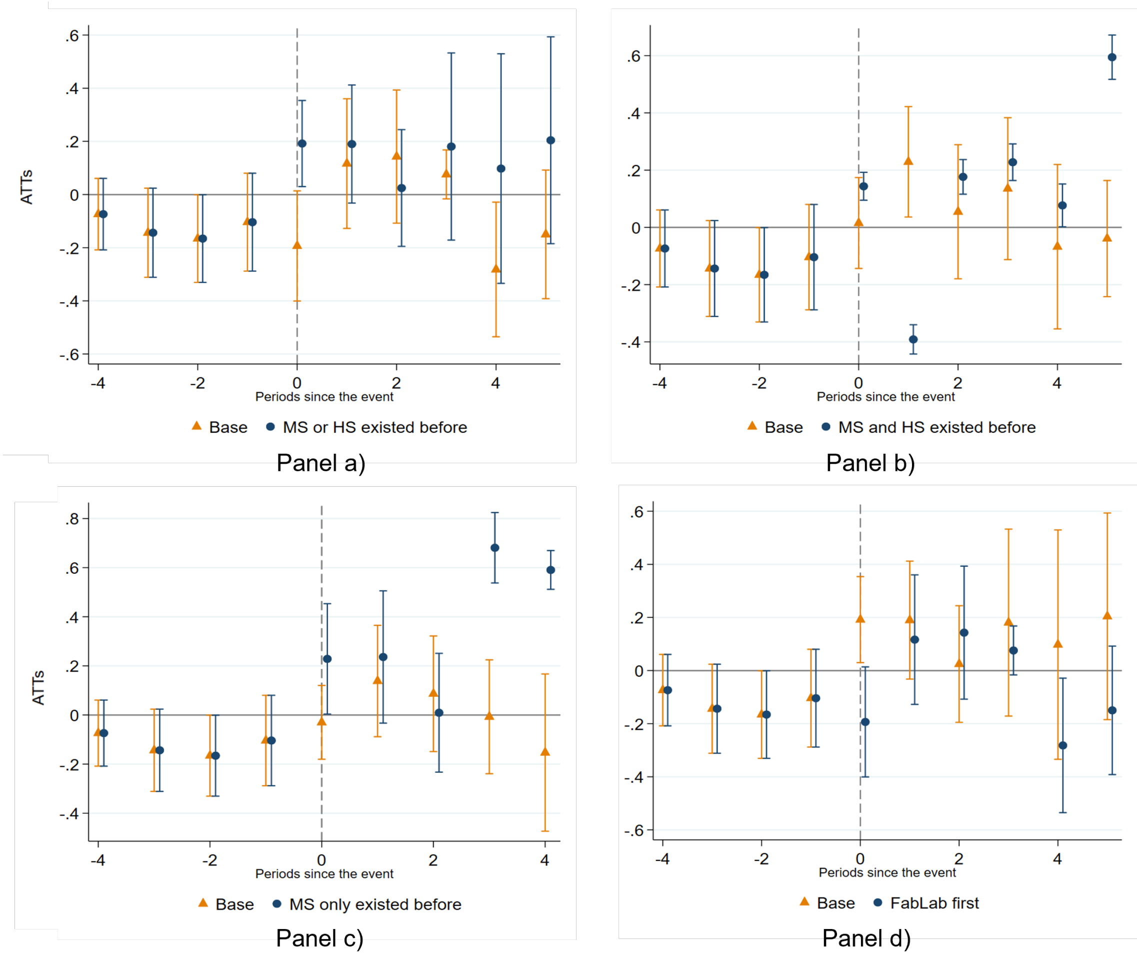


Figure C.15: Event study estimates – applications by individual applicants (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log number of patent applications by individual applicants, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

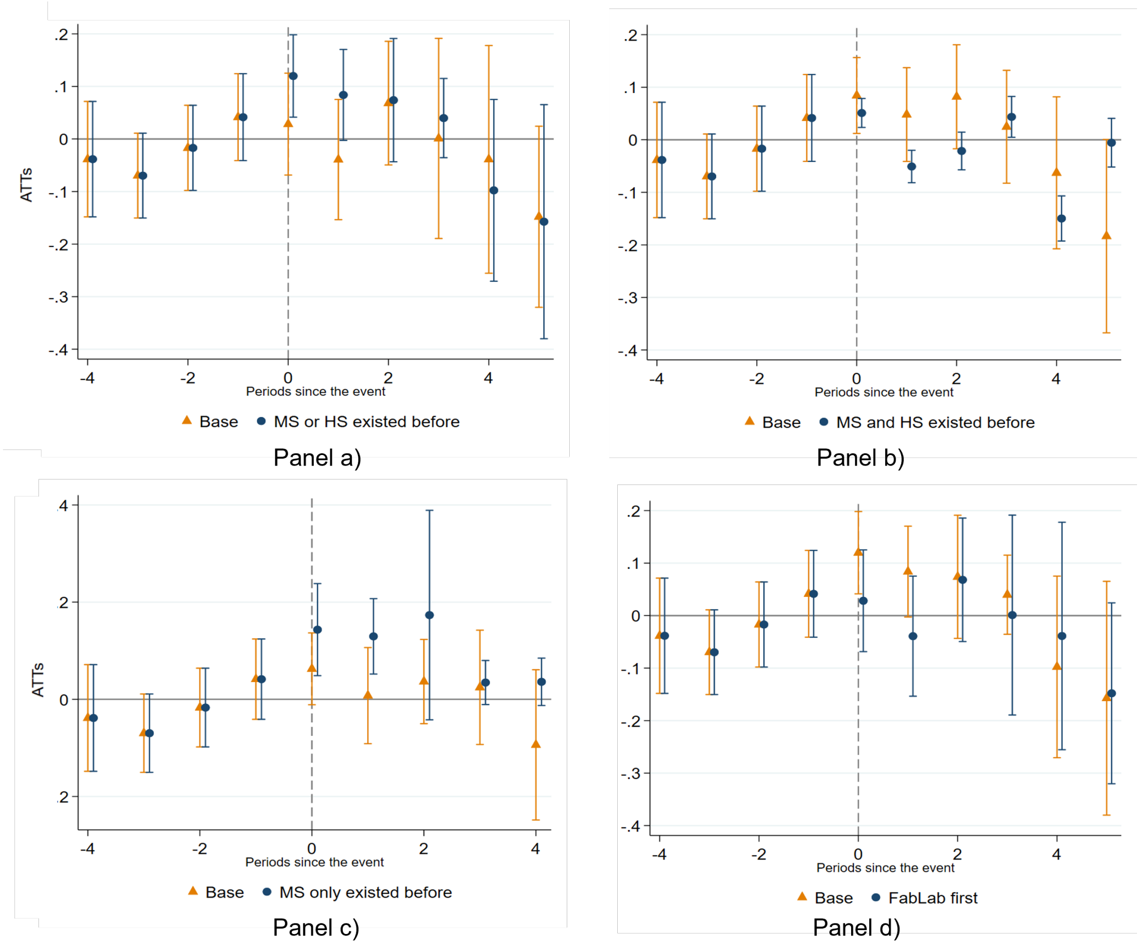


Figure C.16: Event study estimates – applications by organizational applicants (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log number of patent applications by organizational applicants, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

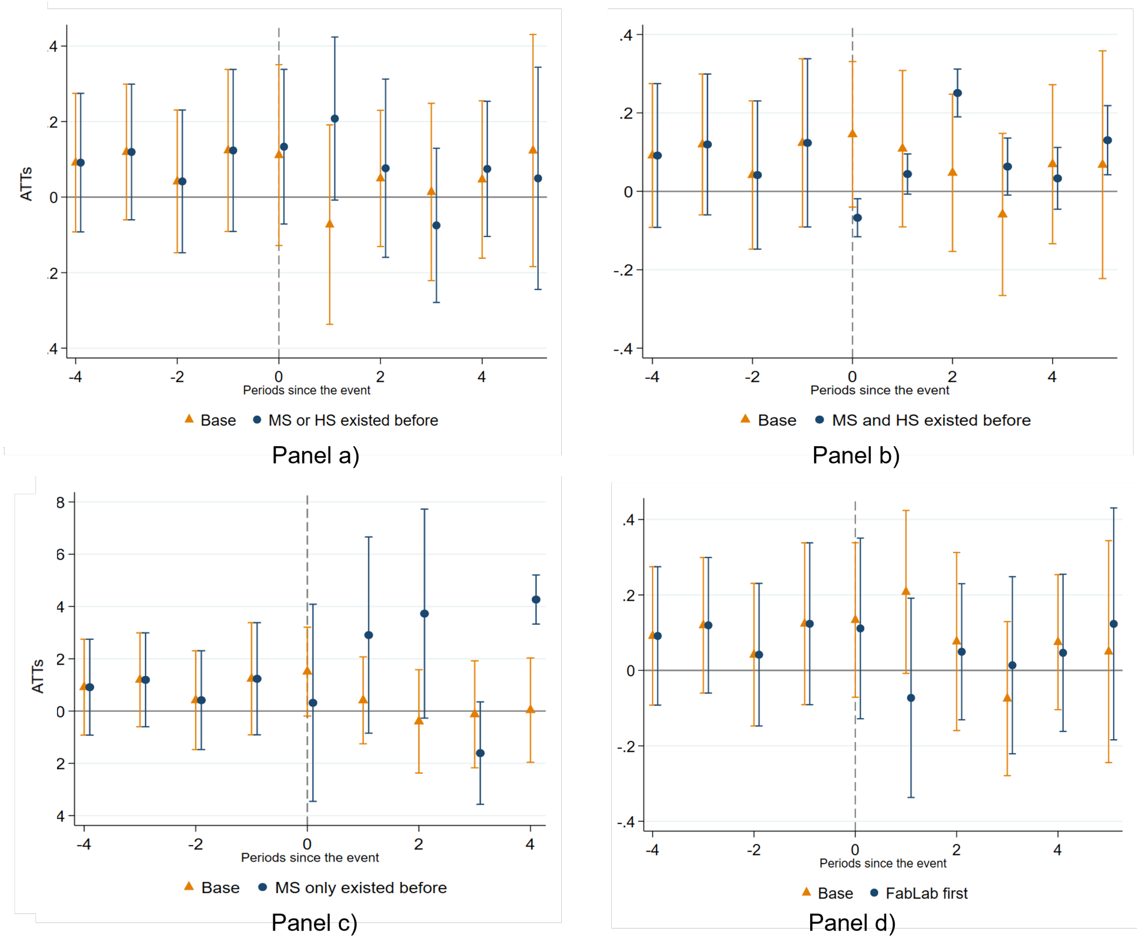


Figure C.17: Event study estimates – AM-related patent applications (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log number of AM-related patent applications, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

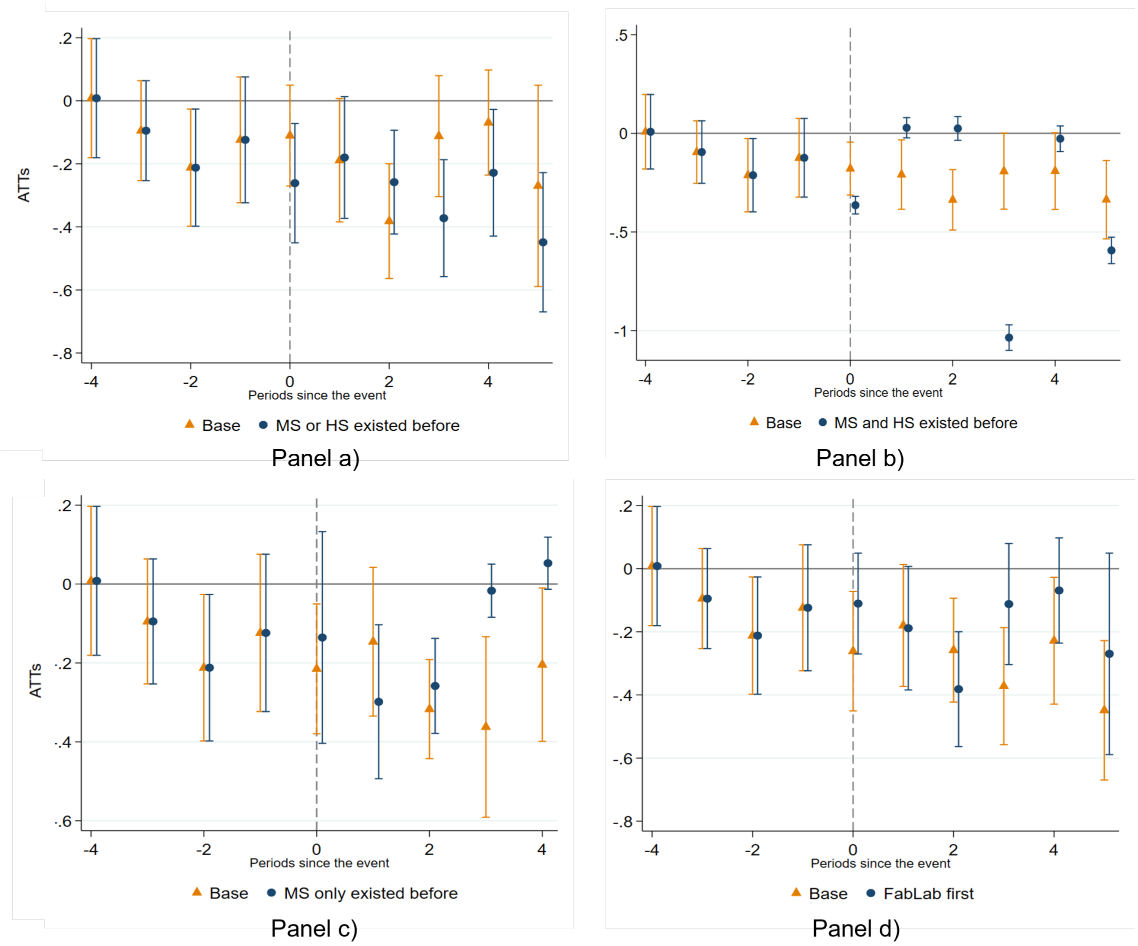


Figure C.18: Event study estimates – utility model applications (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log number of utility model applications, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

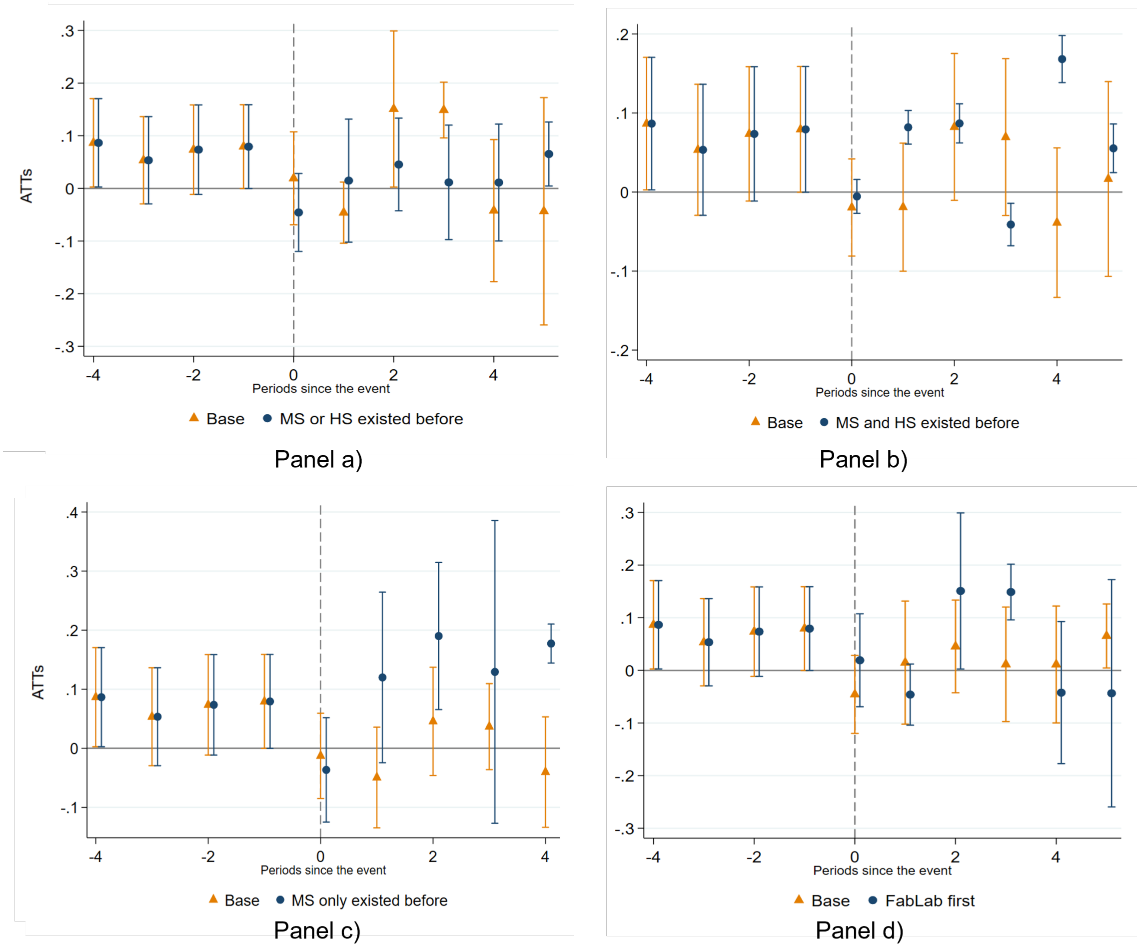


Figure C.19: Event study estimates – average number of forward citations (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log of the average number of forward citations, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

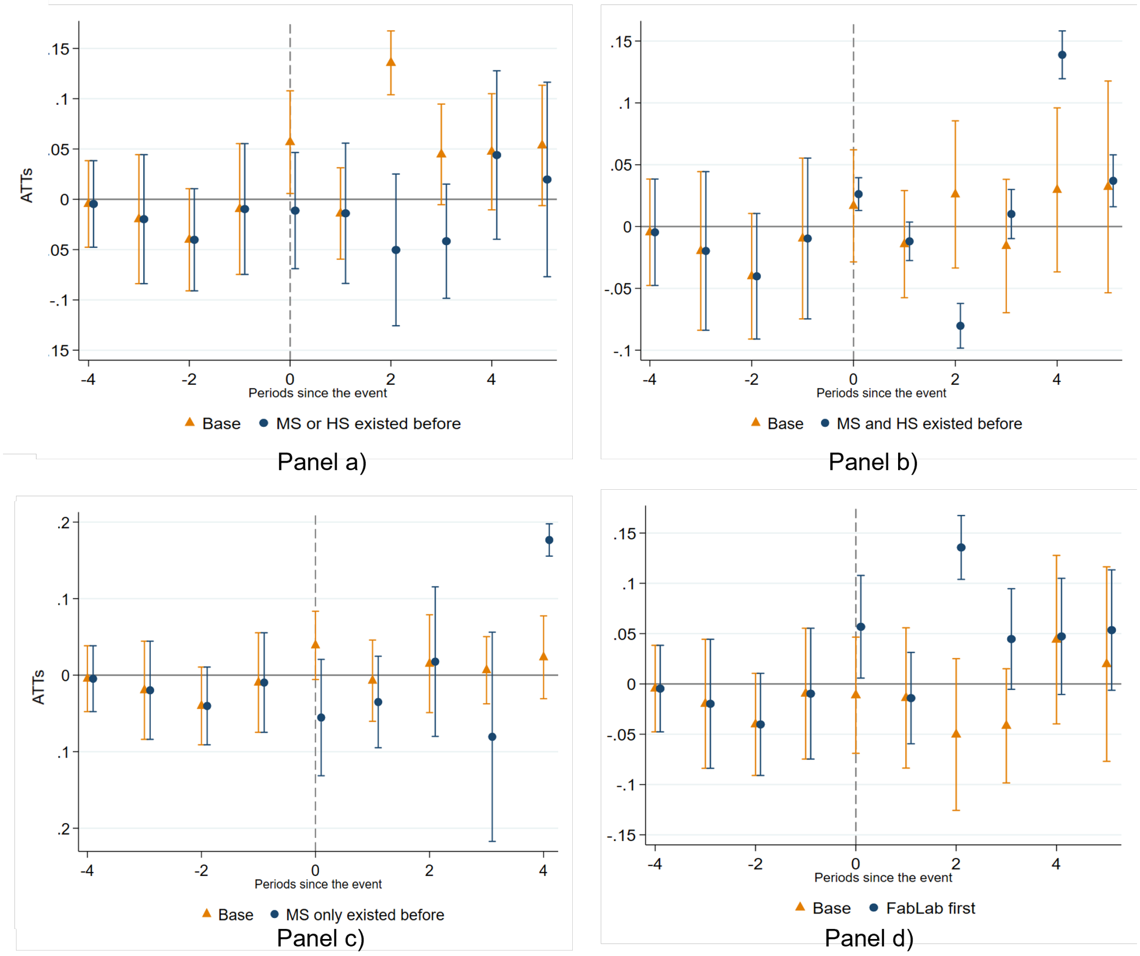


Figure C.20: Event study estimates – average number of areas cited (NUTS3-Year)

Notes: This figure presents event study estimates of the heterogeneous effects of FabLabs in rural NUTS3 regions with pre-existing community workshops on the log of the average number of areas cited, using the method of Borusyak, Jaravel, and Spiess (2024). The unit of analysis is the NUTS3-year, with not-yet-treated and never-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are also accounted for.

C.1.3 Figures – Robustness Checks

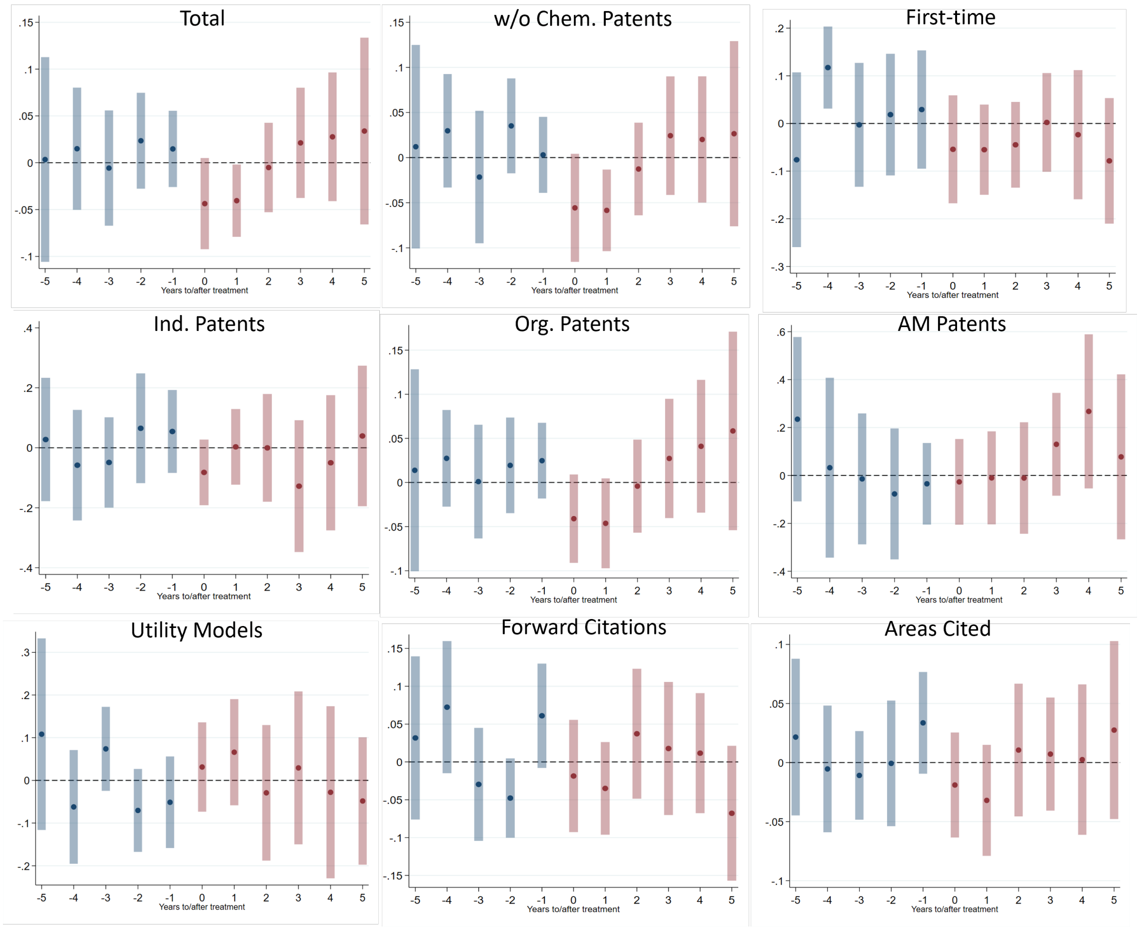


Figure C.21: Event study estimates for the 2010-2019 subsample (FUA-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables for the years 2010-2019. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included.

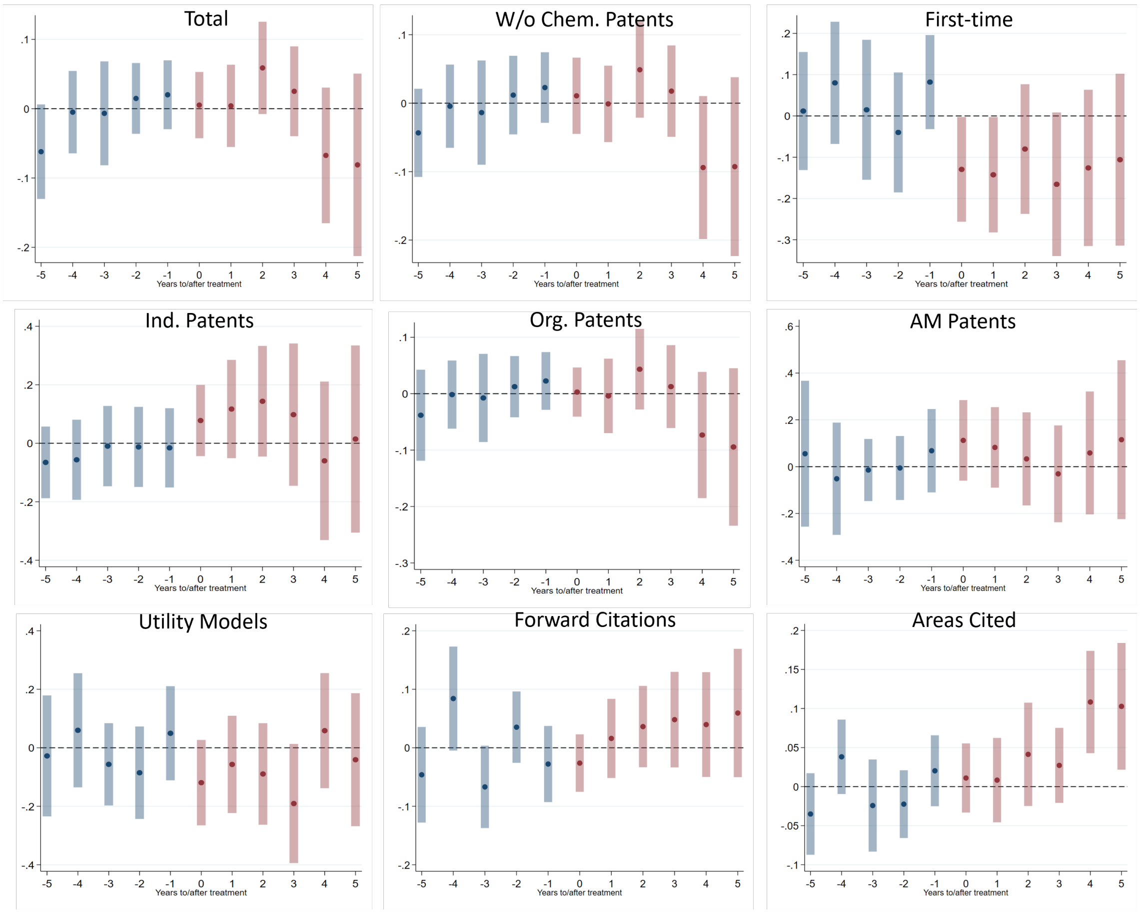


Figure C.22: Event study estimates for the 2010-2019 subsample (NUTS3-Year)

Notes: This figure presents the event study results using the CS method for all log-transformed dependent variables for the years 2010-2019. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included.

C.2 Tables

C.2.1 Tables – Descriptive Analysis

Table C.1: Summary statistics – treated regions

	N	Mean	Median	SD	Min.	Max.
Panel A: dependent variables treated group (FUA dataset)						
Patents	629	1,432.931	837.128	1,724.543	21.738	9,724.714
NoChem. Pat.	629	1,158.441	726.301	1,468.494	16.554	8,927.706
First-time	629	195.595	117.790	231.723	2.5	1,374.501
Pats. by ind.	629	176.133	72.071	280.820	0	2,176.872
Pats. by orgs.	629	1,229.265	713.829	1,488.931	17.163	9,142
Forward cit.	629	9.172	9.381	4.246	0.666	26.831
Areas cited	629	5.999	5.952	1.174	2.943	9.521
AM patents	629	18.148	6.117	37.040	0	373.427
Utility models	629	236.472	75	513.473	2	3,620
Panel B: dependent variables treated group (NUTS3 dataset)						
Patents	1,054	237.65	175.22	222.93	3.50	1,504
NoChem. Pat.	1,054	199.18	151.92	185.11	3.50	1,353
First-time	1,054	25.06	12.19	31.10	0.00	234
Pats. by ind.	1,054	30.62	17.40	39.96	0.00	351
Pats. by orgs.	1,054	202.16	144.41	191.36	3.50	1,333
Forward cit.	1,054	8.89	8.44	5.18	0.41	51
Areas cited	1,054	5.08	4.82	1.35	1.88	14
AM patents	1,054	2.88	0.75	5.52	0.00	63
Utility models	1,023	24.28	11.00	32.65	1.00	215
Panel C: explanatory variables treated group (FUA dataset)						
Population	629	1,005,224	490,283	1,181,751	163,787	5,281,728
GDP	629	51,272.42	31,883.21	50,470.24	3,226.249	246,981
Panel D: explanatory variables treated group (NUTS3 dataset)						
Population	1,054	145,312.90	129,522.50	68,117.57	38,586.00	350,473.00
GDP	1,054	5,096.84	3,825.94	4,443.41	1,103.90	40,725
Urban areas	1,054	0.35	0.00	0.48	0.00	1.00

Notes: The table presents summary statistics for the dependent variables and covariates in the treated FUA-year and NUTS3-year datasets over the period 2003–2019. GDP is expressed in €1,000.

Table C.2: Summary statistics – control regions

	N	Mean	Median	SD	Min.	Max.
Panel A: dependent variables control group (FUA dataset)						
Patents	1,003	524.75	246.74	1,275.91	3.50	10,676
NoChem. Pat.	1,003	446.60	206.32	1,192.14	3.50	10,259
First-time	1,003	65.93	35.90	133.51	0.00	1,287
Pats. by ind.	1,003	60.63	23.75	153.59	0.00	1,933
Pats. by orgs.	1,003	454.67	206.73	1,139.82	3.50	9,366
Forward cit.	1,003	8.37	8.34	4.26	0.52	25
Areas cited	1,003	5.63	5.33	1.56	1.88	12
AM patents	1,003	4.22	1.00	11.40	0.00	178
Utility models	984	49.74	23.00	96.62	1.00	857
Panel B: dependent variables control group (NUTS3 dataset)						
Patents	5,746	267.74	127.60	448.63	2.25	5,892
NoChem. Pat.	5,746	223.15	109.93	392.00	0.50	5,458
First-time	5,746	30.10	11.07	60.68	0.00	863
Patents by ind.	5,746	31.97	12.81	66.06	0.00	1,145
Patents by orgs.	5,746	230.80	108.60	392.56	1.08	5,523
Avg. forward cit.	5,746	8.61	8.24	5.21	0.23	121
Avg. areas cited	5,746	5.34	5.07	1.68	0.95	32
AM patents	5,746	2.78	0.50	8.13	0.00	168
Utility models	5,553	36.43	10.00	168.90	1.00	3,317
Panel C: explanatory variables control group (FUA dataset)						
Population	1,003	303,425.00	201,396.00	363,181.30	88,793	2,537,916
GDP	1,003	18,217.19	12,129.71	21,188.08	2,651.25	175,213
Panel D: explanatory variables control group (NUTS3 dataset)						
Population	5,746	215,955.00	157,648.00	250,424.00	33,944.00	3,669,491
GDP	5,746	7,183.44	4,284.41	11,491.40	833.83	157,131
Urban areas	5,746	0.45	0.00	0.50	0.00	1.00

Notes: The table presents summary statistics for the dependent variables and covariates in the control FUA-year and NUTS3-year datasets over the period 2003–2019. GDP is expressed in €1,000.

C.2.2 Tables – Heterogeneous Effects

Tables – University-Integrated FabLabs

Table C.3: Impact of university-integrated FabLabs (FUA-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	0.00704	-0.00306	-0.223*	0.187	-0.0173	0.0839	-0.0254	-0.0664	-0.0253
	(0.0386)	(0.0386)	(0.0888)	(0.182)	(0.0583)	(0.130)	(0.0866)	(0.0471)	(0.0287)
G2009	0.0712**	0.0173	0.0246	-0.393***	0.143***	-0.00848	0.0207	-0.0428	0.0171
	(0.0233)	(0.0220)	(0.0342)	(0.0415)	(0.0220)	(0.124)	(0.0453)	(0.0279)	(0.0147)
G2012	-0.0361	-0.00925	-0.242**	0.674*	-0.133*	-0.0286	-0.174	0.00982	-0.0371
	(0.0484)	(0.0465)	(0.0809)	(0.267)	(0.0658)	(0.117)	(0.211)	(0.102)	(0.0337)
G2013	-0.0307	-0.0516*	-0.424*	0.264***	-0.0504	0.0132	0.130*	-0.108***	-0.0912*
	(0.0340)	(0.0258)	(0.170)	(0.0727)	(0.0466)	(0.320)	(0.0537)	(0.0200)	(0.0385)
G2014	0.0889	0.105	-0.259	-0.0341	0.0853	0.407	-0.0572	-0.0540	0.0782***
	(0.0937)	(0.0991)	(0.148)	(0.114)	(0.105)	(0.290)	(0.165)	(0.104)	(0.0178)
G2018	-0.228***	-0.373***	0.175**	0.265**	-0.356***	0.0484	0.00845	-0.587***	-0.325***
	(0.0292)	(0.0349)	(0.0558)	(0.0807)	(0.0344)	(0.0590)	(0.0699)	(0.0355)	(0.0249)
N	1,139	1,139	1,139	1,139	1,139	1,139	1,139	1,139	1,139
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
$Controls$	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables, focusing on university-integrated FabLabs. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.4: Impact of university-integrated FabLabs (NUTS3-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	-0.0789	-0.0850	-0.137	-0.119	-0.0632	0.175	-0.160*	-0.0412	0.0312
	(0.0642)	(0.0657)	(0.114)	(0.159)	(0.0710)	(0.0978)	(0.0810)	(0.0607)	(0.0536)
G2009	-0.460***	-0.478***	0.203***	-1.045***	-0.462***	0.412***	-0.282***	-0.318***	-0.254***
	(0.0235)	(0.0234)	(0.0305)	(0.0483)	(0.0233)	(0.0368)	(0.0660)	(0.0191)	(0.0162)
G2011	-0.0714	0.0422	-0.128*	-0.191***	-0.0407	0.835***	-0.155***	-0.00958	0.0722***
	(0.0379)	(0.0399)	(0.0536)	(0.0537)	(0.0352)	(0.0549)	(0.0466)	(0.0200)	(0.0210)
G2012	-0.0855	-0.146	-0.200***	-0.163**	-0.0640	0.150	-0.368***	0.0120	-0.0759
	(0.135)	(0.106)	(0.0506)	(0.0541)	(0.130)	(0.132)	(0.0998)	(0.0722)	(0.101)
G2013	0.0423	0.0221	-0.139	-0.203**	0.147	0.223***	-0.444**	0.0925	0.121
	(0.0314)	(0.0307)	(0.536)	(0.0732)	(0.0859)	(0.0639)	(0.170)	(0.275)	(0.0798)
G2014	-0.0114	-0.00104	-0.305	0.247	-0.0285	-0.107	0.131	-0.0333	0.147
	(0.0918)	(0.0948)	(0.199)	(0.314)	(0.114)	(0.0950)	(0.121)	(0.0333)	(0.0863)
G2015	-0.0864***	-0.0904***	-0.00673	-0.416***	-0.0794***	-0.392***	-0.119*	0.0722**	0.0165
	(0.0207)	(0.0191)	(0.0428)	(0.0511)	(0.0185)	(0.0464)	(0.0513)	(0.0238)	(0.0148)
G2016	0.252***	0.226***	0.338***	0.805***	0.223***	0.0716	-0.218***	0.0566***	0.183***
	(0.0130)	(0.0135)	(0.0293)	(0.0395)	(0.0142)	(0.0374)	(0.0325)	(0.0163)	(0.0132)
G2018	-0.149	-0.188	-0.266	0.171	-0.177	0.368	-0.0153	-0.172	0.0153
	(0.201)	(0.216)	(0.168)	(0.171)	(0.215)	(0.244)	(0.362)	(0.0881)	(0.0900)
N	6,018	6,018	6,018	6,018	6,018	6,018	6,018	6,018	6,018
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables, focusing on university-integrated FabLabs. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Tables – Below-Median Subsample

Table C.5: Impact of FabLabs – below-median subsample (FUA-year)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Total	NoChem	First-time	Ind. Pat.	Org. Pat.	AM Pat.	UM	Fw. Cit.	Areas Cit.
ATT	-0.0207	0.0272	-0.0863	-0.0560	-0.0296	0.158	0.00126	-0.0218	-0.00594
	(0.0485)	(0.0485)	(0.105)	(0.101)	(0.0577)	(0.133)	(0.191)	(0.0329)	(0.303)
G2011	0	0.0550	0.283***	-0.150	0	0.456***	0	0.0730	-0.110***
	(.)	(0.0405)	(0.0569)	(0.0902)	(.)	(0.119)	(.)	(0.0388)	(0.0225)
G2012	-0.0751**	-0.0266	-0.253***	0.251***	-0.148***	-0.0826	-0.628***	0.0380	0.0145
	(0.0267)	(0.0324)	(0.0560)	(0.0752)	(0.0252)	(0.0690)	(0.0676)	(0.0855)	(0.0553)
G2013	0.00395	0.0123	-0.181	0.198*	0.0179	-0.0580	-0.0581	-0.137	-0.0698
	(0.0663)	(0.0745)	(0.215)	(0.0958)	(0.0882)	(0.250)	(0.167)	(0.0726)	(0.0588)
G2014	0.332***	0.524***	-0.244	-0.316	0.454***	0.376	0.503	-0.0261	0.136
	(0.0125)	(0.0129)	(0.135)	(0.197)	(0.0132)	(0.278)	(0.414)	(0.0788)	(0.105)
G2016	-0.105	0.0284	-0.0159	-0.416*	-0.0878	0.223	0.667***	-0.0307	-0.0178
	(0.0658)	(0.0852)	(0.108)	(0.181)	(0.0771)	(0.271)	(0.135)	(0.0848)	(0.0458)
G2017	-0.0960***	-0.0927*	0.0437	0.175	-0.101***	0.384	0	0.0237	0.0163
	(0.0292)	(0.0445)	(0.194)	(0.130)	(0.0225)	(0.199)	(.)	(0.0316)	(0.0150)
G2018	-0.214***	-0.369***	0.185**	0.282**	-0.343***	0.0627	0.0117	-0.245	-0.112
	(0.0323)	(0.0409)	(0.0591)	(0.0884)	(0.0389)	(0.172)	(0.0754)	(0.253)	(0.165)
N	862	879	947	1,100	862	1,190	912	1,328	1,173
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables in the below-median subsample. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.6: Impact of FabLabs – below-median subsample (NUTS3-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	-0.0191	-0.0133	-0.197**	0.0109	-0.0309	0.0737	-0.110	0.0105	0.0311
	(0.0445)	(0.0519)	(0.0746)	(0.111)	(0.0568)	(0.0848)	(0.0841)	(0.0356)	(0.0254)
G2011	-0.000811	0.0316	-0.174	0.189	-0.0596	0.177	-0.320**	-0.0850	-0.0211
	(0.0826)	(0.0966)	(0.167)	(0.358)	(0.134)	(0.194)	(0.110)	(0.111)	(0.0385)
G2012	-0.0856	-0.0159	-0.203***	-0.199***	-0.0645	0.119	-0.153*	0.00958	-0.0832
	(0.125)	(0.0324)	(0.0529)	(0.0484)	(0.112)	(0.132)	(0.0607)	(0.0659)	(0.103)
G2013	-0.222*	-0.278*	-0.192	-0.0711	-0.216	-0.164	0.0763	0.107	0.162*
	(0.0978)	(0.120)	(0.205)	(0.283)	(0.204)	(0.231)	(0.260)	(0.131)	(0.0630)
G2014	0.0362	-0.0344	-0.405**	0.151	0.00896	0.0133	0.0905	0.0452	0.103*
	(0.0758)	(0.0907)	(0.143)	(0.162)	(0.0899)	(0.142)	(0.174)	(0.0416)	(0.0512)
G2015	0	0	-0.650***	-0.00726	0	0.350	0.178	-0.00399	0.0609**
	(.)	(.)	(0.0477)	(0.304)	(.)	(0.269)	(0.333)	(0.0442)	(0.0227)
G2016	0.186*	0.225*	-0.186	0.122	0.187*	-0.128	-0.0409	0.111*	-0.0817
	(0.0859)	(0.106)	(0.195)	(0.181)	(0.0948)	(0.265)	(0.151)	(0.0478)	(0.0500)
G2017	0.0909**	0.0964**	0.0468	0.671***	0.0553	0.319	-1.253***	-0.0347	0.0367
	(0.0324)	(0.0302)	(0.146)	(0.166)	(0.0319)	(0.293)	(0.111)	(0.0732)	(0.0597)
G2018	-0.277	-0.285	-0.317*	-0.0734	-0.280	0.333	-0.0643	-0.0359	-0.0183
	(0.166)	(0.182)	(0.145)	(0.169)	(0.181)	(0.183)	(0.338)	(0.0749)	(0.0718)
N	6,800	6,800	6,800	6,800	6,800	6,800	6,800	6,800	6,800
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables in the below-median subsample. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Tables – Above-Median Subsample

Table C.7: Impact of FabLabs – above-median subsample (FUA-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	0.00447	-0.00882	-0.000720	-0.0962	0.0309	-0.0501	-0.0270	-0.00982	-0.00343
	(0.0230)	(0.0201)	(0.0351)	(0.0716)	(0.0285)	(0.133)	(0.0476)	(0.0673)	(0.0209)
G2009	0.112***	0.0518	0.0556*	-0.427***	0.185***	-0.412***	-0.0401	-0.0181	0.0434*
	(0.0254)	(0.0306)	(0.0225)	(0.0302)	(0.0277)	(0.119)	(0.0373)	(0.0359)	(0.0186)
G2010	-0.0321	-0.0187	-0.0403	-0.0618	-0.00273	0.0463	0.182	-0.00519	-0.0522***
	(0.0631)	(0.0155)	(0.101)	(0.0579)	(0.0699)	(0.166)	(0.129)	(0.00641)	(0.00205)
G2011	0.00374	0.00188	0.0116	0.154**	0.0151	0.0338	-0.0542	0	0.0165
	(0.0401)	(0.0329)	(0.0547)	(0.0541)	(0.0425)	(0.253)	(0.0682)	(.)	(0.0338)
G2012	-0.00102	0.00437	-0.0528	-0.352	0.164*	0.613	-0.140	-0.191***	0
	(0.0275)	(0.0269)	(0.0646)	(0.286)	(0.0750)	(0.320)	(0.101)	(0.0310)	(.)
G2013	0.00719	-0.0397	-0.0477	0.0407	0.0122	-0.141	0.146***	0.132***	-0.0377
	(0.0249)	(0.0307)	(0.0873)	(0.0726)	(0.0241)	(0.408)	(0.0405)	(0.0238)	(0.0259)
G2014	-0.0400	-0.0787*	0.0188	-0.367*	-0.0216	-0.135	-0.144*	0.122	-0.00345
	(0.0321)	(0.0321)	(0.112)	(0.183)	(0.0431)	(0.152)	(0.0586)	(0.229)	(0.0721)
G2015	0.0366	0.0212	0.259***	-0.381**	0.0563	-0.197**	-0.191	0	0
	(0.106)	(0.109)	(0.0585)	(0.139)	(0.111)	(0.0676)	(0.116)	(.)	(.)
G2016	0.0599	0.118*	-0.0813	-0.138	0.0951	-0.113	0.0959	-0.0123	0
	(0.0489)	(0.0465)	(0.0934)	(0.116)	(0.0579)	(0.344)	(0.0648)	(0.0812)	(.)
G2017	-0.194**	-0.0934***	-0.250***	-0.120***	-0.163***	-0.521***	-0.369***	-0.169***	-0.207***
	(0.0623)	(0.0144)	(0.0175)	(0.0310)	(0.0119)	(0.0313)	(0.0326)	(0.0226)	(0.00708)
G2018	-0.170***	-0.249***	0.0425	0.0770	-0.205***	0	-0.0134	0	0
	(0.0156)	(0.0204)	(0.0434)	(0.0942)	(0.0173)	(.)	(0.0425)	(.)	(.)
N	765	748	680	527	765	442	714	276	408
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables in the above-median subsample. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.8: Impact of FabLabs – above-median subsample (NUTS3-year)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Total	NoChem	First-time	Ind. Pat.	Org. Pat.	AM Pat.	UM	Fw. Cit.	Areas Cit.
ATT	0.00230	-0.0204	0.0874	-0.0527	0.00738	0.382*	-0.0975	0.0354	0.0187
	(0.0376)	(0.0355)	(0.0517)	(0.102)	(0.0389)	(0.172)	(0.0817)	(0.0548)	(0.0447)
G2009	-0.355***	-0.365***	0	-0.751***	-0.369***	0	0	0	-0.210***
	(0.0216)	(0.0307)	(.)	(0.0337)	(0.0258)	(.)	(.)	(.)	(0.0209)
G2010	0.0863**	0.0654	0.121*	-0.0730	0.158***	0.901**	0.00358	-0.00532	0.0859***
	(0.0332)	(0.0343)	(0.0476)	(0.0958)	(0.0343)	(0.311)	(0.0986)	(0.0361)	(0.0203)
G2011	-0.107	-0.0425	-0.0480	0.171**	-0.0704	0.345**	0.190	0.0637	0.177***
	(0.0888)	(0.101)	(0.0686)	(0.0659)	(0.0870)	(0.108)	(0.138)	(0.0586)	(0.0237)
G2012	-0.0704	-0.157	0.120***	0.0915	-0.0686	-0.0612	-0.125	0.0128	0.0622***
	(0.0835)	(0.112)	(0.0323)	(0.0864)	(0.0716)	(0.222)	(0.254)	(0.0283)	(0.0118)
G2013	0.0779*	-0.0470	0.422***	0.0671	0.0163	0	-0.00809	0.0874	0.0272
	(0.0391)	(0.0454)	(0.0228)	(0.172)	(0.0662)	(.)	(0.0987)	(0.161)	(0.0228)
G2014	0.0198	0.0304	0.0813	-0.263	0.0252	-0.0433	-0.267	-0.0172	0.0503
	(0.0505)	(0.0331)	(0.0723)	(0.251)	(0.0510)	(0.228)	(0.240)	(0.0547)	(0.0522)
G2015	0.0243	0.0121	0.230	0.0674	0.0279	0.340***	-0.0650	0.274***	0.0238
	(0.0579)	(0.0536)	(0.151)	(0.218)	(0.0530)	(0.0789)	(0.146)	(0.0222)	(0.0230)
G2016	0.0471	0.0382	-0.107	0.206	0.0452	0.364	-0.142	-0.0634	-0.111**
	(0.0579)	(0.0551)	(0.152)	(0.171)	(0.0622)	(0.197)	(0.117)	(0.170)	(0.0356)
G2018	0.100*	0.116***	-0.0352	0.708*	0.0952**	0.234***	-0.124	-0.200***	0.0381*
	(0.0401)	(0.0290)	(0.0758)	(0.352)	(0.0326)	(0.0684)	(0.0759)	(0.0298)	(0.0153)
N	3,026	2,992	2,652	1,887	2,992	1,411	2,397	1,173	1,377
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
$Controls$	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables in the above-median subsample. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Tables – Urban vs. Rural Effects

Table C.9: Impact of FabLabs – urban subsample (NUTS3-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	0.0162	0.000860	0.0402	-0.0460	0.0386	0.312*	-0.0315	0.113**	0.0408
	(0.0331)	(0.0288)	(0.0606)	(0.0965)	(0.0333)	(0.138)	(0.0683)	(0.0413)	(0.0233)
G2010	0.0689*	0.0445	0.141**	-0.168	0.143***	0.898**	-0.0479	0.0959	-0.00519
	(0.0310)	(0.0349)	(0.0545)	(0.0967)	(0.0318)	(0.311)	(0.104)	(0.0876)	(0.0272)
G2011	-0.138**	-0.141***	-0.123*	0.0254	-0.109**	0.512***	-0.173	0.200	0.0106
	(0.0492)	(0.0400)	(0.0531)	(0.119)	(0.0408)	(0.0621)	(0.200)	(0.153)	(0.0399)
G2012	-0.130**	-0.0887	0.0166	-0.206	-0.0750	0.0398	0.390***	-0.00608	-0.000662
	(0.0494)	(0.0520)	(0.0474)	(0.137)	(0.0397)	(0.0872)	(0.0749)	(0.0178)	(0.0103)
G2013	0.168***	-0.0580**	0.401***	-0.340***	0.206***	-1.193***	0.0519	0.397***	0.279***
	(0.0194)	(0.0193)	(0.0287)	(0.0443)	(0.0183)	(0.0453)	(0.0409)	(0.0177)	(0.0126)
G2014	0.00420	-0.00260	-0.00706	-0.225	0.0174	0.321*	0.0690	0.0559	0.0558
	(0.0435)	(0.0356)	(0.0800)	(0.223)	(0.0467)	(0.128)	(0.144)	(0.0418)	(0.0364)
G2015	0.0961	0.101	0.522**	0.355***	0.0704	0.734***	-0.262***	0.110	0.114*
	(0.112)	(0.0807)	(0.176)	(0.0707)	(0.104)	(0.190)	(0.0536)	(0.114)	(0.0563)
G2016	0.0913	0.0954	-0.200	0.162	0.0839	-0.00108	-0.0485	0.128	0.00356
	(0.0591)	(0.0506)	(0.165)	(0.263)	(0.0634)	(0.185)	(0.0657)	(0.107)	(0.0540)
G2017	0.0162	0.00841	0.0485	0.927***	-0.00250	0.214**	-0.296***	-0.0647**	-0.0211
	(0.0180)	(0.0197)	(0.0366)	(0.0682)	(0.0187)	(0.0705)	(0.0420)	(0.0198)	(0.0156)
G2018	0.0504	0.0697*	0.217***	-0.294***	0.0784**	0.405***	-0.789***	0.209***	-0.0863***
	(0.0263)	(0.0289)	(0.0484)	(0.0630)	(0.0256)	(0.0455)	(0.0498)	(0.0215)	(0.0187)
N	2,958	2,958	2,958	2,958	2,958	2,958	2,958	2,958	2,958
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
$Controls$	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables in the urban subsample. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.10: Impact of FabLabs – rural subsample (NUTS3-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	-0.0348	-0.0370	-0.169*	0.0543	-0.0540	0.0422	-0.135	-0.0448	0.0150
	(0.0428)	(0.0463)	(0.0754)	(0.120)	(0.0492)	(0.0956)	(0.0857)	(0.0372)	(0.0328)
G2009	-0.452***	-0.474***	0.217***	-1.027***	-0.461***	0.385***	-0.268***	-0.329***	-0.255***
	(0.0259)	(0.0256)	(0.0347)	(0.0520)	(0.0266)	(0.0391)	(0.0682)	(0.0214)	(0.0181)
G2011	0.0242	0.0957	-0.168	0.226	-0.0307	0.162	-0.261	-0.191**	0.0139
	(0.0778)	(0.0839)	(0.213)	(0.445)	(0.127)	(0.272)	(0.145)	(0.0636)	(0.0456)
G2012	-0.0869	-0.146	-0.206***	-0.145***	-0.0814	0.0986	-0.380**	0.00609	-0.0895
	(0.140)	(0.116)	(0.0460)	(0.0432)	(0.131)	(0.0860)	(0.141)	(0.0682)	(0.103)
G2013	-0.146	-0.171*	-0.175	-0.0279	-0.145	-0.0514	0.0124	0.0453	0.114
	(0.0782)	(0.0791)	(0.202)	(0.199)	(0.109)	(0.189)	(0.189)	(0.104)	(0.0646)
G2014	0.0425	-0.00501	-0.431**	0.254	-0.00159	-0.340*	0.00537	0.000552	0.142
	(0.0891)	(0.0912)	(0.156)	(0.180)	(0.106)	(0.145)	(0.261)	(0.0667)	(0.0783)
G2015	-0.0156	-0.0255	-0.211	-0.205	-0.0120	0.129	0.262	0.0199	0.0212
	(0.0405)	(0.0475)	(0.191)	(0.228)	(0.0341)	(0.276)	(0.294)	(0.0336)	(0.0116)
G2016	0.134	0.162	-0.0987	0.116	0.139	0.00518	-0.0844	0.0459	-0.152**
	(0.0944)	(0.110)	(0.189)	(0.141)	(0.101)	(0.365)	(0.176)	(0.0436)	(0.0572)
G2017	0.106**	0.116***	0.0691	0.697***	0.0570	0.317	-0.956***	-0.0154	0.0652
	(0.0324)	(0.0275)	(0.175)	(0.165)	(0.0383)	(0.365)	(0.285)	(0.0908)	(0.0750)
G2018	-0.0828	-0.0773	-0.220*	0.213	-0.0907	0.281	-0.00509	-0.108	0.0111
	(0.128)	(0.134)	(0.0975)	(0.213)	(0.133)	(0.186)	(0.189)	(0.0655)	(0.0480)
N	3,842	3,842	3,842	3,842	3,842	3,842	3,842	3,842	3,842
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
$Controls$	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables in the rural subsample. The unit of analysis is the NUTS3-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Tables – Regression Results for Pre-Existing Community Workshops (FUA-level)

Table C.11: Total patents & non-chemical patents (FUA-year)

	Total Patents				Non-chemical Patents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HS or MS	MS & HS	MS	FL	HS or MS	MS & HS	MS	FL
τ_0 Base	-0.106***	-0.00835	-0.0249	0.00310	-0.169***	-0.0260	-0.0456	-0.00194
	(0.0246)	(0.0340)	(0.0285)	(0.0253)	(0.0227)	(0.0452)	(0.0407)	(0.0265)
τ_1 Base	-0.0333	0.00428	-0.00979	0.0106	-0.0698*	-0.00689	-0.0371	-0.00273
	(0.0255)	(0.0295)	(0.0271)	(0.0251)	(0.0277)	(0.0335)	(0.0308)	(0.0265)
τ_2 Base	0.0307	0.0622	0.00835	0.0332	0.0480	0.0488	0.00370	0.0125
	(0.0235)	(0.0359)	(0.0292)	(0.0311)	(0.0253)	(0.0394)	(0.0307)	(0.0329)
τ_3 Base	0.0582*	0.0575	-0.00431	0.0295	0.0963**	0.0591	0.00687	0.0261
	(0.0265)	(0.0429)	(0.0305)	(0.0368)	(0.0313)	(0.0485)	(0.0330)	(0.0405)
τ_4 Base	0.0955**	0.0593	0.00634	0.0107	0.0944**	0.0454	0.00785	0.00812
	(0.0343)	(0.0447)	(0.0313)	(0.0355)	(0.0309)	(0.0547)	(0.0319)	(0.0422)
τ_5 Base	0.303***	0.124	0.0547	0.0200	0.373***	0.102	0.0715	0.00173
	(0.0538)	(0.0711)	(0.0429)	(0.0450)	(0.0415)	(0.0824)	(0.0440)	(0.0490)
τ_0 ind.=1	0.00310	-0.0278	0.00670	-0.106***	-0.00194	-0.0355	0.00535	-0.169***
	(0.0253)	(0.0239)	(0.0464)	(0.0246)	(0.0265)	(0.0277)	(0.0482)	(0.0227)
τ_1 ind.=1	0.0106	0.00124	0.0329	-0.0333	-0.00273	-0.0279	0.0388	-0.0698*
	(0.0251)	(0.0281)	(0.0383)	(0.0255)	(0.0265)	(0.0313)	(0.0416)	(0.0277)
τ_2 ind.=1	0.0332	-0.0289	0.0837	0.0307	0.0125	-0.0442	0.0502	0.0480
	(0.0311)	(0.0291)	(0.0496)	(0.0235)	(0.0329)	(0.0310)	(0.0590)	(0.0253)
τ_3 ind.=1	0.0295	-0.0150	0.115	0.0582*	0.0261	-0.00791	0.102	0.0963**
	(0.0368)	(0.0330)	(0.0722)	(0.0265)	(0.0405)	(0.0369)	(0.0844)	(0.0313)
τ_4 ind.=1	0.0107	-0.0235	0.0802	0.0955**	0.00812	-0.00955	0.0665	0.0944**
	(0.0355)	(0.0278)	(0.0825)	(0.0343)	(0.0422)	(0.0310)	(0.111)	(0.0309)
τ_5 ind.=1	0.0200	-0.00946	0.0783	0.303***	0.00173	0.00261	0.00133	0.373***
	(0.0450)	(0.0282)	(0.127)	(0.0538)	(0.0490)	(0.0303)	(0.147)	(0.0415)
Pre1	0.0436	0.0436	0.0436	0.0436	0.0425	0.0425	0.0425	0.0425
	(0.0365)	(0.0365)	(0.0365)	(0.0365)	(0.0366)	(0.0366)	(0.0366)	(0.0366)
Pre2	0.0343	0.0343	0.0343	0.0343	0.0385	0.0385	0.0385	0.0385
	(0.0337)	(0.0337)	(0.0337)	(0.0337)	(0.0347)	(0.0347)	(0.0347)	(0.0347)
Pre3	0.0206	0.0206	0.0206	0.0206	0.0199	0.0199	0.0199	0.0199
	(0.0338)	(0.0338)	(0.0338)	(0.0338)	(0.0345)	(0.0345)	(0.0345)	(0.0345)
Pre4	0.0414	0.0414	0.0414	0.0414	0.0504	0.0504	0.0504	0.0504
	(0.0336)	(0.0336)	(0.0336)	(0.0336)	(0.0336)	(0.0336)	(0.0336)	(0.0336)
N	1,593	1,593	1,593	1,593	1,593	1,593	1,593	1,593
FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed total number of patents and log-transformed number of non-chemical patents. The abbreviations "HS" refer to Hackerspace, "MS" to Makerspace, and "FL" to FabLab. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.12: First-time inventors & individual applicants (FUA-year)

	First-Time				Ind. Patents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HS or MS	MS & HS	MS	FL	HS or MS	MS & HS	MS	FL
τ_0 Base	-0.217*	-0.0492	-0.0364	0.0179	-0.236***	-0.111	-0.00250	-0.0196
	(0.110)	(0.0501)	(0.0465)	(0.0269)	(0.0304)	(0.0595)	(0.0409)	(0.0520)
τ_1 Base	-0.127*	-0.0228	-0.0354	-0.00601	-0.0476	-0.0570	0.0381	0.0266
	(0.0554)	(0.0472)	(0.0383)	(0.0322)	(0.108)	(0.0647)	(0.0497)	(0.0525)
τ_2 Base	-0.0403	0.0139	-0.0367	0.00880	0.114*	-0.0478	0.0182	-0.0922
	(0.0281)	(0.0421)	(0.0315)	(0.0347)	(0.0451)	(0.0848)	(0.0543)	(0.0632)
τ_3 Base	0.223***	0.0391	0.00789	-0.0147	0.118	-0.244**	-0.0761	-0.272***
	(0.0301)	(0.0438)	(0.0370)	(0.0323)	(0.110)	(0.0938)	(0.0761)	(0.0688)
τ_4 Base	0.109***	-0.00541	0.00453	-0.0124	-0.173*	-0.323***	-0.256***	-0.306***
	(0.0329)	(0.0409)	(0.0337)	(0.0367)	(0.0749)	(0.0792)	(0.0592)	(0.0713)
τ_5 Base	0.329***	0.000741	0.0658	-0.0186	0.0119	-0.333***	-0.233**	-0.305***
	(0.0375)	(0.0600)	(0.0416)	(0.0475)	(0.0690)	(0.0835)	(0.0719)	(0.0713)
τ_0 ind.=1	0.0179	0.0401*	0.0135	-0.217*	-0.0196	0.0628	-0.163	-0.236***
	(0.0269)	(0.0186)	(0.0409)	(0.110)	(0.0520)	(0.0382)	(0.0916)	(0.0304)
τ_1 ind.=1	-0.00601	-0.0330	-0.00541	-0.127*	0.0266	0.157***	-0.0399	-0.0476
	(0.0322)	(0.0237)	(0.0606)	(0.0554)	(0.0525)	(0.0352)	(0.0869)	(0.108)
τ_2 ind.=1	0.00880	-0.0286	0.0772	-0.0403	-0.0922	-0.0727	-0.211*	0.114*
	(0.0347)	(0.0294)	(0.0655)	(0.0281)	(0.0632)	(0.0550)	(0.0922)	(0.0451)
τ_3 ind.=1	-0.0147	-0.00885	0.0567	0.223***	-0.272***	-0.136	-0.489***	0.118
	(0.0323)	(0.0291)	(0.0530)	(0.0301)	(0.0688)	(0.0708)	(0.0946)	(0.110)
τ_4 ind.=1	-0.0124	0.0263	0.0176	0.109***	-0.306***	-0.229***	-0.368**	-0.173*
	(0.0367)	(0.0343)	(0.0592)	(0.0329)	(0.0713)	(0.0656)	(0.133)	(0.0749)
τ_5 ind.=1	-0.0186	0.0645	-0.0801	0.329***	-0.305***	-0.179*	-0.344**	0.0119
	(0.0475)	(0.0393)	(0.114)	(0.0375)	(0.0713)	(0.0721)	(0.116)	(0.0690)
Pre1	0.0163	0.0163	0.0163	0.0163	0.0319	0.0319	0.0319	0.0319
	(0.0449)	(0.0449)	(0.0449)	(0.0449)	(0.0606)	(0.0606)	(0.0606)	(0.0606)
Pre2	-0.0181	-0.0181	-0.0181	-0.0181	0.00975	0.00975	0.00975	0.00975
	(0.0434)	(0.0434)	(0.0434)	(0.0434)	(0.0602)	(0.0602)	(0.0602)	(0.0602)
Pre3	0.0107	0.0107	0.0107	0.0107	-0.0359	-0.0359	-0.0359	-0.0359
	(0.0408)	(0.0408)	(0.0408)	(0.0408)	(0.0563)	(0.0563)	(0.0563)	(0.0563)
Pre4	0.0209	0.0209	0.0209	0.0209	0.0162	0.0162	0.0162	0.0162
	(0.0365)	(0.0365)	(0.0365)	(0.0365)	(0.0660)	(0.0660)	(0.0660)	(0.0660)
N	1,593	1,593	1,593	1,593	1,593	1,593	1,593	1,593
FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed number of applications by first-time inventors and by individual applicants. The abbreviations "HS" refer to Hackerspace, "MS" to Makerspace, and "FL" to FabLab. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table C.13: Organizational applicants & AM-related patents (FUA-year)

	Org. Patents				AM patents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HS or MS	MS & HS	MS	FL	HS or MS	MS & HS	MS	FL
τ_0 Base	-0.0527*	0.00290	-0.0163	-0.0000383	-0.0762	-0.00736	0.0802	0.0569
	(0.0236)	(0.0351)	(0.0301)	(0.0269)	(0.0824)	(0.102)	(0.112)	(0.101)
τ_1 Base	-0.00763	0.0103	-0.0102	0.00584	-0.149	-0.0233	0.0186	0.0404
	(0.0445)	(0.0313)	(0.0285)	(0.0257)	(0.133)	(0.109)	(0.115)	(0.0886)
τ_2 Base	0.0488*	0.0704	0.0191	0.0417	-0.143**	-0.0425	-0.0413	0.0601
	(0.0240)	(0.0377)	(0.0307)	(0.0326)	(0.0535)	(0.113)	(0.0863)	(0.0900)
τ_3 Base	0.0870*	0.0727	0.0145	0.0427	0.00880	0.129	0.0689	0.144
	(0.0385)	(0.0470)	(0.0333)	(0.0388)	(0.0656)	(0.141)	(0.0791)	(0.106)
τ_4 Base	0.155***	0.0958*	0.0335	0.0334	0.831**	0.324	0.409***	0.212
	(0.0296)	(0.0488)	(0.0347)	(0.0373)	(0.254)	(0.176)	(0.108)	(0.116)
τ_5 Base	0.395***	0.175*	0.0886	0.0449	0.591	0.220	0.307**	0.172
	(0.0562)	(0.0813)	(0.0493)	(0.0490)	(0.315)	(0.210)	(0.118)	(0.127)
τ_0 ind.=1	-0.0000383	-0.0325	0.00752	-0.0527*	0.0569	0.124	-0.0581	-0.0762
	(0.0269)	(0.0271)	(0.0470)	(0.0236)	(0.101)	(0.0920)	(0.146)	(0.0824)
τ_1 ind.=1	0.00584	-0.00989	0.0350	-0.00763	0.0404	0.0729	-0.0136	-0.149
	(0.0257)	(0.0271)	(0.0389)	(0.0445)	(0.0886)	(0.0695)	(0.126)	(0.133)
τ_2 ind.=1	0.0417	-0.0145	0.0928	0.0488*	0.0601	0.164*	0.161	-0.143**
	(0.0326)	(0.0279)	(0.0540)	(0.0240)	(0.0900)	(0.0745)	(0.103)	(0.0535)
τ_3 ind.=1	0.0427	0.00177	0.124	0.0870*	0.144	0.109	0.235	0.00880
	(0.0388)	(0.0318)	(0.0771)	(0.0385)	(0.106)	(0.0989)	(0.196)	(0.0656)
τ_4 ind.=1	0.0334	-0.00523	0.114	0.155***	0.212	0.302***	0.0320	0.831**
	(0.0373)	(0.0283)	(0.0882)	(0.0296)	(0.116)	(0.0813)	(0.289)	(0.254)
τ_5 ind.=1	0.0449	0.00662	0.115	0.395***	0.172	0.246*	-0.00604	0.591
	(0.0490)	(0.0296)	(0.141)	(0.0562)	(0.125)	(0.105)	(0.334)	(0.325)
Pre1	0.0466	0.0466	0.0466	0.0466	0.0224	0.0224	0.0224	0.0224
	(0.0404)	(0.0404)	(0.0404)	(0.0404)	(0.128)	(0.128)	(0.128)	(0.128)
Pre2	0.0303	0.0303	0.0303	0.0303	0.0308	0.0308	0.0308	0.0308
	(0.0372)	(0.0372)	(0.0372)	(0.0372)	(0.127)	(0.127)	(0.127)	(0.127)
Pre3	0.0221	0.0221	0.0221	0.0221	0.103	0.103	0.103	0.103
	(0.0356)	(0.0356)	(0.0356)	(0.0356)	(0.116)	(0.116)	(0.116)	(0.116)
Pre4	0.0483	0.0483	0.0483	0.0483	0.0969	0.0969	0.0969	0.0969
	(0.0332)	(0.0332)	(0.0332)	(0.0332)	(0.108)	(0.108)	(0.108)	(0.108)
N	1,593	1,593	1,593	1,593	1,593	1,593	1,593	1,593
FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed number of applications by organizations and AM-related patents. The abbreviations "HS" refer to Hackerspace, "MS" to Makerspace, and "FL" to FabLab. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.14: Utility model applications (FUA-year)

	(1) HS or MS	(2) HS & MS	(3) MS	(4) FL
τ_1 Base	-0.0175	-0.0497	-0.000511	-0.0371
	(0.0754)	(0.0530)	(0.0420)	(0.0351)
τ_2 Base	0.126*	-0.0244	0.00581	-0.0657
	(0.0576)	(0.0587)	(0.0433)	(0.0437)
τ_3 Base	-0.112	-0.0999	-0.0619	-0.0772
	(0.0810)	(0.0623)	(0.0519)	(0.0492)
τ_4 Base	-0.256	-0.0469	-0.0424	-0.00546
	(0.184)	(0.0819)	(0.0741)	(0.0543)
τ_5 Base	-0.242	-0.111	-0.0128	-0.0144
	(0.193)	(0.0953)	(0.0717)	(0.0632)
τ_0 ind.=1	-0.223*	-0.127	-0.0469	-0.0652
	(0.0918)	(0.142)	(0.0627)	(0.0955)
τ_1 ind.=1	-0.0371	-0.00103	-0.103	-0.0175
	(0.0351)	(0.0272)	(0.0559)	(0.0754)
τ_2 ind.=1	-0.0657	-0.0528	-0.124	0.126*
	(0.0437)	(0.0396)	(0.0792)	(0.0576)
τ_3 ind.=1	-0.0772	-0.0483	-0.128	-0.112
	(0.0492)	(0.0456)	(0.0655)	(0.0810)
τ_4 ind.=1	-0.00546	-0.0435	-0.0530	-0.256
	(0.0543)	(0.0602)	(0.0928)	(0.184)
τ_5 ind.=1	-0.0144	0.0295	-0.171	-0.242
	(0.0632)	(0.0657)	(0.112)	(0.193)
Pre1	-0.0923	-0.0923	-0.0923	-0.0923
	(0.0747)	(0.0747)	(0.0747)	(0.0747)
Pre2	-0.0297	-0.0297	-0.0297	-0.0297
	(0.0594)	(0.0594)	(0.0594)	(0.0594)
Pre3	0.00246	0.00246	0.00246	0.00246
	(0.0431)	(0.0431)	(0.0431)	(0.0431)
Pre4	-0.0283	-0.0283	-0.0283	-0.0283
	(0.0378)	(0.0378)	(0.0378)	(0.0378)
N	1,593	1,593	1,593	1,593
FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

*Notes:*This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed number of utility model applications. The abbreviations “HS” refer to Hackerspace, “MS” to Makerspace, and “FL” to FabLab. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.15: Patent quality measures (FUA-year)

	Fw. Citations				Areas Cited			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HS or MS	MS & HS	MS	FL	HS or MS	MS & HS	MS	FL
τ_0 Base	-0.00658	0.0204	0.00737	0.0222	-0.0237	0.00447	0.00410	0.00635
	(0.0272)	(0.0332)	(0.0264)	(0.0244)	(0.0349)	(0.0248)	(0.0239)	(0.0159)
τ_1 Base	0.0717*	0.00777	0.0372	0.00226	-0.0584**	-0.00897	0.00138	0.0104
	(0.0339)	(0.0315)	(0.0255)	(0.0247)	(0.0200)	(0.0204)	(0.0175)	(0.0175)
τ_2 Base	0.00492	0.0906	0.0261	0.0803	-0.00424	0.0337	0.0236	0.0360
	(0.0482)	(0.0551)	(0.0306)	(0.0441)	(0.0451)	(0.0304)	(0.0213)	(0.0252)
τ_3 Base	-0.0357	0.0720	0.0163	0.0753	-0.00563	0.0257	0.00242	0.0220
	(0.0661)	(0.0497)	(0.0344)	(0.0415)	(0.0382)	(0.0227)	(0.0193)	(0.0202)
τ_4 Base	0.0920*	0.0944	0.0534	0.0610	-0.00505	0.0338	0.00243	0.0158
	(0.0364)	(0.0560)	(0.0281)	(0.0410)	(0.0175)	(0.0230)	(0.0186)	(0.0209)
τ_5 Base	0.0330	0.0169	0.0184	0.0135	0.110	0.0700	0.0230	0.0173
	(0.0244)	(0.0779)	(0.0260)	(0.0531)	(0.118)	(0.0446)	(0.0321)	(0.0264)
τ_0 ind.=1	0.0222	0.0115	0.0386	-0.00658	0.00635	-0.00479	-0.00401	-0.0237
	(0.0244)	(0.0230)	(0.0453)	(0.0272)	(0.0159)	(0.0167)	(0.0193)	(0.0349)
τ_1 ind.=1	0.00226	0.0260	-0.0393	0.0717*	0.0104	0.0147	-0.00663	-0.0584**
	(0.0247)	(0.0258)	(0.0446)	(0.0339)	(0.0175)	(0.0173)	(0.0303)	(0.0200)
τ_2 ind.=1	0.0803	0.0176	0.152	0.00492	0.0360	0.0188	0.0398	-0.00424
	(0.0441)	(0.0247)	(0.0982)	(0.0482)	(0.0252)	(0.0182)	(0.0482)	(0.0451)
τ_3 ind.=1	0.0753	0.0267	0.144	-0.0357	0.0220	0.000461	0.0493	-0.00563
	(0.0415)	(0.0219)	(0.0805)	(0.0661)	(0.0202)	(0.0229)	(0.0318)	(0.0382)
τ_4 ind.=1	0.0610	0.0265	0.104	0.0920*	0.0158	-0.0177	0.0420	-0.00505
	(0.0410)	(0.0260)	(0.118)	(0.0364)	(0.0209)	(0.0225)	(0.0390)	(0.0175)
τ_5 ind.=1	0.0135	0.0156	0.00949	0.0330	0.0173	-0.0128	0.0550	0.110
	(0.0531)	(0.0250)	(0.167)	(0.0244)	(0.0264)	(0.0219)	(0.0590)	(0.118)
Pre1	0.0670*	0.0670*	0.0670*	0.0670*	0.0248	0.0248	0.0248	0.0248
	(0.0299)	(0.0299)	(0.0299)	(0.0299)	(0.0181)	(0.0181)	(0.0181)	(0.0181)
Pre2	0.0288	0.0288	0.0288	0.0288	-0.00182	-0.00182	-0.00182	-0.00182
	(0.0232)	(0.0232)	(0.0232)	(0.0232)	(0.0137)	(0.0137)	(0.0137)	(0.0137)
Pre3	0.0469	0.0469	0.0469	0.0469	-0.00172	-0.00172	-0.00172	-0.00172
	(0.0240)	(0.0240)	(0.0240)	(0.0240)	(0.0194)	(0.0194)	(0.0194)	(0.0194)
Pre4	0.0380	0.0380	0.0380	0.0380	0.00884	0.00884	0.00884	0.00884
	(0.0323)	(0.0323)	(0.0323)	(0.0323)	(0.0181)	(0.0181)	(0.0181)	(0.0181)
N	1,593	1,593	1,593	1,593	1,593	1,593	1,593	1,593
FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed number of forward citations per patent and of average areas cited per patent. The abbreviations “HS” refer to Hackerspace, “MS” to Makerspace, and “FL” to FabLab. The unit of analysis is the FUA-year, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Tables – Regression Results for Pre-Existing Community Workshops (Urban NUTS3-level)

Table C.16: Total patents & w/o chemical patents (urban, NUTS3-year)

	Total Patents				Non-chemical Patents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HS or MS	MS & HS	MS	FL	HS or MS	MS & HS	MS	FL
τ_0 Base	-0.133***	-0.146***	-0.122***	-0.138***	-0.123***	-0.135***	-0.121***	-0.137***
	(0.0256)	(0.0236)	(0.0311)	(0.0221)	(0.0297)	(0.0249)	(0.0322)	(0.0219)
τ_1 Base	-0.0958***	-0.117***	-0.0666**	-0.0840**	-0.0870***	-0.103***	-0.0647**	-0.0772**
	(0.0266)	(0.0309)	(0.0246)	(0.0266)	(0.0225)	(0.0293)	(0.0229)	(0.0237)
τ_2 Base	0.00809	-0.0204	0.0216	-0.00970	0.0133	-0.0137	0.0189	-0.0134
	(0.0356)	(0.0384)	(0.0271)	(0.0370)	(0.0304)	(0.0313)	(0.0266)	(0.0318)
τ_3 Base	-0.0302	-0.0693*	0.00580	-0.0345	-0.0584*	-0.0878**	-0.0123	-0.0351
	(0.0353)	(0.0323)	(0.0324)	(0.0322)	(0.0257)	(0.0288)	(0.0312)	(0.0308)
τ_4 Base	-0.0362	-0.126*	-0.0165	-0.124*	-0.0619	-0.155*	-0.0279	-0.135**
	(0.0388)	(0.0622)	(0.0359)	(0.0549)	(0.0329)	(0.0607)	(0.0389)	(0.0487)
τ_5 Base	0.0471	-0.0256	0.0165	-0.0752	0.0495	-0.0274	0.0259	-0.0687
	(0.0427)	(0.0905)	(0.0568)	(0.100)	(0.0396)	(0.0782)	(0.0594)	(0.0810)
τ_0 ind.=1	-0.138***	-0.111***	-0.175***	-0.133***	-0.137***	-0.121***	-0.157***	-0.123***
	(0.0221)	(0.0246)	(0.0312)	(0.0256)	(0.0219)	(0.0300)	(0.0304)	(0.0297)
τ_1 ind.=1	-0.0840**	-0.0201	-0.161**	-0.0958***	-0.0772**	-0.0276	-0.135**	-0.0870***
	(0.0266)	(0.0180)	(0.0549)	(0.0266)	(0.0237)	(0.0175)	(0.0462)	(0.0225)
τ_2 ind.=1	-0.00970	0.0419*	-0.0716	0.00809	-0.0134	0.0272	-0.0622	0.0133
	(0.0370)	(0.0192)	(0.0729)	(0.0356)	(0.0318)	(0.0297)	(0.0535)	(0.0304)
τ_3 ind.=1	-0.0345	0.0706**	-0.140**	-0.0302	-0.0351	0.0706**	-0.141***	-0.0584*
	(0.0322)	(0.0274)	(0.0449)	(0.0353)	(0.0308)	(0.0372)	(0.0312)	(0.0257)
τ_4 ind.=1	-0.124*	0.0131	-0.261*	-0.0362	-0.135**	0.0232	-0.294***	-0.0619
	(0.0549)	(0.0251)	(0.105)	(0.0388)	(0.0487)	(0.0363)	(0.0877)	(0.0329)
τ_5 ind.=1	-0.0752	-0.0217	-0.147	0.0471	-0.0687	-0.00358	-0.156	0.0495
	(0.100)	(0.0413)	(0.226)	(0.0427)	(0.0810)	(0.0282)	(0.183)	(0.0396)
Pre1	-0.107**	-0.107**	-0.107**	-0.107**	-0.0980**	-0.0980**	-0.0980**	-0.0980**
	(0.0391)	(0.0391)	(0.0391)	(0.0391)	(0.0357)	(0.0357)	(0.0357)	(0.0357)
Pre2	-0.0459	-0.0459	-0.0459	-0.0459	-0.0631	-0.0631	-0.0631	-0.0631
	(0.0363)	(0.0363)	(0.0363)	(0.0363)	(0.0351)	(0.0351)	(0.0351)	(0.0351)
Pre3	-0.0153	-0.0153	-0.0153	-0.0153	-0.0115	-0.0115	-0.0115	-0.0115
	(0.0451)	(0.0451)	(0.0451)	(0.0451)	(0.0414)	(0.0414)	(0.0414)	(0.0414)
Pre4	-0.0312	-0.0312	-0.0312	-0.0312	-0.0318	-0.0318	-0.0318	-0.0318
	(0.0477)	(0.0477)	(0.0477)	(0.0477)	(0.0460)	(0.0460)	(0.0460)	(0.0460)
N	2,941	2,941	2,941	2,941	2,941	2,941	2,941	2,941
FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed total number of patents and of non-chemical patents. The abbreviations "HS" refer to Hackerspace, "MS" to Makerspace, and "FL" to FabLab. The unit of analysis is the NUTS3-year in the urban subsample, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.17: First-time inventors & individual applicants (urban, NUTS3-year)

	First-Time				Ind. Patents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HS or MS	MS & HS	MS	FL	HS or MS	MS & HS	MS	FL
τ_0 Base	-0.143**	-0.140***	-0.0809	-0.0864	0.0893	-0.0552	0.0391	-0.169**
	(0.0441)	(0.0351)	(0.0644)	(0.0480)	(0.0970)	(0.0830)	(0.0662)	(0.0530)
τ_1 Base	0.0961***	0.0213	0.0923*	-0.00284	-0.108	-0.157**	-0.0508	-0.103**
	(0.0291)	(0.0538)	(0.0428)	(0.0336)	(0.0752)	(0.0595)	(0.0472)	(0.0341)
τ_2 Base	0.305**	0.175	0.223**	0.0276	0.130	0.101	0.133	0.0973
	(0.108)	(0.0914)	(0.0744)	(0.0555)	(0.110)	(0.102)	(0.0698)	(0.0616)
τ_3 Base	-0.213***	-0.206***	-0.110*	-0.0581	-0.294	-0.170	-0.104	0.146*
	(0.0354)	(0.0345)	(0.0515)	(0.0516)	(0.181)	(0.144)	(0.127)	(0.0602)
τ_4 Base	0.0828	-0.0781	0.117*	-0.0749	-0.215	-0.262	-0.229**	-0.292*
	(0.0923)	(0.109)	(0.0581)	(0.0855)	(0.123)	(0.167)	(0.0835)	(0.143)
τ_5 Base	0.144*	0.155	0.140	0.151	-0.0703	-0.0892	-0.158	-0.205*
	(0.0676)	(0.119)	(0.0760)	(0.113)	(0.288)	(0.217)	(0.260)	(0.0990)
τ_0 ind.=1	-0.0864	-0.0278	-0.186***	-0.143**	-0.169**	-0.0847***	-0.336**	0.0893
	(0.0480)	(0.0161)	(0.0429)	(0.0441)	(0.0530)	(0.0224)	(0.110)	(0.0970)
τ_1 ind.=1	-0.00284	0.0853***	-0.129	0.0961***	-0.103**	0.0255	-0.279***	-0.108
	(0.0336)	(0.0232)	(0.0734)	(0.0291)	(0.0341)	(0.0216)	(0.0514)	(0.0752)
τ_2 ind.=1	0.0276	0.100**	-0.0595	0.305**	0.0973	0.138***	0.0487	0.130
	(0.0555)	(0.0328)	(0.0719)	(0.108)	(0.0616)	(0.0228)	(0.0760)	(0.110)
τ_3 ind.=1	-0.0581	0.0765***	-0.193***	-0.213***	0.146*	0.239***	0.0539	-0.294
	(0.0516)	(0.0209)	(0.0457)	(0.0354)	(0.0602)	(0.0320)	(0.106)	(0.181)
τ_4 ind.=1	-0.0749	0.169***	-0.319	0.0828	-0.292*	-0.251***	-0.333	-0.215
	(0.0855)	(0.0259)	(0.168)	(0.0923)	(0.143)	(0.0614)	(0.276)	(0.123)
τ_5 ind.=1	0.151	0.136**	0.172	0.144*	-0.205*	-0.268***	-0.121	-0.0703
	(0.113)	(0.0505)	(0.255)	(0.0676)	(0.0990)	(0.0373)	(0.220)	(0.288)
Pre1	-0.0210	-0.0210	-0.0210	-0.0210	-0.0817	-0.0817	-0.0817	-0.0817 8
	(0.0757)	(0.0757)	(0.0757)	(0.0757)	(0.101)	(0.101)	(0.101)	(0.101)
Pre2	-0.0242	-0.0242	-0.0242	-0.0242	-0.0233	-0.0233	-0.0233	-0.0233
	(0.0566)	(0.0566)	(0.0566)	(0.0566)	(0.0555)	(0.0555)	(0.0555)	(0.0555)
Pre3	0.0510	0.0510	0.0510	0.0510	0.0230	0.0230	0.0230	0.0230
	(0.0604)	(0.0604)	(0.0604)	(0.0604)	(0.0682)	(0.0682)	(0.0682)	(0.0682)
Pre4	-0.0123	-0.0123	-0.0123	-0.0123	-0.0455	-0.0455	-0.0455	-0.0455
	(0.0762)	(0.0762)	(0.0762)	(0.0762)	(0.0467)	(0.0467)	(0.0467)	(0.0467)
N	2,941	2,941	2,941	2,941	2,941	2,941	2,941	2,941
FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed number of applications by first-time inventors and by individual applicants. The abbreviations "HS" refer to Hackerspace, "MS" to Makerspace, and "FL" to FabLab. The unit of analysis is the NUTS3-year in the urban subsample, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.18: Organizational applicants & AM-related patents (urban, NUTS3-year)

	Org. Patents				AM patents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HS or MS	MS & HS	MS	FL	HS or MS	MS & HS	MS	FL
τ_0 Base	-0.141***	-0.134***	-0.124***	-0.116***	0.378**	0.224*	0.361***	0.154*
	(0.0237)	(0.0233)	(0.0296)	(0.0233)	(0.140)	(0.112)	(0.0996)	(0.0660)
τ_1 Base	-0.0791**	-0.0942**	-0.0555*	-0.0674*	0.165	0.152	0.247**	0.227***
	(0.0261)	(0.0322)	(0.0259)	(0.0293)	(0.137)	(0.120)	(0.0940)	(0.0506)
τ_2 Base	0.00191	-0.0185	0.0219	0.00323	0.286*	0.0474	0.230*	-0.0940
	(0.0332)	(0.0384)	(0.0282)	(0.0408)	(0.129)	(0.146)	(0.0997)	(0.145)
τ_3 Base	-0.0135	-0.0493	0.0242	-0.0108	0.415*	0.188	0.418***	0.101
	(0.0393)	(0.0338)	(0.0349)	(0.0321)	(0.165)	(0.155)	(0.115)	(0.120)
τ_4 Base	-0.0292	-0.111	0.00616	-0.0872	0.473	0.231	0.374	0.0460
	(0.0360)	(0.0641)	(0.0398)	(0.0635)	(0.278)	(0.252)	(0.196)	(0.225)
τ_5 Base	0.0489	-0.0148	0.0411	-0.0338	0.348	0.176	0.428	0.254
	(0.0386)	(0.0927)	(0.0426)	(0.103)	(0.401)	(0.289)	(0.288)	(0.277)
τ_0 ind.=1	-0.116***	-0.104**	-0.130***	-0.141***	0.154*	0.301***	-0.0617	0.378**
	(0.0233)	(0.0344)	(0.0259)	(0.0237)	(0.0660)	(0.0642)	(0.0795)	(0.140)
τ_1 ind.=1	-0.0674*	-0.0180	-0.127*	-0.0791**	0.227***	0.321***	0.0524	0.165
	(0.0293)	(0.0241)	(0.0590)	(0.0261)	(0.0506)	(0.0339)	(0.0990)	(0.137)
τ_2 ind.=1	0.00323	0.0519*	-0.0551	0.00191	-0.0940	0.146	-0.382	0.286*
	(0.0408)	(0.0205)	(0.0785)	(0.0332)	(0.145)	(0.0978)	(0.223)	(0.129)
τ_3 ind.=1	-0.0108	0.0921**	-0.114**	-0.0135	0.101	0.424***	-0.221	0.415*
	(0.0321)	(0.0318)	(0.0429)	(0.0393)	(0.120)	(0.0795)	(0.153)	(0.165)
τ_4 ind.=1	-0.0872	0.0592*	-0.234	-0.0292	0.0460	0.225	-0.133	0.473
	(0.0635)	(0.0300)	(0.121)	(0.0360)	(0.225)	(0.228)	(0.385)	(0.278)
τ_5 ind.=1	-0.0338	0.0315	-0.121	0.0489	0.254	0.527	-0.112	0.348
	(0.103)	(0.0402)	(0.234)	(0.0386)	(0.277)	(0.395)	(0.368)	(0.401)
Pre1	-0.104*	-0.104*	-0.104*	-0.104*	-0.0109	-0.0109	-0.0109	-0.0109
	(0.0410)	(0.0410)	(0.0410)	(0.0410)	(0.171)	(0.171)	(0.171)	(0.171)
Pre2	-0.0429	-0.0429	-0.0429	-0.0429	0.0220	0.0220	0.0220	0.0220
	(0.0398)	(0.0398)	(0.0398)	(0.0398)	(0.163)	(0.163)	(0.163)	(0.163)
Pre3	-0.0216	-0.0216	-0.0216	-0.0216	-0.0284	-0.0284	-0.0284	-0.0284
	(0.0477)	(0.0477)	(0.0477)	(0.0477)	(0.116)	(0.116)	(0.116)	(0.116)
Pre4	-0.0378	-0.0378	-0.0378	-0.0378	-0.0252	-0.0252	-0.0252	-0.0252
	(0.0520)	(0.0520)	(0.0520)	(0.0520)	(0.114)	(0.114)	(0.114)	(0.114)
N	2,941	2,941	2,941	2,941	2,941	2,941	2,941	2,941
FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed number of applications by organizational applicants and of AM-related patents. The abbreviations "HS" refer to Hackerspace, "MS" to Makerspace, and "FL" to FabLab. The unit of analysis is the NUTS3-year in the urban subsample, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.19: Utility model applications (urban, NUTS3-year)

	(1) HS or MS	(2) HS & MS	(3) MS	(4) FL
τ_0 Base	-0.423**	-0.271*	-0.307***	-0.0630
	(0.145)	(0.125)	(0.0856)	(0.0677)
τ_1 Base	-0.228	-0.103	-0.163	0.00196
	(0.122)	(0.112)	(0.0835)	(0.0705)
τ_2 Base	-0.102	-0.0181	-0.0926	0.0170
	(0.165)	(0.129)	(0.102)	(0.0573)
τ_3 Base	-0.500***	-0.357**	-0.298*	-0.0163
	(0.133)	(0.123)	(0.127)	(0.101)
τ_4 Base	-0.108	-0.0111	-0.0157	0.129
	(0.104)	(0.139)	(0.0799)	(0.107)
τ_5 Base	0.0112	0.120	0.131	0.289
	(0.187)	(0.203)	(0.113)	(0.186)
τ_0 ind.=1	-0.0630	-0.0470	0.0487	-0.423**
	(0.0677)	(0.0301)	(0.141)	(0.145)
τ_1 ind.=1	0.00196	-0.0798***	0.115	-0.228
	(0.0705)	(0.0229)	(0.163)	(0.122)
τ_2 ind.=1	0.0170	-0.0791	0.132	-0.102
	(0.0573)	(0.0411)	(0.111)	(0.165)
τ_3 ind.=1	-0.0163	0.0660	-0.0987	-0.500***
	(0.101)	(0.0601)	(0.188)	(0.133)
τ_4 ind.=1	0.129	0.123	0.135	-0.108
	(0.107)	(0.117)	(0.178)	(0.104)
τ_5 ind.=1	0.289	0.280**	0.301	0.0112
	(0.186)	(0.0897)	(0.415)	(0.187)
Pre1	-0.0249	-0.0249	-0.0249	-0.0249
	(0.0752)	(0.0752)	(0.0752)	(0.0752)
Pre2	-0.00150	-0.00150	-0.00150	-0.00150
	(0.108)	(0.108)	(0.108)	(0.108)
Pre3	0.0115	0.0115	0.0115	0.0115
	(0.0918)	(0.0918)	(0.0918)	(0.0918)
Pre4	0.0604	0.0604	0.0604	0.0604
	(0.0785)	(0.0785)	(0.0785)	(0.0785)
N	2,941	2,941	2,941	2,941
FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed number of utility model applications. The abbreviations "HS" refer to Hackerspace, "MS" to Makerspace, and "FL" to FabLab. The unit of analysis is the NUTS3-year in the urban subsample, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.20: Patent quality measures (urban, NUTS3-year)

	Fw. Citations				Areas Cited			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HS or MS	MS & HS	MS	FL	HS or MS	MS & HS	MS	FL
τ_0 Base	0.00526	0.0181	-0.0140	0.00485	-0.0763***	-0.0591***	-0.0613***	-0.0341**
	(0.0458)	(0.0430)	(0.0275)	(0.0300)	(0.0179)	(0.0158)	(0.0147)	(0.0125)
τ_1 Base	-0.0212	0.0882	-0.00567	0.117	-0.0328	-0.0262	-0.00300	0.00351
	(0.0277)	(0.0825)	(0.0209)	(0.0818)	(0.0288)	(0.0224)	(0.0187)	(0.0200)
τ_2 Base	-0.0488**	0.0305	-0.0488*	0.0520	-0.0310	-0.0216	-0.00739	0.0132
	(0.0167)	(0.0455)	(0.0224)	(0.0479)	(0.0235)	(0.0191)	(0.0273)	(0.0213)
τ_3 Base	0.0199	0.0453	0.0313	0.0714	0.0121	0.0166	0.0318	0.0459**
	(0.0390)	(0.0588)	(0.0294)	(0.0468)	(0.0213)	(0.0195)	(0.0191)	(0.0148)
τ_4 Base	0.0420	0.0884	0.0230	0.0763	0.0232	0.0274	0.0420*	0.0519
	(0.0587)	(0.0660)	(0.0462)	(0.0412)	(0.0142)	(0.0312)	(0.0206)	(0.0327)
τ_5 Base	0.170**	0.103*	0.106*	0.0106	0.0720***	0.0620*	0.0297	0.00622
	(0.0600)	(0.0499)	(0.0467)	(0.0475)	(0.0166)	(0.0256)	(0.0349)	(0.0192)
τ_0 ind.=1	0.00485	-0.0297**	0.0557	0.00526	-0.0341**	-0.0307*	-0.0248	-0.0763***
	(0.0300)	(0.00943)	(0.0381)	(0.0458)	(0.0125)	(0.0130)	(0.0219)	(0.0179)
τ_1 ind.=1	0.117	-0.0179	0.261***	-0.0212	0.00351	0.0234*	-0.0410	-0.0328
	(0.0818)	(0.0298)	(0.0268)	(0.0277)	(0.0200)	(0.0112)	(0.0362)	(0.0288)
τ_2 ind.=1	0.0520	-0.0489	0.173***	-0.0488**	0.0132	0.0281**	-0.00475	-0.0310
	(0.0479)	(0.0454)	(0.0364)	(0.0167)	(0.0213)	(0.00997)	(0.0274)	(0.0235)
τ_3 ind.=1	0.0714	0.0518	0.0909***	0.0199	0.0459**	0.0672***	0.0247	0.0121
	(0.0468)	(0.0351)	(0.0261)	(0.0390)	(0.0148)	(0.0181)	(0.0181)	(0.0213)
τ_4 ind.=1	0.0763	-0.00547	0.158***	0.0420	0.0519	0.0702***	0.0336	0.0232
	(0.0412)	(0.0712)	(0.0376)	(0.0587)	(0.0327)	(0.0130)	(0.0636)	(0.0142)
τ_5 ind.=1	0.0106	0.0252	-0.00885	0.170**	0.00622	-0.0232	0.0455**	0.0720***
	(0.0475)	(0.0547)	(0.0819)	(0.0600)	(0.0192)	(0.0303)	(0.0149)	(0.0166)
Pre1	-0.0548	-0.0548	-0.0548	-0.0548	-0.0549	-0.0549	-0.0549	-0.0549
	(0.0505)	(0.0505)	(0.0505)	(0.0505)	(0.0282)	(0.0282)	(0.0282)	(0.0282)
Pre2	0.00469	0.00469	0.00469	0.00469	-0.0398	-0.0398	-0.0398	-0.0398
	(0.0565)	(0.0565)	(0.0565)	(0.0565)	(0.0259)	(0.0259)	(0.0259)	(0.0259)
Pre3	-0.0444	-0.0444	-0.0444	-0.0444	-0.0582	-0.0582	-0.0582	-0.0582
	(0.0387)	(0.0387)	(0.0387)	(0.0387)	(0.0320)	(0.0320)	(0.0320)	(0.0320)
Pre4	0.00651	0.00651	0.00651	0.00651	-0.0716**	-0.0716**	-0.0716**	-0.0716**
	(0.0332)	(0.0332)	(0.0332)	(0.0332)	(0.0253)	(0.0253)	(0.0253)	(0.0253)
N	2,941	2,941	2,941	2,941	2,941	2,941	2,941	2,941
FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the method by Borusyak, Jaravel, and Spiess (2024) for the log-transformed number of forward citations per patent and of average areas cited per patent. The abbreviations "HS" refer to Hackerspace, "MS" to Makerspace, and "FL" to FabLab. The unit of analysis is the NUTS3-year in the urban subsample, with not-yet-treated and never treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

C.2.3 Tables – Robustness Checks

Table C.21: Regression results for the 2010-2019 subsample (FUA-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	-0.00904	-0.0148	-0.0456	0.00592	-0.00423	0.0461	0.00439	0.00378	-0.00717
	(0.0215)	(0.0228)	(0.0432)	(0.0726)	(0.0259)	(0.0880)	(0.0635)	(0.0312)	(0.0213)
GAverage	-0.0227	-0.0323	-0.0492	-0.0288	-0.0231	0.0247	0.0179	-0.0144	-0.00789
	(0.0195)	(0.0210)	(0.0359)	(0.0537)	(0.0228)	(0.0760)	(0.0524)	(0.0288)	(0.0202)
G2011	-0.0121	-0.00871	0.0564	0.157	-0.00376	0.0800	0.0196	0.117**	-0.0292
	(0.0383)	(0.0374)	(0.0491)	(0.0854)	(0.0379)	(0.200)	(0.0994)	(0.0396)	(0.0364)
G2012	-0.0279	-0.0130	-0.160*	0.357	-0.0319	0.153	-0.207	0.00266	0.0130
	(0.0368)	(0.0362)	(0.0738)	(0.267)	(0.0926)	(0.202)	(0.154)	(0.0717)	(0.0485)
G2013	0.0140	-0.00211	-0.130	0.103	0.0250	-0.0616	-0.0375	-0.107	-0.0564
	(0.0368)	(0.0452)	(0.138)	(0.1000)	(0.0498)	(0.206)	(0.137)	(0.0744)	(0.0352)
G2014	0.0260	-0.00432	-0.0857	-0.305*	0.0409	0.211	0.0914	0.0276	0.0834
	(0.0563)	(0.0600)	(0.111)	(0.134)	(0.0611)	(0.173)	(0.199)	(0.0866)	(0.0744)
G2015	-0.000418	-0.00659	0.189**	-0.162*	0.00914	-0.0695	-0.136	-0.0273	0.00201
	(0.0720)	(0.0729)	(0.0643)	(0.0780)	(0.0758)	(0.236)	(0.0889)	(0.0462)	(0.0180)
G2016	-0.00247	0.0111	-0.0674	-0.256*	-0.00169	-0.0257	0.243*	-0.0214	-0.0158
	(0.0459)	(0.0475)	(0.0730)	(0.128)	(0.0474)	(0.167)	(0.102)	(0.0636)	(0.0375)
G2017	-0.0846*	-0.0796	-0.155	-0.0270	-0.0827*	-0.0822	-0.143	0.0340	-0.0144
	(0.0417)	(0.0465)	(0.130)	(0.164)	(0.0372)	(0.158)	(0.134)	(0.0333)	(0.0273)
G2018	-0.193***	-0.308***	0.100	0.180	-0.274***	-0.0103	0.00988	-0.219	-0.102
	(0.0388)	(0.0577)	(0.0545)	(0.129)	(0.0723)	(0.0725)	(0.0432)	(0.254)	(0.165)
G2019	-0.171***	-0.187***	-0.221***	0.174***	-0.220***	-0.115	-0.0238	-0.118***	0.0556***
	(0.00939)	(0.0144)	(0.0258)	(0.0378)	(0.0155)	(0.167)	(0.0386)	(0.0186)	(0.00944)
N	930	930	930	930	930	930	930	930	930
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables for the years 2010-2019. The unit of analysis is the FUA-year, with not-yet-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the FUA level. Year and FUA fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C.22: Regression results for the 2010-2019 subsample (NUTS3-year)

	(1) Total	(2) NoChem	(3) First-time	(4) Ind. Pat.	(5) Org. Pat.	(6) AM Pat.	(7) UM	(8) Fw. Cit.	(9) Areas Cit.
ATT	0.000812	-0.00514	-0.111	0.0736	-0.00751	0.0836	-0.0903	0.0225	0.0377
	(0.0276)	(0.0287)	(0.0571)	(0.0812)	(0.0327)	(0.0782)	(0.0625)	(0.0302)	(0.0216)
G2011	-0.0252	0.0198	-0.144	0.177	-0.0521	0.262	-0.228*	-0.0657	0.0138
	(0.0604)	(0.0676)	(0.141)	(0.297)	(0.0930)	(0.197)	(0.113)	(0.0949)	(0.0338)
G2012	-0.0699	-0.101	-0.129	-0.169***	-0.0509	0.0779	-0.172	0.0141	-0.0463
	(0.0965)	(0.0827)	(0.0741)	(0.0453)	(0.0874)	(0.112)	(0.176)	(0.0505)	(0.0730)
G2013	-0.0978	-0.149*	-0.0880	-0.0580	-0.0897	-0.197	0.0424	0.0944	0.143*
	(0.0780)	(0.0694)	(0.193)	(0.181)	(0.102)	(0.224)	(0.154)	(0.105)	(0.0572)
G2014	0.0251	0.00177	-0.194	-0.00350	0.0113	0.00445	0.0398	0.0320	0.0911*
	(0.0487)	(0.0481)	(0.108)	(0.158)	(0.0557)	(0.125)	(0.140)	(0.0372)	(0.0438)
G2015	0.0355	0.0304	0.0797	0.0331	0.0270	0.352	0.0656	0.0520	0.0538**
	(0.0595)	(0.0547)	(0.207)	(0.180)	(0.0525)	(0.204)	(0.209)	(0.0602)	(0.0185)
G2016	0.116*	0.132*	-0.142	0.144	0.116*	0.00590	-0.0633	0.0801	-0.0833
	(0.0545)	(0.0626)	(0.124)	(0.132)	(0.0585)	(0.212)	(0.101)	(0.0548)	(0.0456)
G2017	0.0836**	0.0889***	0.0554	0.695***	0.0431	0.312	-0.854***	-0.0253	0.0416
	(0.0290)	(0.0269)	(0.142)	(0.162)	(0.0310)	(0.291)	(0.250)	(0.0733)	(0.0605)
G2018	-0.0607	-0.0567	-0.144	0.147	-0.0625	0.335*	-0.0845	-0.0593	0.000358
	(0.107)	(0.113)	(0.0980)	(0.196)	(0.113)	(0.157)	(0.196)	(0.0673)	(0.0428)
N	3,970	3,970	3,970	3,970	3,970	3,970	3,970	3,970	3,970
FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents the results of the staggered DiD analysis following the CS method for all log-transformed dependent variables for the years 2010-2019. The unit of analysis is the NUTS3-year, with not-yet-treated observations serving as the control group. Population and GDP are included as control variables, and standard errors are clustered at the NUTS3 level. Year and NUTS3 fixed effects are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

C.3 Questionnaires and Results of the Online Surveys

C.3.1 Manager Survey – Questionnaire

Q1: I have read the information above and would like to participate in the survey.

- Yes
- No

General Information

Q2: What year did the FabLab you primarily manage first open?

Q3: Was your FabLab previously a Maker-Space (e.g., it did not adhere to the Fab Charta before or did not offer 3D-printing)?

- Yes
- No
- Other (please specify)

Foundation and Financing

Q4: What were the reasons for specifically founding a "FabLab"? (Select all that apply)

- Conviction of the mission and idea of FabLabs (i.e., openness)
- Focus on "Digital Fabrication" (i.e., 3D-printing, CNC-milling)
- Participation in the global FabLab network
- Teaching Opportunities (i.e., Fab Academy)
- Other (please specify)

Q5: How is your FabLab financed? (Select all that apply)

- University funding (department)
- University funding (student group)
- Membership fees (association)
- Private Funding/Donations
- Government Grants (please specify)
- Other sources (please specify)

Networking and Collaboration

Q6: Is your FabLab part of a local innovation or entrepreneurship ecosystem (i.e., Erlanger Innovations- und Gründerzentrum)?

- Yes
- No

Q7: What is the name of this innovation network or ecosystem?

Q8: With which external institutions do you collaborate e.g., for funding, collaborative research etc.? (Select all that apply)

- Universities
- NGOs
- Large enterprises
- Small-/medium enterprises or start-ups
- Political Entities (e.g., EU, Federal level, municipality)
- None
- Other (please specify)

Q9: What are the purposes of these collaborations? (Select all that apply)

- Joint research projects
- Exchange of resources and/or know-how
- Organization of joint workshops and trainings
- Providing internships and project opportunities for students
- Funding opportunities for the FabLab
- Other (please specify)

Management and openness

Q10: Who primarily manages the FabLab? (Select all that apply)

- Volunteers
- Paid staff
- Other (please specify)

Q11: How do you promote your FabLab? (Select all that apply)

- By participating in fairs and conferences
- Through public relations (e.g., local newspapers)
- By collaborating with local firms
- By collaborating with universities
- Through social media marketing (e.g., Facebook or Instagram)

- Other (please specify)

Q12: Is your FabLab open to everyone?

- Yes
- No
- Other (please specify)

Operation and Usage

Q13: How many days per week is the Lab usually open to everyone?

- 1 day
- 2 days
- 3 days
- 4 days
- 5 days
- 6 days
- 7 days

Q14: How many users does your FabLab have on average per week?

- 0-5
- 6-20
- 21-30
- 31-40
- 41-50
- More than 50

Q15: How frequently do the following groups typically use the FabLab per week?

- Students
  - Never
  - Sometimes
  - About half the time
  - Most of the time

  - Always
- Faculty Members
  - Never
  - Sometimes
  - About half the time
  - Most of the time
  - Always
- Inventors/Entrepreneurs
  - Never
  - Sometimes
  - About half the time
  - Most of the time
  - Always
- Firms
  - Never
  - Sometimes
  - About half the time
  - Most of the time
  - Always
- Hobbyists
  - Never
  - Sometimes
  - About half the time
  - Most of the time
  - Always
- School kids
  - Never
  - Sometimes
  - About half the time
  - Most of the time
  - Always

- Other (please specify)
  - Never
  - Sometimes
  - About half the time
  - Most of the time
  - Always

Outputs and Documentation

Q16: What are the outcomes of the users' projects? (Select all that apply)

- Wrote a qualification paper (seminar, thesis, PhD)
- Published a scientific paper based on a FabLab project
- Created a functional prototype
- Launched a product or business
- Completed a personal project
- Wrote a software
- Applied for a patent or utility model
- Collaborated on a non-profit project
- Other (please specify)

Q17: How are projects and activities in the FabLab documented?

- Not at all
- Via an internal sharing platform
- Via an external sharing platform
- Other (please specify)

Activities and Collaborations

Q18: How often do collaborations occur between users?

- Never
- Sometimes
- About half the time

- Most of the time
- Always
- Unknown

Q19: Do you actively promote knowledge exchange among the users?

- Yes
- No

Q20: How do you promote knowledge exchange among users? (Select all that apply)

- By organizing regular networking meetings
- By organizing workshops and trainings
- By inviting to document projects and activities using a communication platform (internal/external)
- Other (please specify)

Offerings and Services

Q21: What offerings are available in your FabLab? (Select all that apply)

- Technical courses (e.g., handling machines)
- Courses or support on entrepreneurship
- Courses or support on patent/utility model applications
- Courses for schools
- Co-working space
- Fab Academy courses
- Contract Work
- Other (please specify)

Q22: How much do you charge on average for the participation in technical courses?

- 1€-10€
- 11€-20€
- more than 20€
- Other (please specify)

Challenges and improvements

Q23: What are the primary challenges you face as a FabLab manager? (Select all that apply)

- Securing adequate funding to sustain the lab (e.g., for rent, equipment, and tools)
- Having enough volunteers to operate the lab (e.g., to offer trainings)
- Attracting and retaining users
- Other (please specify)

Q24: Do you have any additional insights or suggestions about the FabLab that you would like to share?

C.3.2 User Survey – Questionnaire

Q1: I have read the information above and would like to participate in the survey.

- Yes
- No

Usage

Q2: Do you use other FabLabs as well?

- Yes
- No

Q3: How did you learn about the FabLab? (Select all that apply)

- Internet research
- Recommendation from friends / professors / colleagues
- Newspaper article
- Social media
- Other (please specify)

Q4: How far is the FabLab from your place of residence?

- 0-5 km

- 5-10 km
- 10-20 km
- >20 km

Q5: What do you use the FabLab for? (Select all that apply)

- Personal projects or hobbies
- Working on seminar / final theses for your studies
- Learning new skills or technologies
- Prototyping new products or inventions
- Entrepreneurial activities (e.g., starting a business)
- Community involvement (e.g., participating in events, networking)
- Collaborating on projects
- Other (please specify)

Q6: How often do you usually visit the FabLab?

- Daily
- Several times a week
- Once a week
- Every two weeks
- Once a month
- Other (please specify)

Q7: Do you usually work with others at the FabLab?

- No, I always work alone
- No, but I exchange ideas with others
- Yes, others help, but I work on my own ideas
- Yes, I collaborate on projects with others

Q8: What I value most about the FabLab is...

- ... the open access to the tools and technology
- ... the community and networking

- ... learning new skills
- ...other (please specify)

Q9: Have you participated in workshops or trainings offered by the FabLab?

- Yes
- No

Projects

Q10: What are the outcomes of your FabLab activities? (Select all that apply)

- Created a functional prototype
- Developed new skills or knowledge
- Launched a product or business
- Completed a personal project
- Applied for a patent or utility model
- Collaborated on a non-profit project
- Enhanced my professional network
- No specific outcomes yet
- Wrote a qualification paper (i.e., seminar, thesis, PhD)
- Wrote a software program
- Other (please specify)

Q11: How do you share the outcomes of your FabLab activities? (Select all that apply)

- Via a code sharing platform (e.g., GitHub)
- Via an internal (sharing) platform
- Via a product sharing platform (e.g., Thingiverse)
- By applying for utility models/patents
- By documenting it on the FabLab website
- Not at all
- Other (please specify)

Q12: Which category best describes the subject area of your end results from activities at the FabLab? (Select all that apply)

- Engineering and technology
- Art and design
- Environmental sustainability
- Software
- Medical applications/products/processes
- Other (please specify)

Demographics

Q13: How old are you?

- Under 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older

Q14: What is your gender?

- Male
- Female
- Non-binary / third gender
- Prefer not to say

Q15: What is the highest level of education you have achieved?

- Less than high school
- High school graduate

- Bachelor degree
- Master degree / Diploma
- Professional degree
- Doctorate
- Other (please specify)

Q16: Which of the following best describes your current employment status?

- Employed at a company
- Employed in public service (e.g., public administration, education, healthcare)
- Self-employed
- Unemployed looking for work
- Unemployed not looking for work
- Retired
- Student
- Disabled

Q17: How many employees work at your company?(If unsure, please provide a rough estimate.)

- 1-10
- 11-50
- 51-200
- 201-500
- More than 500

Q18: Which of the following best describes your current profession?

- Agriculture and forestry, fisheries
- Architecture and civil engineering
- Art and design
- Business administration
- Crafts (e.g., carpenter, electrician), installation, maintenance, repair

- Education
- Electrical engineerin
- Healthcare
- Information technology and mathematics
- Law
- Mechanical engineering
- Natural sciences
- Production and logistics
- Public service
- Sales, office and administration
- Service industry
- Social sciences
- Sports/sport sciences
- Other

Overall experience and improvements

Q19: In your opinion, how could the FabLab improve? (Select all that apply)

- More advanced tools and machines
- Implementation of knowledge management technology (i.e., wiki, intranet)
- More workshops and trainings
- Longer operating hours
- More events with firms / hosting of Venture Capital pitches
- Adjustment of usage/membership fees
- More community events and activities
- Other (please specify)

Q20: Please rate your overall experience with the FabLab from 1 to 5 building blocks, where 1 means "Bad" and 5 means "Great."

- 1

- 2
- 3
- 4
- 5

Q21: Could you please specify the main reasons for your dissatisfaction? (Select all that apply)

- Limited access to tools and machines
- Insufficient hours of operation
- High costs or fees
- Poor maintenance of facilities or equipment
- Other (please specify)

Q22: Is there anything else you would like to share about your experience at the FabLab or suggestions for improvement?

C.3.3 Online Survey Results

Manager Survey – Results

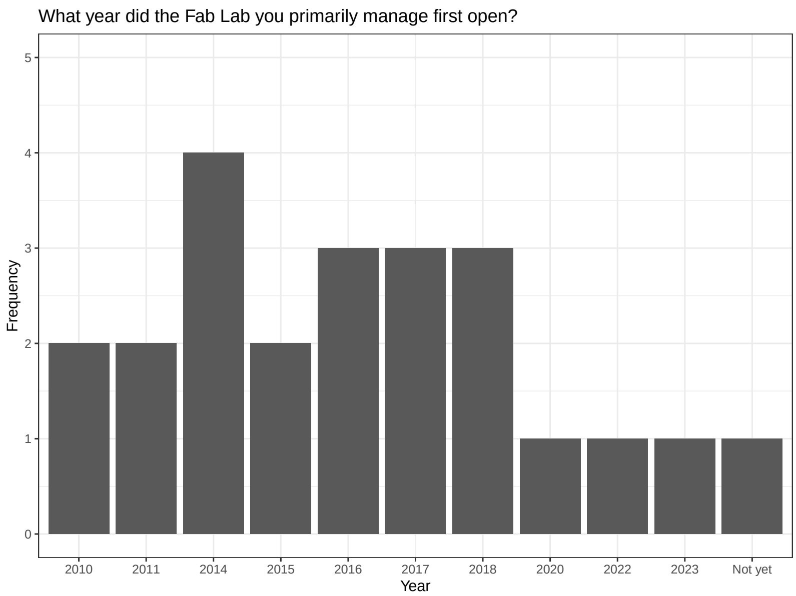


Figure C.23: Manager survey - question 2

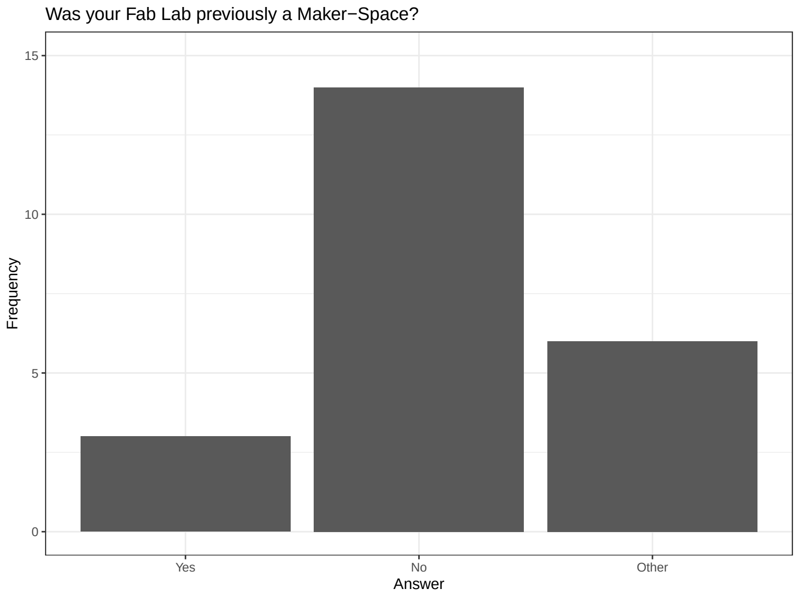


Figure C.24: Manager survey - question 3

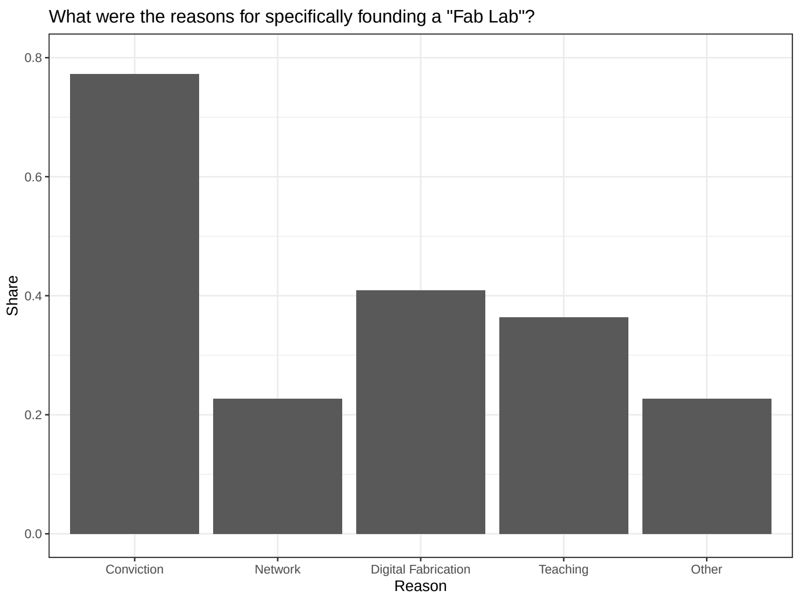


Figure C.25: Manager survey - question 4

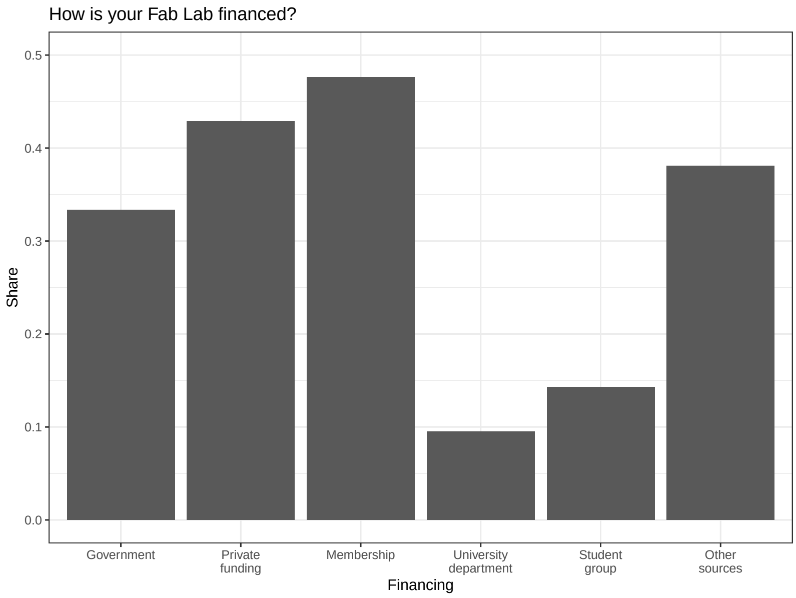


Figure C.26: Manager survey - question 5

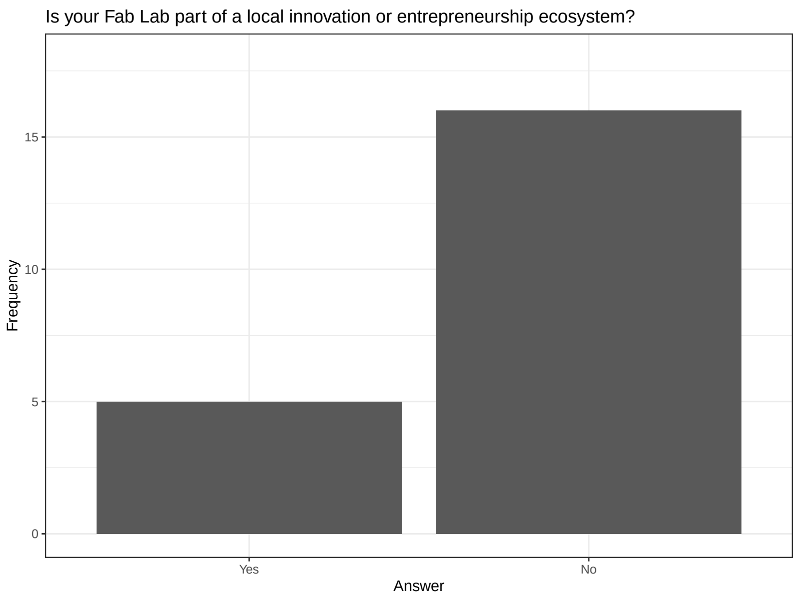


Figure C.27: Manager survey - question 6

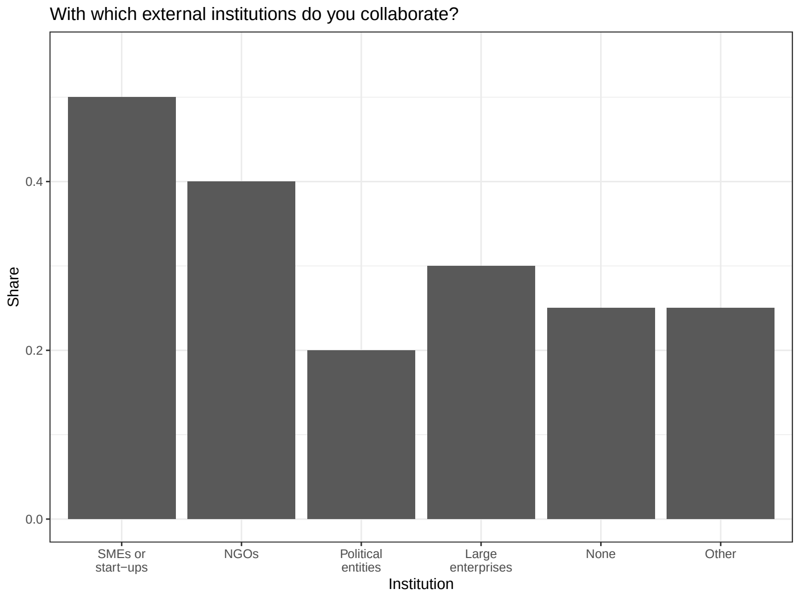


Figure C.28: Manager survey - question 8

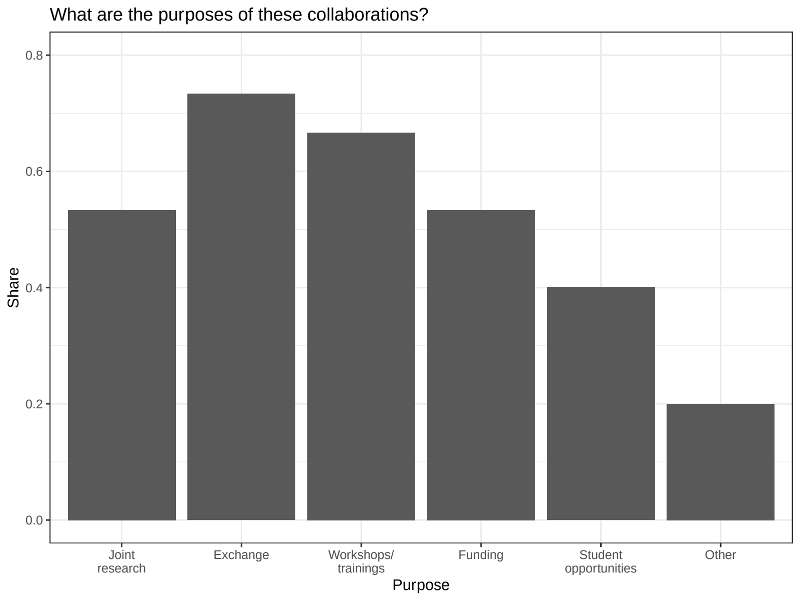


Figure C.29: Manager survey - question 9

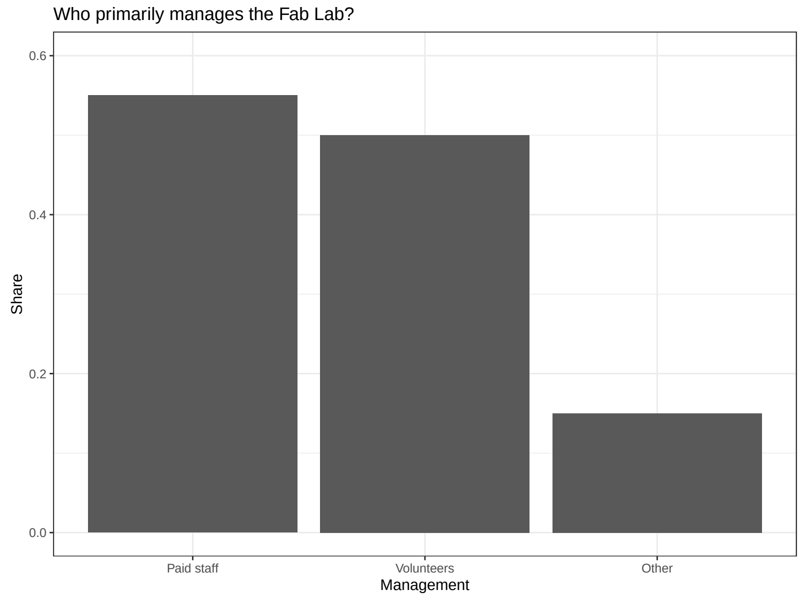


Figure C.30: Manager survey - question 10

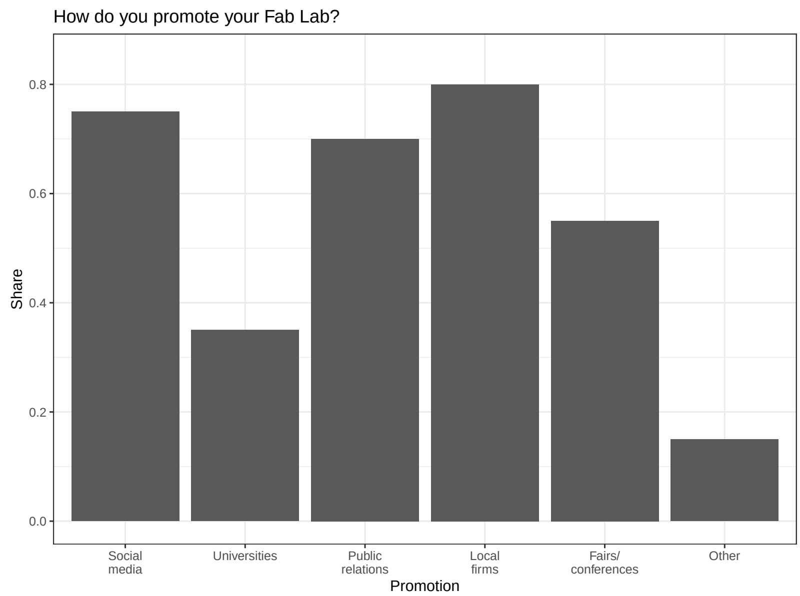


Figure C.31: Manager survey - question 11

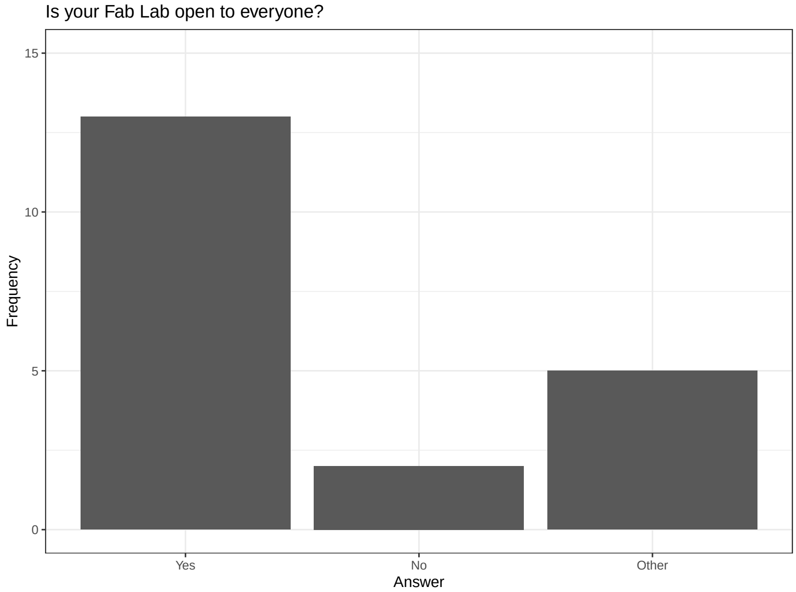


Figure C.32: Manager survey - question 12

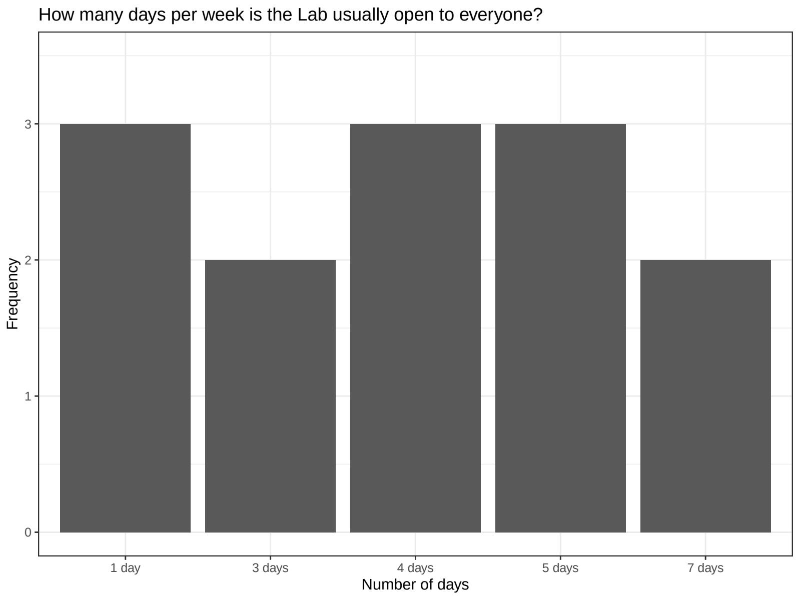


Figure C.33: Manager survey - question 13

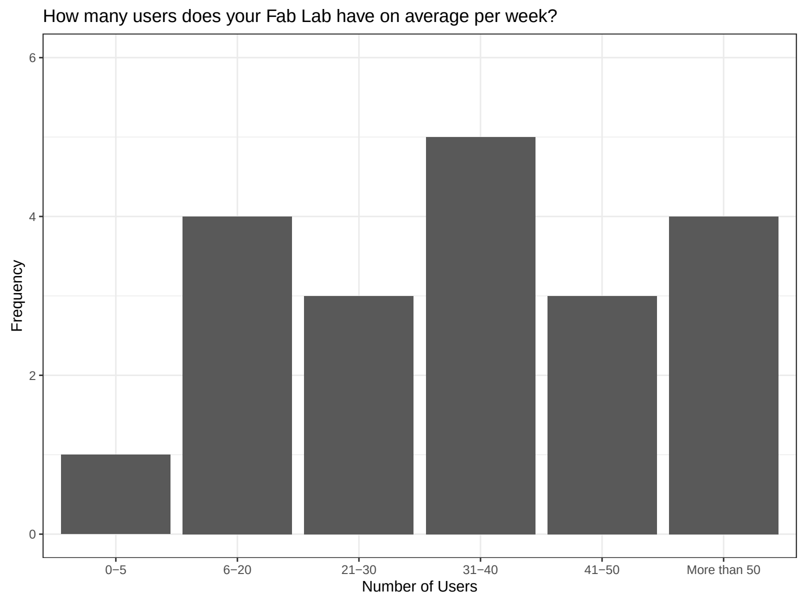


Figure C.34: Manager survey - question 14

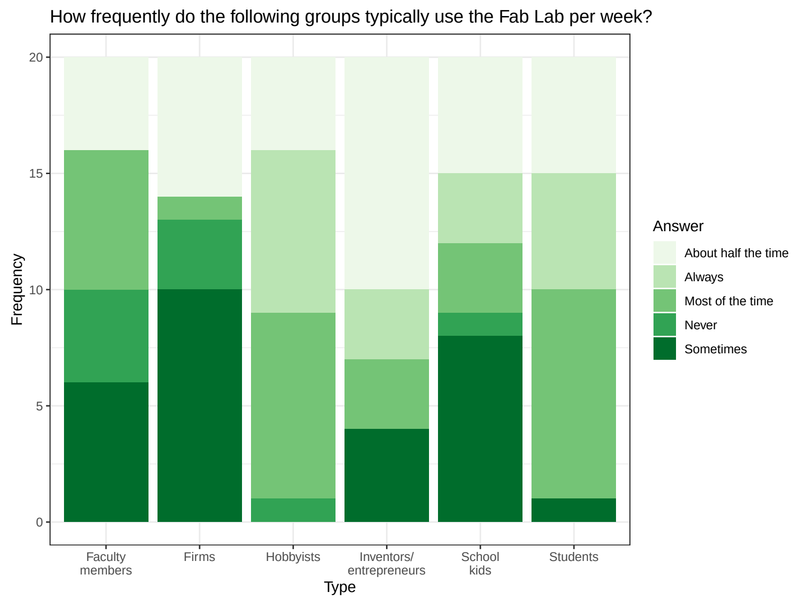


Figure C.35: Manager survey - question 15

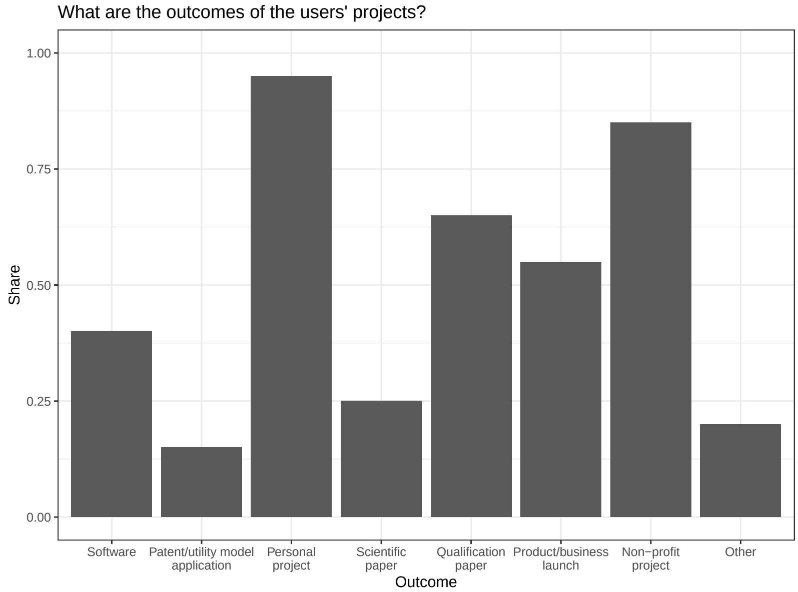


Figure C.36: Manager survey - question 16

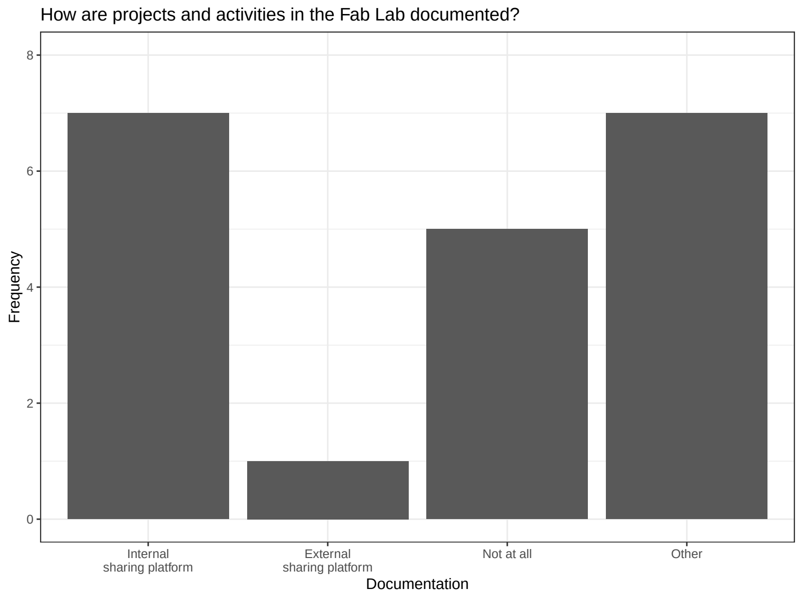


Figure C.37: Manager survey - question 17

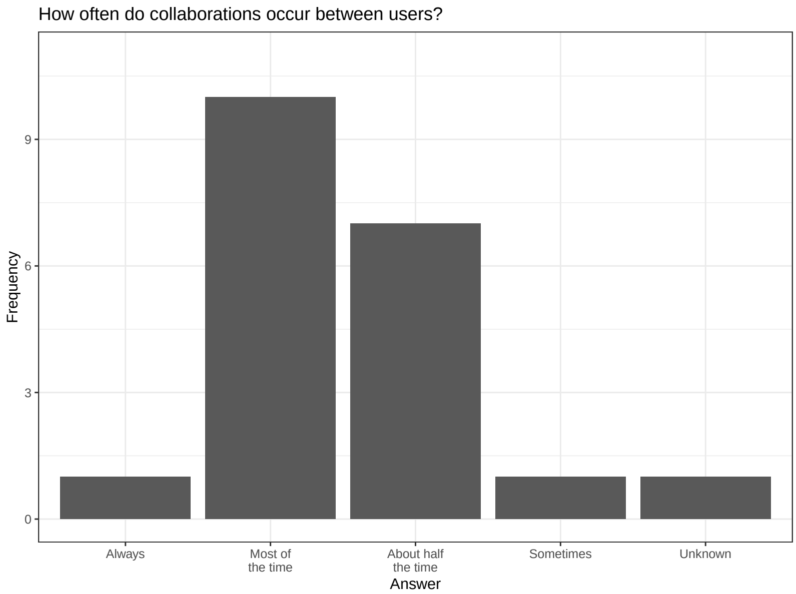


Figure C.38: Manager survey - question 18

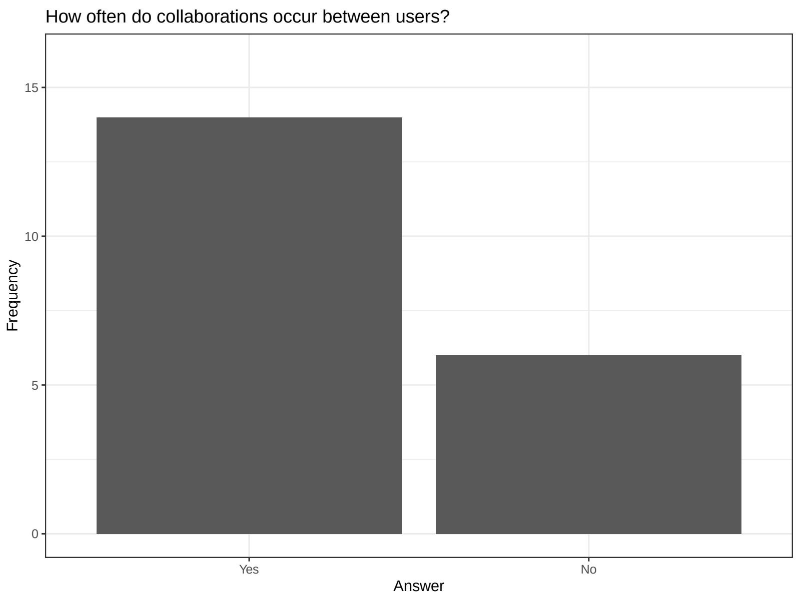


Figure C.39: Manager survey - question 19

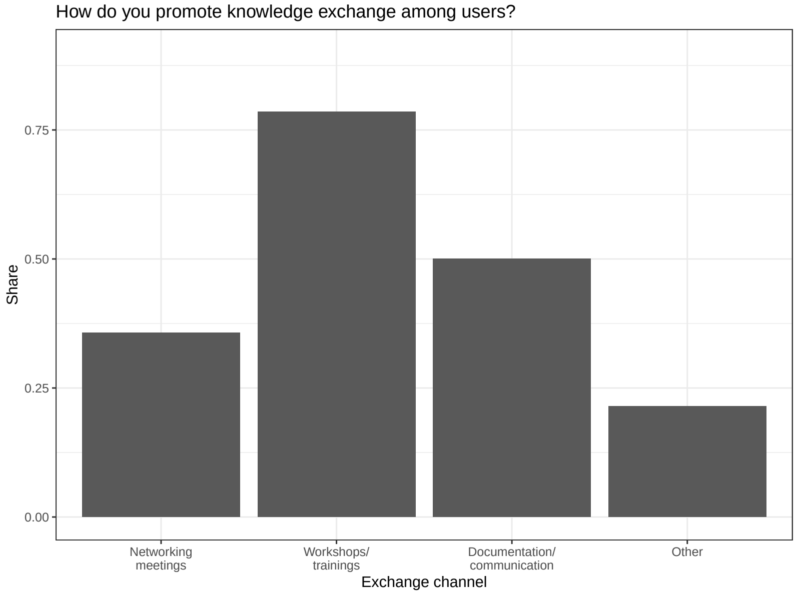


Figure C.40: Manager survey - question 20

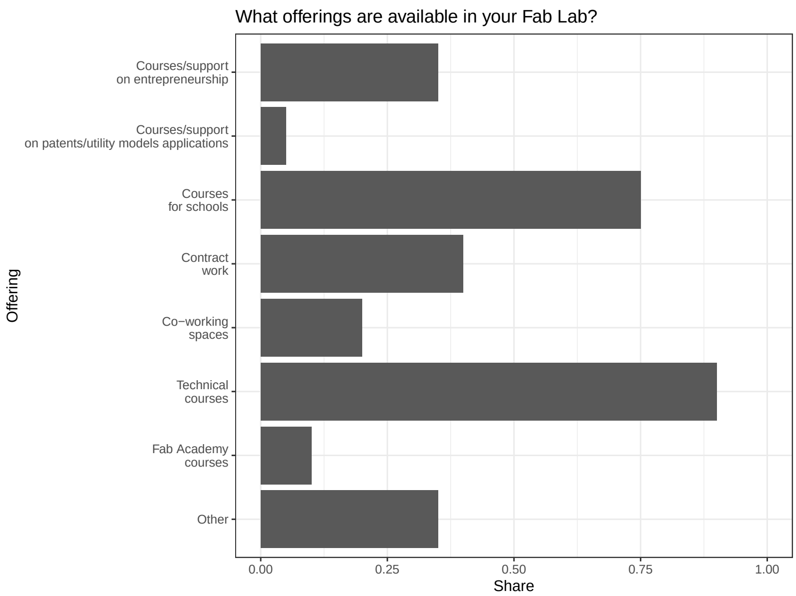


Figure C.41: Manager survey - question 21

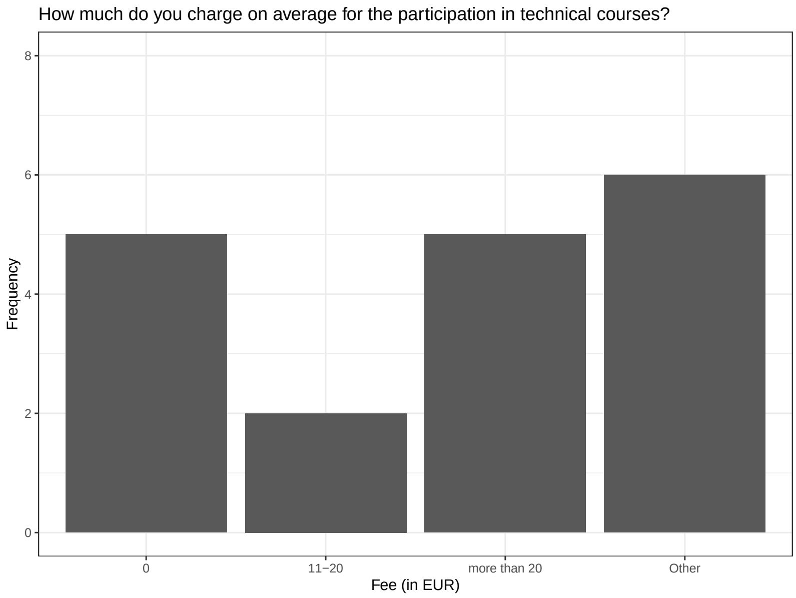


Figure C.42: Manager survey - question 22

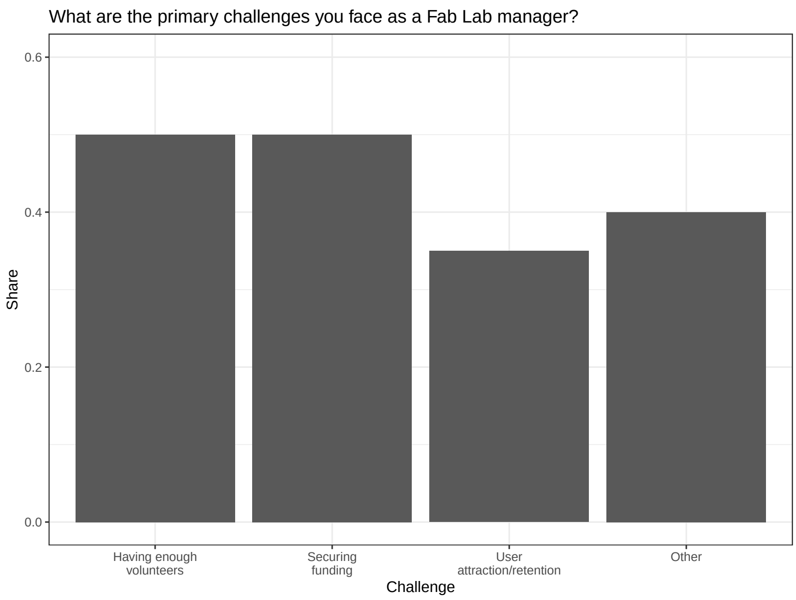


Figure C.43: Manager survey - question 23

User Survey – Results

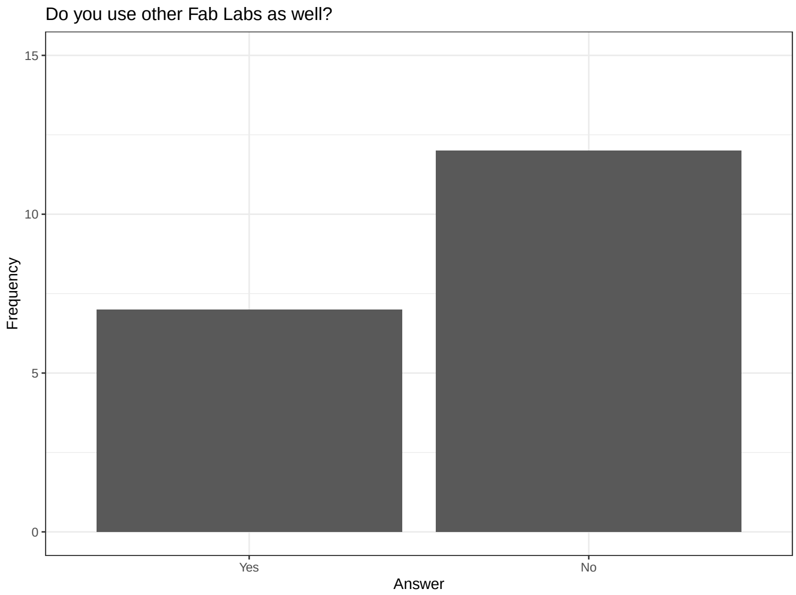


Figure C.44: User survey - question 2

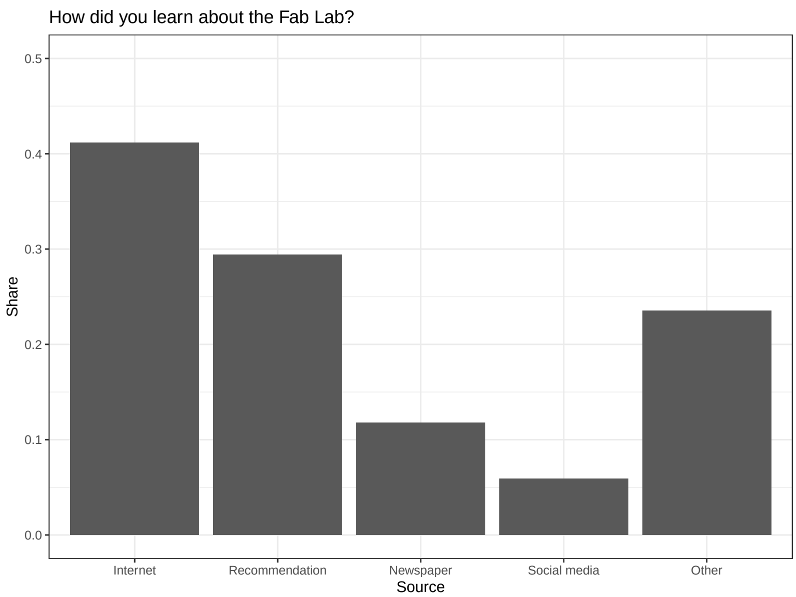


Figure C.45: User survey - question 3

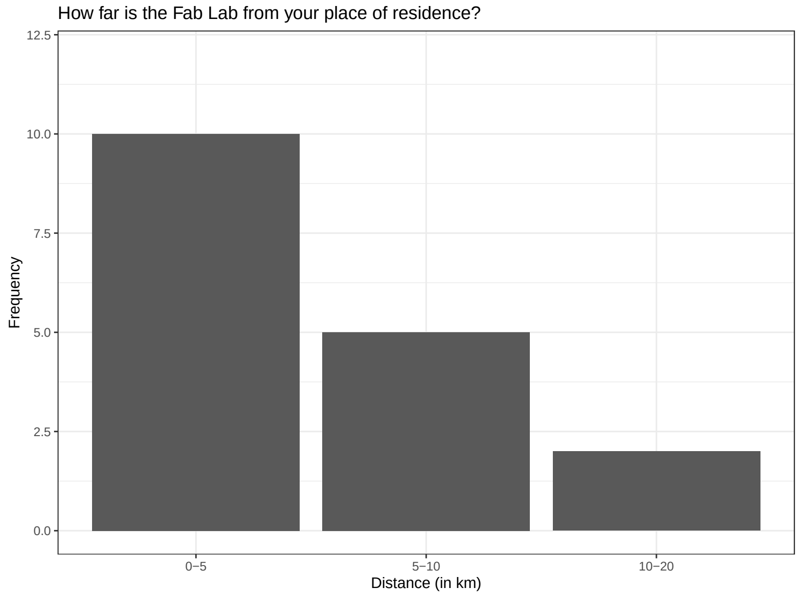


Figure C.46: User survey - question 4

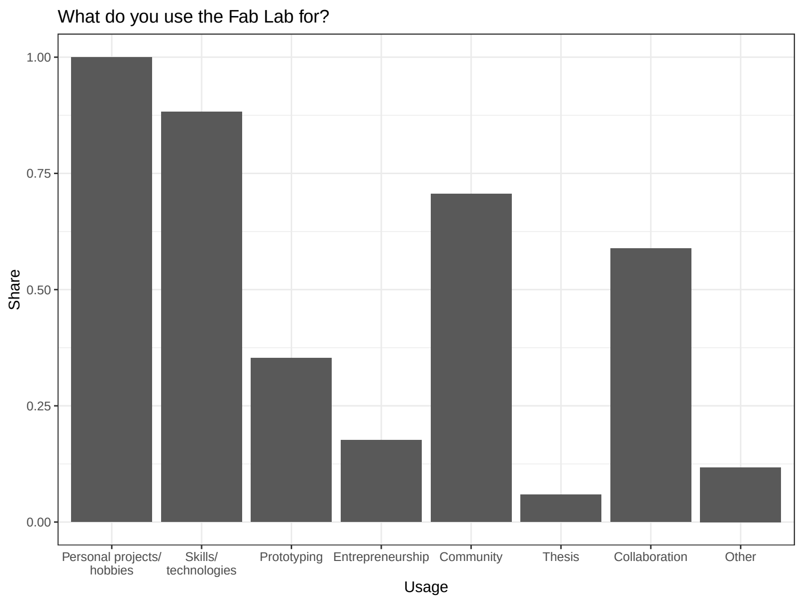


Figure C.47: User survey - question 5

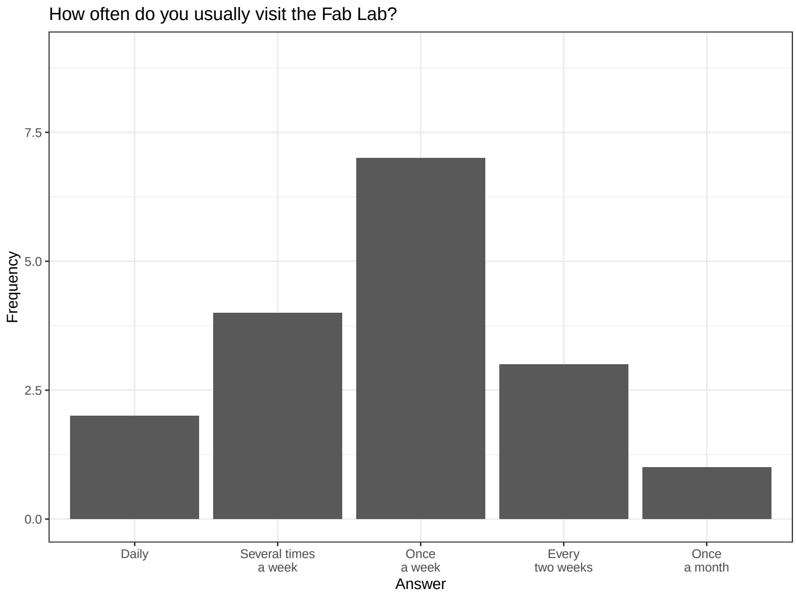


Figure C.48: User survey - question 6

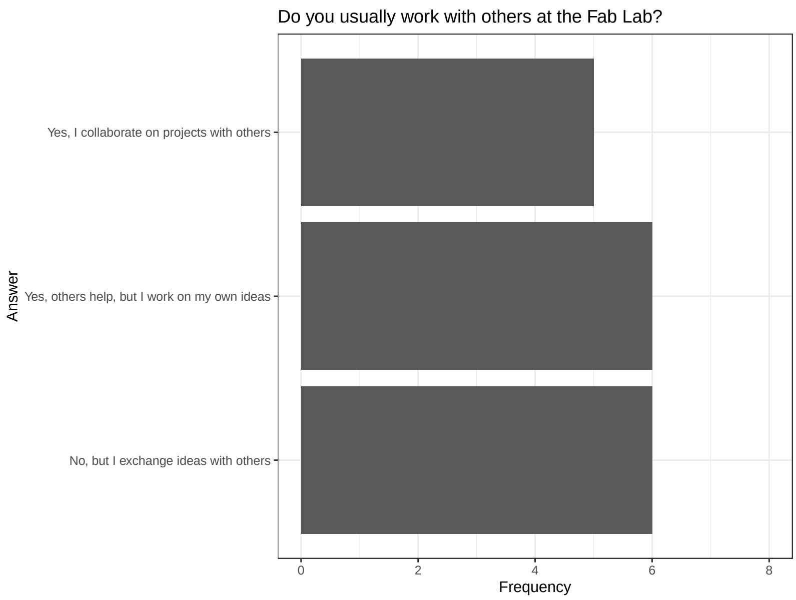


Figure C.49: User survey - question 7

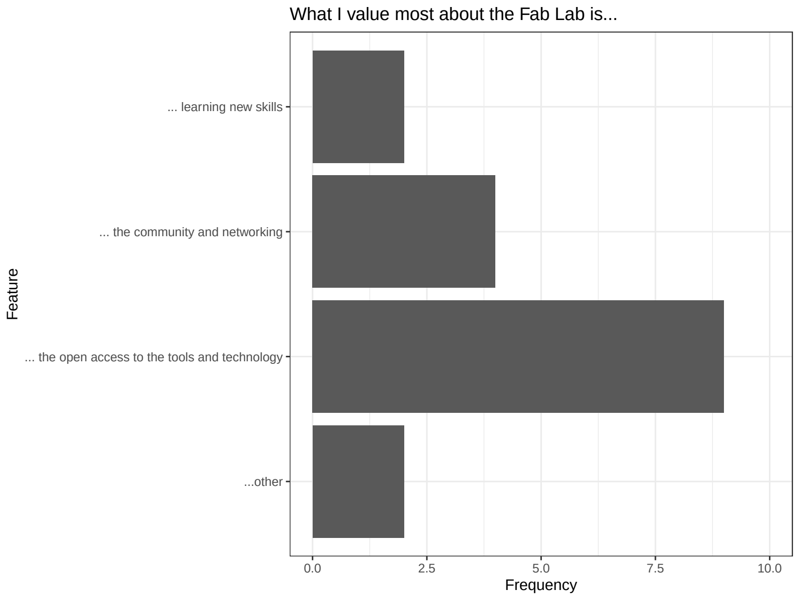


Figure C.50: User survey - question 8

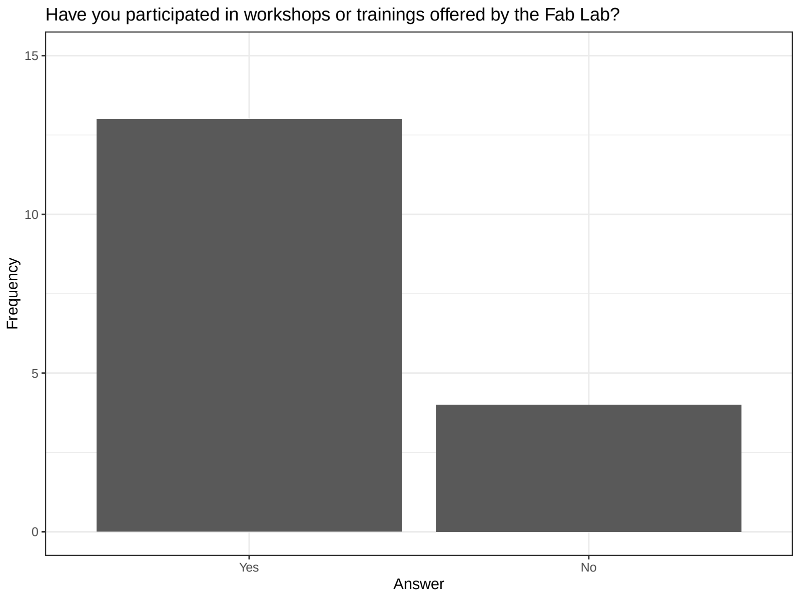


Figure C.51: User survey - question 9

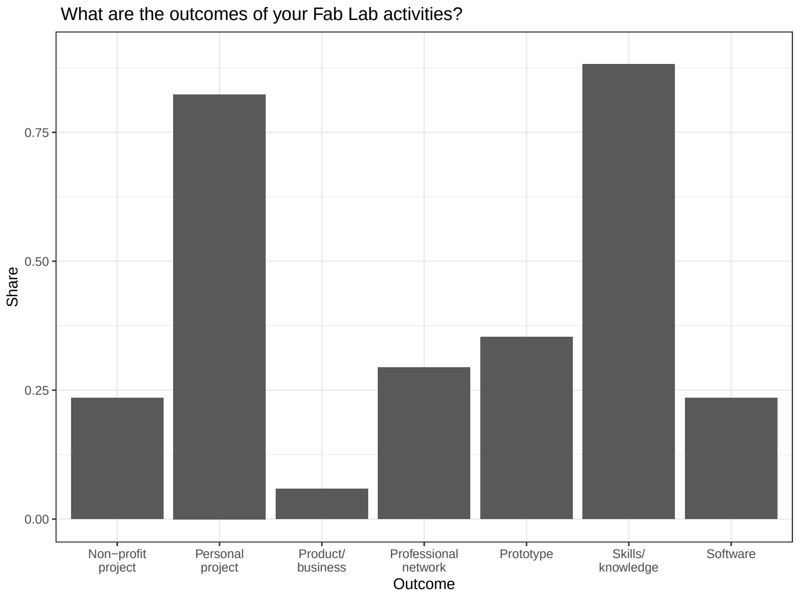


Figure C.52: User survey - question 10

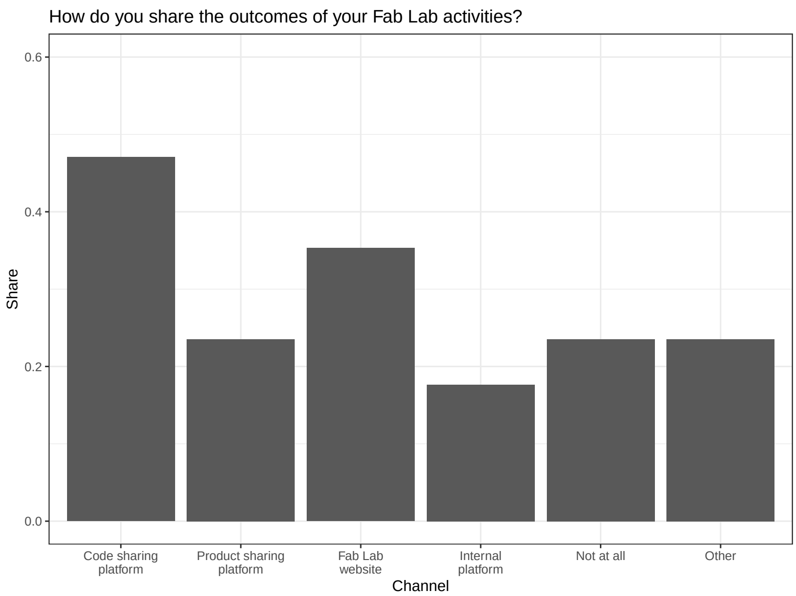


Figure C.53: User survey - question 11

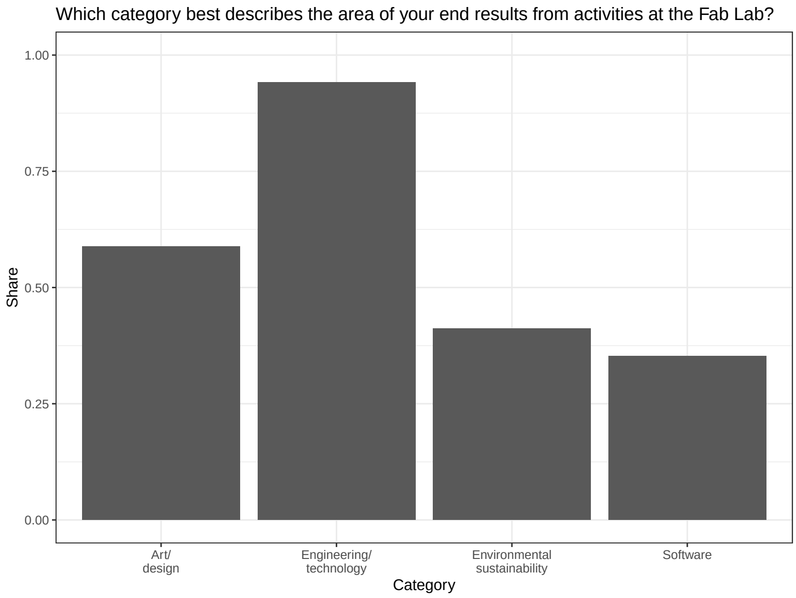


Figure C.54: User survey - question 12

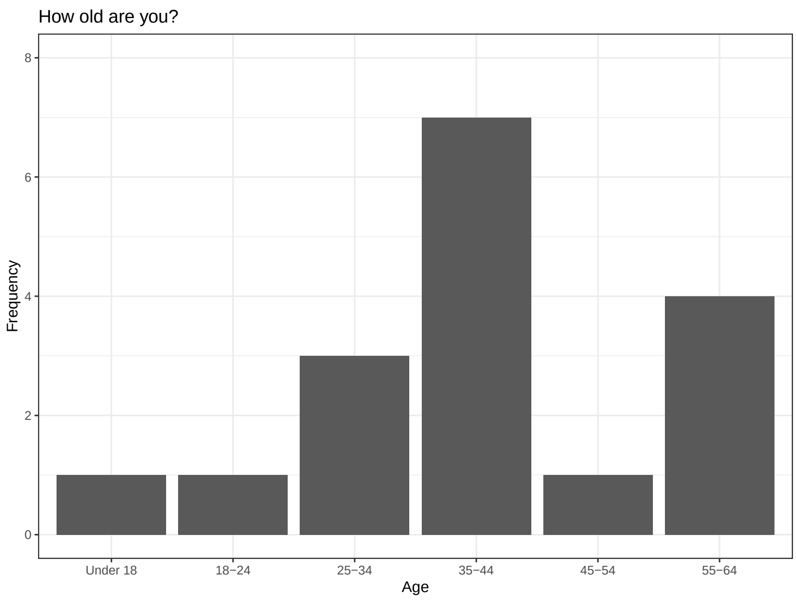


Figure C.55: User survey - question 13

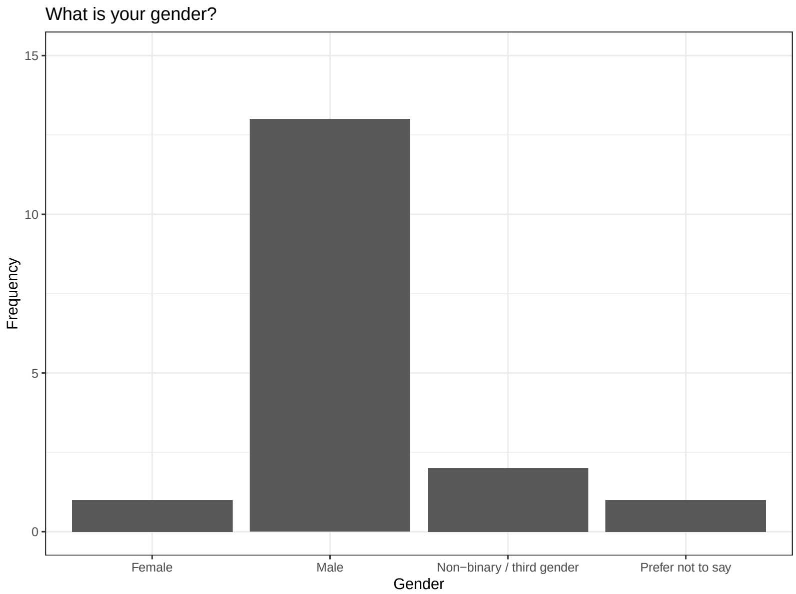


Figure C.56: User survey - question 14

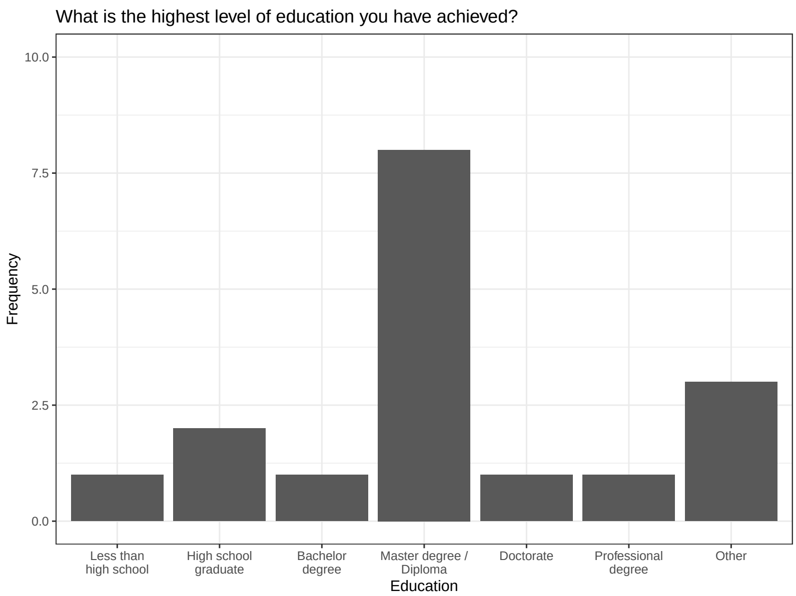


Figure C.57: User survey - question 15

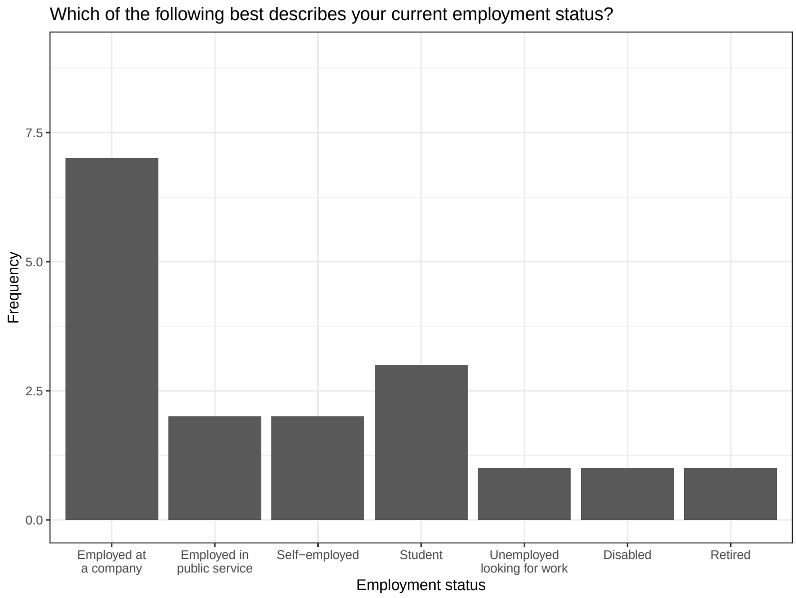


Figure C.58: User survey - question 16

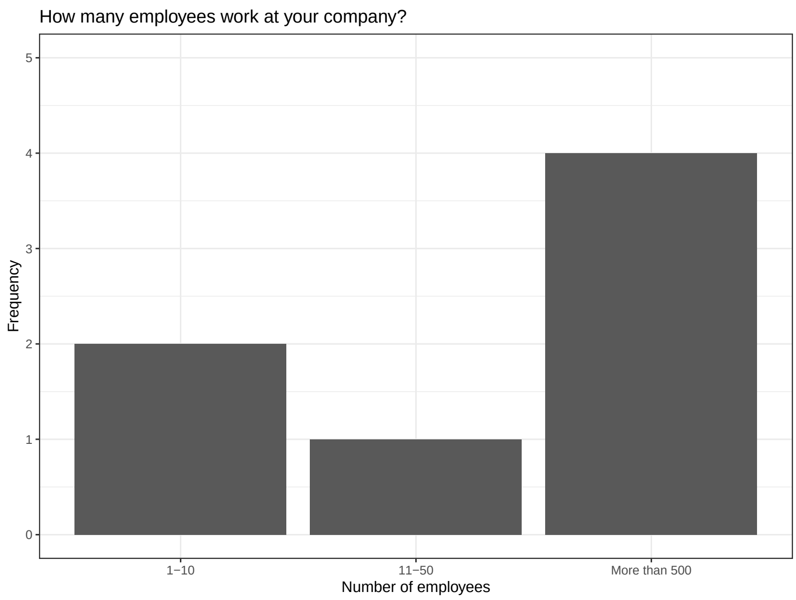


Figure C.59: User survey - question 17

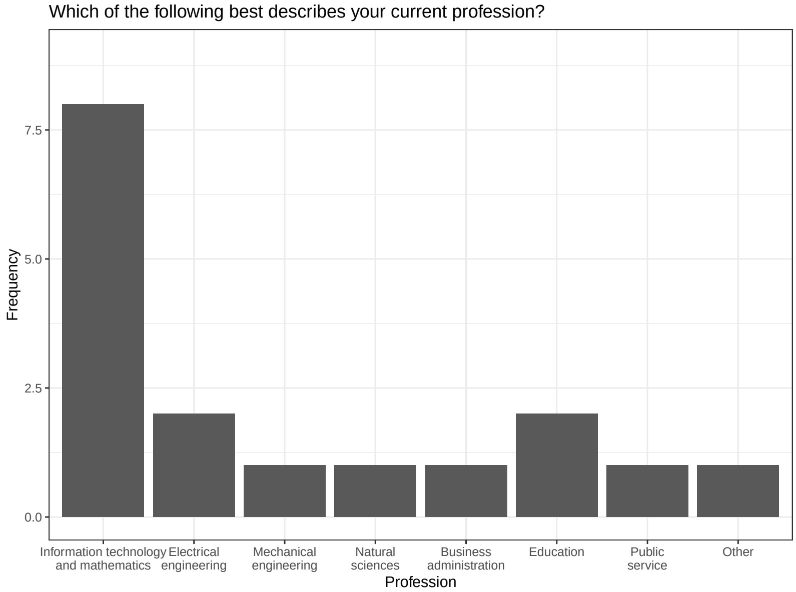


Figure C.60: User survey - question 18

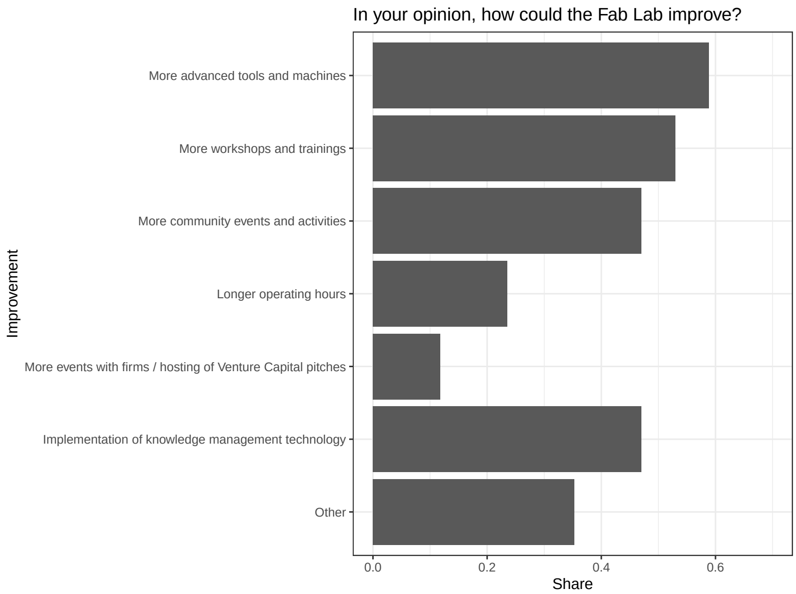


Figure C.61: User survey - question 19

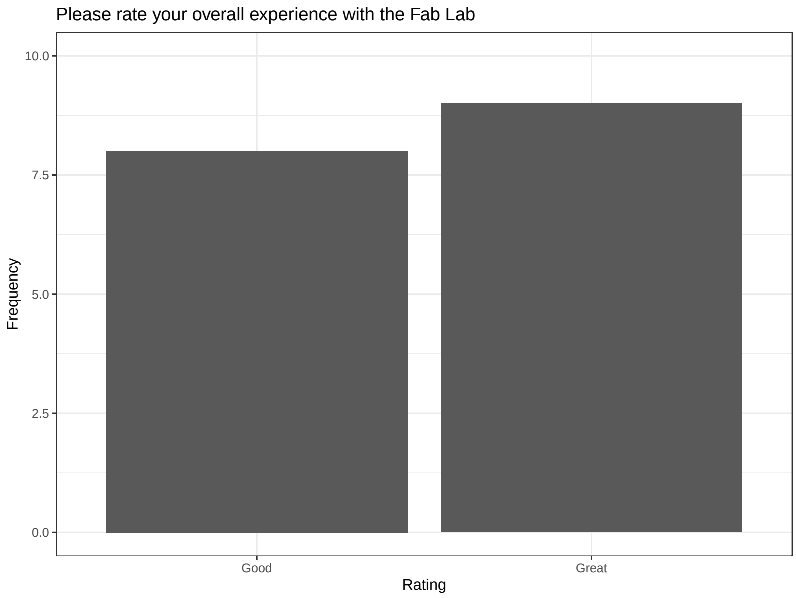


Figure C.62: User survey - question 20

C.4 Data

This section elaborates on the inventor disambiguation process. In November 2024, I received a dataset from Patrick Lehnert designed to disambiguate inventor names in the PATSTAT dataset. This dataset includes all German, Austrian, and Swiss inventors listed in PATSTAT, along with their respective addresses, personal identifiers, and regional identifiers. The identifiers include the PATSTAT person ID (*person-id*), which is derived from the inventor's name, address, and country code as associated with each patent application. However, variations in spelling, minor changes in names or addresses, and data inconsistencies can result in multiple *person-id* entries for the same individual. The regional identifies comprise the NUTS3 code, the AGS and the DEBKG-ID, which identify a municipality.

To address these issues, the OECD and EPO provide harmonized identifiers, including the PATSTAT standardized name ID (*psn-id*) and the harmonized applicant name ID (*han-id*). Building on these existing identifiers in PATSTAT, as well as the names and addresses of inventors, Patrick Lehnert developed two additional variables to improve identification. The first variable, *name-only-id*, uniquely identifies an individual based solely on their name and the provided IDs. The second variable, *name-city-id*, further distinguishes individuals within a *name-only-id* group by incorporating address information.

The original dataset comprises 3,584,821 observations, including 1,403,008 unique *name-only-id* entries and 2,007,109 unique *name-city-id* entries. Notably, 1,917,001 observations lack address information. Empty entries were removed, which accounted for 10 observations. The objective of the disambiguation process is to accurately identify individuals throughout the PATSTAT dataset, including their addresses. To utilize the dataset for inventor name disambiguation, the primary challenge was determining when to use the *name-only-id* and when to rely on the *name-city-id*. To achieve this, the following measures were designed to harmonize the dataset and fill missing address fields where appropriate.

First, I calculated the number of unique *name-city-id* entries within each *name-only-id* group and examined samples for different counts. This analysis informed the development of the following rules for selecting the appropriate identifier.

In general, if a *name-only-id* contained more than three *name-city-id* entries, the *name-city-id* was used for identification. Conversely, for *name-only-id* groups with three or fewer *name-city-id* entries, the *name-only-id* was deemed sufficient.

A first step involved additional refinement to correct inconsistencies in address information. During the dataset review, I identified 26,275 observations where English names of German cities were used, leading to unnecessary *name-city-id* entries within the corresponding *name-only-id* group. These additional entries often contained very few observations or even a single observation, despite the name and address being identical. The affected cities included *Cologne, Constance, Nuremberg, and Munich*. To resolve this, English city names were replaced with their German counterparts in the *name-city* string variable. This string variable, created by Patrick Lehnert, consists of an alphabetically ordered combination of the first name, surname, and city of the individual. Following Patrick Lehnert's methodology, the modified *name-city* strings were re-ordered alphabetically and grouped to generate a revised numerical identifier, *name-city2-id*. This step reduced the number of unique *name-city-id* entries from 2,007,109 to 1,990,305.

The next stage of disambiguation aimed to harmonize missing address fields while preserving the integrity of individual identifications. In cases where a *name-only-id* contained more than three distinct *name-city2-id* entries, it was frequently observed that one *name-city2-id* had an empty address field but non-missing NUTS3, AGS and DEKKG-ID variables. These values often matched those of another *name-city2-id* with a non-empty address field within the same *name-only-id*. Based on the assumption that individuals within a *name-only-id* (which already implies identical names) are identical if their regional identifiers match, the following measures were implemented. Within each *name-only-id*, I identified observations with missing address fields but non-missing NUTS3, AGS and DEBKG-ID variables. If these matched a *name-city2-id* within the same *name-only-id* that had a complete address and the same NUTS3, AGS, and DEBKG-ID values, the *name-city2-id* of the fully specified observation was reassigned to the one with missing address data. This operation was only applied to *name-only-id* groups with more than three unique *name-city2-id* entries. The resulting variable, *name-city3-id*, was created by overwriting empty rows in *name-city2-id* with the newly assigned values. In total, 102,523 observations were reassigned, reducing the number of unique *name-city3-id* to 1,958,473.

To further refine the dataset, an additional step reassigned remaining records with missing addresses to the most frequent *name-city3-id* within the corresponding *name-only-id*. This step was applied only to *name-only-id* groups with more than three unique *name-city3-id* entries and where the number of missing address entries did not exceed three. The variable *max-city3-id* was introduced to capture the most frequently occurring*name-city3-id* within each *name-only-id*. The new identifier, *name-city4-id*, resulted in a total of 1,938,454 unique IDs.

Finally, an additional verification step ensured that *name-only-id* groups with three distinct *name-city4-id* entries were not mistakenly assigned to multiple individuals. In many cases, the three n*name-city4-id* represented the same person using different address references – for example, a personal address versus an employer's address. To distinguish such cases from genuinely different individuals, geographic distance between NUTS3 regions was calculated. If the centroids of the NUTS3 regions exceeded 45 km, the corresponding *name-city4-id* were assumed to belong to distinct individuals. This threshold, based on the BBSR report on commuting distances in Germany, which identifies a maximum commuting distance of 50 km within Germany.[1]

For cases where only two distinct *name-city4-id* were associated with a *name-only-id*, additional criteria were applied. If one *name-city4-id* had a missing address, the two entries were assumed to belong to the same individual and the *name-only-id* was used identification. If both *name-city4-id* entries had non-missing addresses and were located within the same NUTS3 region, they were also treated as the same individual. In all other instances, I used the *name-city4-id* for identification.

[1]https:/ / www.bbsr.bund.de/ BBSR/ DE/ startseite/ topmeldungen/ pendeln-2021.html, accessed November 20, 2024.

D

Appendix to Chapter 4

Value Estimates for U.S. Patent Grants

D.1 Tables

D.1.1 Tables – Fixing the SNR

Table D.1: Winsorized regression analyses between capital gains and cumulative value correlates

	KPSS			FHK		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	22.31***	18.72***	40.66***	4.929***	1.460***	5.970***
	(0.156)	(0.156)	(0.0472)	(0.0215)	(0.0194)	(0.00858)
Adj. R^2	0.0799	0.241	0.966	0.345	0.560	0.957
Within R^2	0.0799	0.241	0.000759	0.345	0.560	0.461
N	855,618	855,618	833,641	855,614	855,614	833,637
#patents	YES	YES	YES	YES	YES	YES
Value correlates	NO	YES	YES	NO	YES	YES
FXT	NO	NO	YES	NO	NO	YES

Notes: The dependent variables are the original KPSS capital gains and our SNR-corrected FHK estimates. The dependent variables are winsorized at the 0.1% and 99.9% levels. Standard errors in parentheses. Patent characteristics are aggregated over simultaneously issued patents. The patent characteristics comprise the cumulative GDP covered by the international patent family, the number of U.S. patent literature references weighted by three-year citations, the number of claims in the grant document, the number of non-patent literature (NPL) references, the number of different technical areas cited by the focal patent, and the number of USPTO citations within 3 years and 10 years of filing. We also include the quadratic transformations of these patent characteristics. FXT: firm fixed effects x grant year. *** p<0.01, ** p<0.05, * p<0.

D.1.2 Tables – Hedonic Regressions

Table D.2: Log-linear specifications for singletons (*HED-SNG*)

log(variable)	(1)	(2)	(3)	(4)
GDP (patent family)	-0.021***	0.034***	0.029***	0.027***
	(0.004)	(0.005)	(0.005)	(0.002)
PL refs (weighted by 3yr-cites)	0.004***	0.007***	0.022***	0.006***
	(0.001)	(0.002)	(0.002)	(0.001)
claims (grant)	-0.020***	0.010***	0.008***	0.002
	(0.003)	(0.003)	(0.003)	(0.001)
NPL refs	-0.010***	-0.023***	-0.033***	0.005***
	(0.002)	(0.002)	(0.002)	(0.001)
areas cited	-0.064***	-0.047***	-0.064***	-0.001
	(0.004)	(0.004)	(0.004)	(0.002)
N	354,633	354,633	354,633	353,324
R^2	0.002	0.030	0.064	0.868
Year FE	NO	YES	YES	YES
Tech. class FE	NO	NO	YES	YES
Firm FE	NO	NO	NO	YES

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table D.3: Linear specifications for aggregates (*HED-AGG*)

Variable	(1)	(2)	(3)
#patents	1.287***	5.519***	3.111***
	(0.011)	(0.020)	(0.015)
GDP (patent family)	-0.074***	-0.501***	-0.249***
	(0.000)	(0.002)	(0.001)
(GDP (patent family))^2		0.008***	0.004***
		(0.000)	(0.000)
PL refs (weighted by 3yr-cites)	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
(PL refs)^2		-0.000***	-0.000***
		(0.000)	(0.000)
claims (grant)	0.074***	0.109***	0.060***
	(0.000)	(0.001)	(0.000)
(claims (grant))^2		-0.001***	-0.000***
		(0.000)	(0.000)
NPL refs	0.046***	0.028***	0.003***
	(0.000)	(0.000)	(0.000)
(NPL refs)^2		0.000***	0.000***
		(0.000)	(0.000)
areas cited	0.100***	0.282***	0.036***
	(0.002)	(0.004)	(0.003)
(areas cited)^2		-0.006***	0.003***
		(0.000)	(0.000)
N	678,743	678,743	677,438
R^2	0.530	0.583	0.822
Year FE	YES	YES	YES
Firm FE	NO	NO	YES

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table D.4: Specifications for aggregates using upper percentile coding (*HED-TOP*)

Variable	(1)	(2)	(3)
TOP5 GDP (patent family)	0.457***	-0.737***	-0.385***
	(0.018)	(0.022)	(0.015)
TOP5 PL refs (weighted by 3yr-cites)	3.753***	1.639***	1.324***
	(0.047)	(0.046)	(0.031)
TOP5 claims (grant)	3.933***	-1.202***	-0.758***
	(0.031)	(0.035)	(0.024)
TOP5 NPL refs	7.115***	1.956***	1.245***
	(0.042)	(0.044)	(0.029)
TOP5 areas cited	4.752***	2.277***	2.096***
	(0.043)	(0.040)	(0.027)
TOP25 GDP (patent family)		-1.273***	-0.366***
		(0.009)	(0.007)
TOP25 PL refs (weighted by 3yr-cites)		1.540***	1.117***
		(0.019)	(0.013)
TOP25 claims (grant)		2.695***	1.703***
		(0.012)	(0.009)
TOP25 NPL refs		1.377***	0.104***
		(0.014)	(0.010)
N	678,743	678,743	677,438
R^2	0.329	0.465	0.785
Year FE	NO	YES	YES
Firm FE	NO	NO	YES

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table D.5: Log-linear specifications using PatVal survey data (*HED-PATVAL*)

log(variable)	(1)	(2)	(3)
GDP (patent family)	1.019***	0.801***	0.943***
	(0.144)	(0.145)	(0.156))
PL refs (weighted by 3yr-cites)	0.087***	0.225***	0.260***
	(0.031)	(0.035)	(0.037)
claims (grant)	0.228***	0.239***	0.235***
	(0.064)	(0.064)	(0.064)
NPL refs	0.193***	0.084*	0.100**
	(0.040)	(0.043)	(0.044)
areas cited	0.403***	0.226**	0.255***
	(0.089)	(0.093)	(0.094)
N	3,874	3,874	3,874
R^2	0.062	0.107	0.118
Year FE	NO	YES	YES
Tech. class FE	NO	NO	YES

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

D.1.3 Tables – Citation Analysis

Table D.6: Citation analysis using hedonic value estimates (*HED-TOP* and *HED-SNG*)

	HED-TOP			HED-SNG		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(1+cites)	0.212***	0.185***	0.101***	0.112***	0.086***	0.001***
	(0.014)	(0.013)	(0.012)	(0.005)	(0.005)	(0.000)
N	2,703,446	2,703,446	2,685,320	2,703,446	2,703,446	2,685,320
R^2	0.154	0.233	0.814	0.177	0.274	0.957
Firm size	NO	YES	YES	NO	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
FE	CXT	CXT	CXT, FXT	CXT	CXT	CXT, FXT

Notes: The dependent variable is log(patent value estimate). See KPSS (2017), section III, Table II for their original estimates. Standard errors in parentheses, clustered by grant year. Firm size is approximated using the market value of the firm in the week preceding the grant event. CXT: tech classification x grant year. FXT: firm fixed effects x grant year. *** p<0.01, ** p<0.05, * p<0.1.

D.1.4 Tables – Applications

Table D.7: Distribution of Firm Size Categories

	Frequency	Percent	Cumulative
Patent Term Extension			
Patent term extension	364	0.01	0.01
No extension	2,703,241	99.99	100
Standard Essential Patents			
SEP	13,865	0.51	0.51
No SEP	2,689,740	99.49	100
Firm Size Category			
Missing	52,500	1.94	1.94
≤250 Employees	42,723	1.58	3.52
251–500 Employees	29,385	1.09	4.61
501–2,500 Employees	106,994	3.96	8.57
2,501–10,000 Employees	233,074	8.62	17.19
10,001–50,000 Employees	679,023	25.12	42.30
50,001–250,000 Employees	1,148,169	42.47	84.77
>250,000 Employees	411,737	15.23	100.00
N	2,703,605		

List of Figures

List of Tables

Bibliography

Abadie, A. (2005). Semiparametric difference-in-difference estimators. *Review of Economic Studies*, *72*, 1–19.

Abrams, D. S., Akcigit, U., Oz, G., & Pearce, J. G. (2019). *The patent troll: Benign middleman or stick-up artist?* (Tech. rep.). National Bureau of Economic Research.

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, *31*(7), 1069–1085.

Adams, A., Albin, S., Amunts, K., Asakawa, T., Bernard, A., Bjaalie, J. G., Chakli, K., Deshler, J. O., De Koninck, Y., Ebell, C. J., et al. (2020). International brain initiative: An innovative framework for coordinated global brain research efforts. *Neuron*, *105*(2), 212–216.

Agha, L. (2014). The effects of health information technology on the costs and quality of medical care. *Journal of health economics*, *34*, 19–30.

Agrawal, A., & Cockburn, I. (2002). University research, industrial r&d, and the anchor tenant hypothesis. *NBER Working Papers*, (9212).

Agrawal, A., Cockburn, I., & McHale, J. (2006). Gone but not forgotten: Knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, *6*(5), 571–591.

Akcigit, U., Grigsby, J., Nicholas, T., & Stantcheva, S. (2021). Taxation and innovation in the twentieth century. *The Quarterly Journal of Economics*, *137*(1), 329–385.

Alcacer, J., & Chung, W. (2007). Location strategies and knowledge spillovers. *Management Science*, *53*(5), 760–776.

Andrews, M. (2023). Bar talk: Informal social networks, alcohol prohibition, and invention. *SSRN Electronic Journal*.

Argente, D., Baslandze, S., Hanley, D., & Moreira, S. (2020). *Patents to products: Product innovation and firm dynamics* (Discussion Paper No. DP14692). CEPR.

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

Asaduzzaman, M., Mashiyat, A. S., Roy, C. K., & Schneider, K. A. (2013). Answering questions about unanswered questions of stack overflow. *2013 10th Working Conference on Mining Software Repositories (MSR)*, 97–100.

Asheim, B. T., & Gertler, M. S. (2009). The Geography of Innovation: Regional Innovation Systems. In *The Oxford Handbook of Innovation*. Oxford University Press.

Assenova, V. A. (2020). Early-stage venture incubation and mentoring promote learning, scaling, and profitability among disadvantaged entrepreneurs. *Organization Science*, *31*(6), 1560–1578.

Audretsch, D. B. (1998). Agglomeration and the location of innovative activity. *Oxford Review of Economic Policy*, *14*(2), 18–29.

Audretsch, D. B., & Belitski, M. (2020). The role of R&D and knowledge spillovers in innovation and productivity. *European Economic Review*, *123*.

Audretsch, D. B., Hülsbeck, M., & Lehmann, E., E. (2012). Regional competitiveness, university spillovers, and entrepreneurial activity. *Small Business Economics*, *39*, 587–601.

Azoulay, P., Graff Zivin, J. S., & Manso, G. (2011). Incentives and creativity: Evidence from the academic life sciences. *The RAND Journal of Economics*, *42*(3), 527–554.

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, *1*(1), 377–386.

Bailey, D. E., Faraj, S., Hinds, P. J., Leonardi, P. M., & von Krogh, G. (2022). We are all theorists of technology now: A relational perspective on emerging technology and organizing. *Organization Science*, *33*(1), 1–18.

Baldwin, C., & von Hippel, E. (2011). Modeling a paradigm shift: From producer innovation to user and open collaborative innovation. *Organization Science*, *22*(6), 1399–1417.

Bao, J. (2025). Do makerspaces affect entrepreneurship? If so, who, how, and when? *Strategic Management Journal*, *46*, 502–533.

Baruffaldi, S., & Poege, F. (2025). Like stars: How firms learn at scientific conferences. *Management Science*, *71*(3), 1865–1888.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & van Reenen, J. (2019). Who becomes an inventor in America? The importance of exposure to innovation. *The Quarterly Journal of Economics*, *134*(2), 647–713.

Berends, H., van der Bij, H., Debackere, K., & Weggeman, M. (2006). Knowledge sharing mechanisms in industrial research. *R&D Management*, *36*.

Berkes, E., & Nencka, P. (2020). Knowledge access: The effects of carnegie libraries on innovation. *SSRN Electronic Journal*.

Berman, B. (2012). 3D printing: The new industrial revolution. *Business Horizons*, *55*(2), 155–162.

Blikstein, P. (2013). Digital fabrication and "making" in education: The democratization of invention. In J. Walter-Herrmann & C. Büching (Eds.), *Fablabs: Of machines, makers, and inventors*. Transcript Publishers.

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104–1144.

Bogers, M., Zobel, A.-K., Afuah, A., Almirall, E., Brunswicker, S., Dahlander, L., Frederiksen, L., Gawer, A., Gruber, M., Haefliger, S., Hagedoorn, J., Hilgers, D., Laursen, K., Magnusson, M. G., Majchrzak, A., McCarthy, I. P., Moeslein, K. M., Nambisan, S., Piller, F. T., . . . Wal, A. L. J. (2017). The open innovation research landscape: Established perspectives and emerging themes across different levels of analysis. *Industry and Innovation*, *24*(1), 8–40.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, *91*(6), 3253–3285.

Boudreau, K. (2010). Open platform strategies and innovation: Granting access vs. devolving control. *Management Science*, *56*(10), 1849–1872.

Boutillier, S., Capdevila, I., Morel, L., & Dupont, L. (2020). Collaborative spaces promoting creativity and innovation. *Journal of Innovation Economics Management*, *31*(1), 1–9.

Brachtendorf, L., Gaessler, F., & Harhoff, D. (2023). Truly standard-essential patents? a semantics-based analysis. *Journal of Economics & Management Strategy*, *32*(1), 132–157.

Bray, M. J., & Lee, J. N. (2000). University revenues from technology transfer: Licensing fees vs. equity positions. *Journal of Business venturing*, *15*(5-6), 385–392.

Breschi, S. (2001). Knowledge spillovers and local innovation systems: A critical survey. *Industrial and Corporate Change*, *10*(4), 975–1005.

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography*, *9*(4), 439–468.

Breschi, S., Lissoni, F., & Montobbio, F. (2005). The geography of knowledge spillovers: Conceptual issues and measurement problems. In *Clusters, Networks and Innovation*. Oxford University Press.

Browder, R. E., Aldrich, H. E., & Bradley, S. W. (2019). The emergence of the maker movement: Implications for entrepreneurship research. *Journal of Business Venturing*, *34*(3), 459–476.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020). Language models are few-shot learners.

Brynjolfsson, E., & Hitt, L. M. (2003). Computing productivity: Firm-level evidence. *Review of economics and statistics*, *85*(4), 793–808.

Budish, E., Roin, B. N., & Williams, H. (2015). Do firms underinvest in long-term research? evidence from cancer clinical trials. *American Economic Review*, *105*(7), 2044–2085.

Bundesministerium für Digitales und Verkehr (BMDV). (2021). Regionalstatistische raumtypologie [Accessed: 2025-01-02]. https://bmdv.bund.de/SharedDocs/DE/Artikel/G/regionalstatistische-raumtypologie.html

Bundesministerium für Gesundheit. (2024). Chronik zum coronavirus sars-cov-2 | maßnahmen [Zugriff am 29. Dezember 2024]. https://www.bundesgesundheitsministerium.de/coronavirus/chronik-coronavirus.html

Burt, R. S. (2004). Structural holes and good ideas. *American journal of sociology*, *110*(2), 349–399.

Calefato, F., Lanubile, F., & Novielli, N. (2018). How to ask for technical help? evidence-based guidelines for writing questions on stack overflow. *Information and Software Technology*, *94*, 186–207.

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

Cantarero-Prieto, D., Leon, P. L., Blazquez-Fernandez, C., Juan, P. S., & Cobo, C. S. (2020). The economic cost of dementia: A systematic review. *Dementia*, *19*(8), 2637–2657.

Carayannis, E., & Campbel, D. F. (2009). ‘Mode 3’ and ‘Quadruple Helix’: Toward a 21st century fractal innovation ecosystem. *International Journal of Technology Management*, *46*(3/4).

Carroll, W. M. (2019). The global burden of neurological disorders. *The Lancet Neurology*, *18*(5), 418–419.

Catalini, C. (2018). Microgeography and the direction of inventive activity. *Management Science*, *64*(9), 4348–4364.

CB Insights. (2021). State of healthcare Q3’21 report [Cited: January 9, 2022. In: CB Insights, New York, [about 1 screen]].

Chae, Y., & Davidson, T. (2023). Large language models for text classification: From zero-shot learning to fine-tuning. *Open Science Foundation*.

Che, Y.-K., & Yoo, S.-W. (2001). Optimal incentives for teams. *American Economic Review*, *91*(3), 525–541.

Clancy, M. (2023). How many inventions are patented? [Accessed: 2025-01-04]. https://mattsclancy.substack.com/p/how-many-inventions-are-patented

Clavie, B., Ciceu, A., Naylor, F., Soulie, G., & Brightwell, T. (2023). Large language models in the workplace: A case study on prompt engineering for job type classification. https://arxiv.org/pdf/2303.07142.pdf

Cockburn, I., & Henderson, R. (1994). Racing to invest? the dynamics of competition in ethical drug discovery. *Journal of Economics & Management Strategy*, *3*(3), 481–519.

Cohen, W. M., Goto, A., Nagata, A., Nelson, R. R., & Walsh, J. P. (2002). R&d spillovers, patents and the incentives to innovate in japan and the united states. *Research policy*, *31*(8-9), 1349–1367.

Colombo, M. G., Grilli, L., & Piva, E. (2006). In search of complementary assets: The determinants of alliance formation of high-tech start-ups. *Research policy*, *35*(8), 1166–1199.

Cuntz, A., & Peuckert, J. (2023). From hackers to start-ups: Innovation commons and local entrepreneurial activity. *Research Policy*, *53*(104675).

Cutcher-Gershenfeld, J., Gershenfeld, A., & Gershenfeld, N. (2018). Digital fabrication and the future of work. *Perspectives on Work*, *22*.

Czarnitzki, D., Doherr, T., Hussinger, K., Schliessler, P., & Toole, A. A. (2017). Individual versus institutional ownership of universitydiscovered inventions. *USPTO Economic Working Paper Series*, (2017-07).

Dario, P., Carrozza, M. C., Guglielmelli, E., Laschi, C., Menciassi, A., Micera, S., & Vecchi, F. (2005). Robotics as a future and emerging technology: Biomimetics, cybernetics, and neuro-robotics in european projects. *IEEE Robotics & Automation Magazine*, *12*(2), 29–45.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Harvard Business School Press.

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–96.

Denicol, J., Davies, A., & Krystallis, I. (2020). What are the causes and cures of poor megaproject performance? a systematic literature review and research agenda. *Project Management Journal*, *51*(3), 328–345.

Deutsches Patent- und Markenamt. (2022). Gebrauchsmuster - Jahresbericht 2022 [Accessed: 2025-01-30]. https://www.dpma.de/digitaler_jahresbericht/2022/jb22_de/gebrauchsmuster.html

Deutsches Patent- und Markenamt. (2025a). Gebrauchsmuster [Accessed: 2025-01-02]. https://www.dpma.de/gebrauchsmuster/

Deutsches Patent- und Markenamt. (2025b). Merkblatt zur Gebrauchsmusteranmeldung [Accessed: 2025-01-02]. https://www.dpma.de/docs/formulare/gebrauchsmuster/g6181.pdf

Di Stefano, G., & Micheli, M. R. (2022). To stem the tide: Organizational climate and the locus of knowledge transfer. *Organization Science*.

Diez, T. (2012). Personal fabrication: FabLabs as platforms for citizen-based innovation, from microcontrollers to cities. *Nexus Network Journal*, *14*(3), 457–468.

Dijkstra, L., Poelman, H., & Veneri, P. (2019). The EU-OECD definition of a functional urban area. *OECD Regional Development Working Papers*, (11).

Dougherty, D. (2012). The Maker Movement. *Innovations: Technology, governance, globalization*, *7*(3), 11–14.

Dranove, D., Forman, C., Goldfarb, A., & Greenstein, S. (2014). The trillion dollar conundrum: Complementarities and health information technology. *American Economic Journal: Economic Policy*, *6*(4), 239–70.

European Patent Office. (2020). *Patents and additive manufacturing: Trends in 3D printing technologies* (tech. rep.). European Patent Office.

European Patent Office. (2023). *Innovation trends in additive manufacturing: Patents in 3D printing technologies* (tech. rep.). European Patent Office.

Eurostat. (2024a). Definition of NUTS [Accessed December 3, 2024]. https://ec.europa.eu/eurostat/documents/3859598/15193590/KS_GQ_22_010_EN_N.pdf

Eurostat. (2024b). NUTS3 shapefiles [Accessed December 5, 2024]. https://ec.europa.eu/eurostat/de/web/gisco/geodata/statistical_units/territorial_units_statistics

Fab Foundation. (2023). The FabFoundation [Accessed: 2024-02-05]. https://fabfoundation.org/

FabAcademy. (2024). FabAcademy [Accessed: 2024-04-15]. https://fabacademy.org/

FabLab Foundation. (2012). The FabCharter [Accessed: 2024-02-05]. http://fab.cba.mit.edu/about/charter/

Falotico, E., Vannucci, L., Ambrosano, A., Albanese, U., Ulbrich, S., Vasquez Tieck, J. C., Hinkel, G., Kaiser, J., Peric, I., Denninger, O., et al. (2017). Connecting artificial brains to robots in a comprehensive simulation framework: The neurorobotics platform. *Frontiers in neurorobotics*, *11*, 2.

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge collaboration in online communities. *Organization science*, *22*(5), 1224–1239.

Faraj, S., von Krogh, G., Monteiro, E., & Lakhani, K. R. (2016). Online community as space for knowledge flows. *Information Systems Research*, *27*(4), 668–684.

Federal Office for Building and Regional Planning (BBSR). (2024). File mapping districts to municipalities [Accessed February 2024]. https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/downloads/download_referenzen.html;jsessionid=C4CF328B69A342BED9161DFF484B0502.live11312

Federal Office for Cartography and Geodesy (BKG). (2024). NUTS3 shapefiles [Accessed December 12, 2024]. https://mis.bkg.bund.de/trefferanzeige?docuuid=D38F5B40_9209_4DC0_82BC_57FB575D29D7#detail_overview

Federal Statistical Office. (2024). Definition of nuts [Accessed December 3, 2024]. https://www.destatis.de/Europa/DE/Methoden_Metadaten/Klassifikationen/UebersichtKlassifikationen_NUTS.html

Fleischmann, K., Hielscher, S., & Merritt, T. (2016). Making things in FabLabs: A case study on sustainability and co-creation. *Digital Creativity*, *27*(2), 113–131.

Ford, D., Harkins, A., & Parnin, C. (2017). Someone like me: How does peer parity influence participation of women on stack overflow? *2017 IEEE symposium on visual languages and human-centric computing (VL/HCC)*, 239–243.

Franke, N., & Lüthje, C. (2020). User Innovation. In *Oxford Research Encyclopedia of Business and Management*. Oxford University Press.

Franke, N., & Shah, S. (2003). How communities support innovative activities: An exploration of assistance and sharing among end-users. *Research Policy*, *32*(1), 157–178.

Freedman, S., Lin, H., & Prince, J. (2018). Information technology and patient health: Analyzing outcomes, populations, and mechanisms. *American Journal of Health Economics*, *4*(1), 51–79.

Frégnac, Y. (2017). Big data and the industrialization of neuroscience: A safe roadmap for understanding the brain? *Science*, *358*(6362), 470–477.

Funke, M., Schularick, M., & Trebesch, C. (2023). Populist leaders and the economy. *American Economic Review*, *113*(12), 3249–3288.

Furman, J. L., Kyle, M. K., Cockburn, I., & Henderson, R. M. (2006). Public & private spillovers, location and the productivity of pharmaceutical research. *NBER WORKING PAPER SERIES*, (12509). https://www.nber.org/system/files/working_papers/w12509/w12509.pdf

Furman, J. L., & MacGarvie, M. J. (2007). Academic Science and the Birth of Industrial Research Laboratories in the US Pharmaceutical Industry. *Journal of Economic Behavior & Organization*, *63*(4), 756–776.

Furman, J. L., Nagler, M., & Watzinger, M. (2021). Disclosure and subsequent innovation: Evidence from the patent depository library program. *American Economic Journal: Economic Policy*, *13*(4), 239–270.

Furman, J. L., & Stern, S. (2011). Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *American Economic Review*, *101*(5), 1933–63.

Gallus, J., Jung, O., & Lakhani, K. (2019). Managerial recognition as an incentive for innovation platform engagement: A field experiment and interview study at nasa. *Harvard Business School working paper series# 20-059*.

Gallus, J., Jung, O. S., & Lakhani, K. R. (2020). Recognition incentives for internal crowdsourcing: A field experiment at nasa. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (20-059).

Gambardella, A., Raasch, C., & von Hippel, E. (2017). The user innovation paradigm: Impacts on makrets and welfare. *Management Science*, *63*(5), 1450–1468.

Garcia-Ruiz, M.-E., Lena-Acebo, F.-J., & Rocha Blanco, R. (2023). Early stages of the Fablab movement: A new path for an open innovation model. *Sustainability*, *15*(8907).

Gebhardt, A., & Hötter, J.-S. (2016). *Additive manufacturing: 3D printing for prototyping and manufacturing*. Carl Hanser Verlag GmbH & Co. KG.

Gershenfeld, N. (2007). Neil Gershenfeld: The beckoning promise of personal fabrication [Accessed: 2024-10-14].

Gershenfeld, N. (2012). How to make almost anysthing: The digital fabrication revolution. *Foreign Affairs*, *91*(6), 43–57.

Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or The undefinable tacitness of being (there). *Journal of Economic Geography*, *3*(1), 75–99.

Giorcelli, M. (2019). The long-term effects of management and technology transfers. *American Economic Review*, *109*(1), 121–152.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., Garcia-Fontes, W., Geuna, A., Gonzales, R., Harhoff, D., et al. (2007). Inventors and invention processes in Europe: Results from the PatVal-EU survey. *Research policy*, *36*(8), 1107–1127.

Goldfarb, A., Taska, B., & Teodoridis, F. (2020). Artificial intelligence in health care? evidence from online job postings. *AEA Papers and Proceedings*, *110*, 400–404.

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

Grabher, G., & Ibert, O. (2014). Distance as asset? knowledge collaboration in hybrid virtual communities. *Journal of Economic geography*, *14*(1), 97–123.

Graham, S. J., & Hedge, D. (2015). Disclosing patents' secrets. *Science*, *347*(6219), 236–237.

Granstrand, O., & Holgersson, M. (2020). Innovation ecosystems: A conceptual review and a new definition. *Technovation*, *90-91*(102098).

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

Grillner, S., Ip, N., Koch, C., Koroshetz, W., Okano, H., Polachek, M., Poo, M.-m., & Sejnowski, T. J. (2016). Worldwide initiatives to advance brain research. *Nature neuroscience*, *19*(9), 1118–1122.

Gross, D. P. (2019). Consequences of invention secrecy: Evidence from the USPTO Patent Secrecy Program in World War II. *National Bureau of Economic Research Working Paper*.

Haas, M. R., Criscuolo, P., & George, G. (2015). Which problems to solve? online knowledge sharing and attention allocation in organizations. *Academy of Management Journal*, *58*(3), 680–711.

Hagedoorn, J. (2002). Inter-firm r&d partnerships: An overview of major trends and patterns since 1960. *Research policy*, *31*(4), 477–492.

Halbinger, M. A. (2018). The role of makerspaces in supporting consumer innovation and diffusion: An empirical analysis. *Research Policy*, *47*, 2028–2036.

Hall, B. H. (2004). Exploring the patent explosion. *The Journal of Technology Transfer*, *30*, 35–48.

Hall, B. H., & Helmers, C. (2024). *The economics of innovation and intellectual property*. Oxford University Press.

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

Hall, B. H., & Ziedonis, R. H. (2001). The patent paradox revisited: An empirical study of patenting in the u.s. semiconductor industry, 1979-1995. *The RAND Journal of Economics*, *32*(1), 101–128.

Hallen, B. L., Cohen, S. L., & Bingham, C. B. (2020). Do accelerators work? if so, how? *Organization Science*, *31*(2), 378–414.

Harhoff, D., Henkel, J., & von Hippel, E. (2003). Profiting from voluntary information spillovers: How users benefit by freely revealing their innovations. *Research Policy*, *32*(10), 1753–1769.

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

Harhoff, D., Scherer, F. M., & Vopel, K. (1998). Exploring the tail of patent value distributions.

Heckman, J. J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, *66*(5), 1017–1098.

Heckman, J. J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, *64*(4), 605–654.

Heikkilä, J., & Verba, M. (2018). The role of utility models in patent filing strategies: Evidence from european countries. *Scientometrics*, *116*, 689–719.

Henkel, J., & von Hippel, E. (2004). Welfare implications of user innovation. *The Journal of Technology Transfer*, *30*, 73–87.

Higham, K., de Rassenfosse, G., & Jaffe, A. B. (2021). Patent quality: Towards a systematic framework for analysis and measurement. *Research Policy*, *50*(4), 104215.

Hochschulrektorenkonferenz. (2024). Maßnahmen und Sonderregelungen der Bundesländer [Accessed: 2024-12-29]. https://www.hrk.de/themen/hochschulsystem/covid-19-pandemie-und-die-hochschulen/massnahmen-und-sonderregelungen-der-bundeslaender/

Holmstrom, B. (1989). Agency costs and innovation. *Journal of Economic Behavior & Organization*, *12*(3), 305–327.

Hsu, D. H., Hsu, P.-H., Zhou, T., & Ziedonis, A. A. (2021). Benchmarking u.s. university patent value and commercialization efforts: A new approach. *Research Policy*, *50*(1), 104076.

Huang, P., Ceccagnoli, M., Forman, C., & Wu, D. (2022). IT knowledge spillovers, absorptive capacity, and productivity: Evidence from enterprise software. *Information Systems Research*, *33*(3), 908–934.

Hwang, E. H., Singh, P. V., & Argote, L. (2015). Knowledge sharing in online communities: Learning to cross geographic and hierarchical boundaries. *Organization Science*, *26*(6), 1593–1611.

Jaffe, A. B. (1989). Real effects of academic research. *American Economic Review*, *79*(5), 957–970.

Jeppesen, L. B. (2021). Reprint of: Social movements and free innovation. *Research Policy*, *50*(104288).

Jones, B. F. (2009). The burden of knowledge and the “death of the renaissance man”: Is innovation getting harder? *The Review of Economic Studies*, *76*(1), 283–317.

Jones, B. F. (2021). The rise of research teams: Benefits and costs in economics. *Journal of Economic Perspectives*, *35*(2), 191–216.

Jong, J. P., von Hippel, E., Gault, F., Kuusisto, J., & Raasch, C. (2015). Market failure in the diffusion of consumer-developed innovations: Patterns in Finland. *Research Policy*, *44*(10), 1856–1865.

Jordà, Ò., Schularick, M., & Taylor, A. M. (2017). Macrofinancial history and the new business cycle facts. *NBER macroeconomics annual*, *31*(1), 213–263.

Jungblut, M. (2020). Rückblick 2020 – Chronologie eines Schuljahrs in der Coronakrise [Accessed: 2024-12-29]. https://www.deutschlandfunk.de/rueckblick-2020-chronologie-eines-schuljahrs-in-der-100.html

Karagianis, L. (2006). Fab Lab.

Kline, P., Petkova, N., Williams, H., & Zidar, O. (2019). Who Profits from Patents? Rent-Sharing at Innovative Firms. *The Quarterly Journal of Economics*, *134*(3), 1343–1404.

Klöckner, J., & Specht, F. (2023). Fluch des Erfolgs: Deutschland hat Ruf als Erfinderland verloren. *Handelsblatt*. https://www.handelsblatt.com/politik/deutschland/studie-fluch-des-erfolgs-deutschland-hat-ruf-als-erfinderland-verloren/29165830.html

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth*. *The Quarterly Journal of Economics*, *132*(2), 665–712.

Koppel, O. (2023). Der Erfindergeist in Deutschland lahmt. *Institut der deutschen Wirtschaft (IW)*.

Korinek, A. (2023). Generative AI for economic research: Use cases and implications for economists. *Journal of Economic Literature*, *61*(4), 1281–1317.

Kreyer, A., & Wang, L. X. (2022). Collaborating neuroscience online: The case of the human brain project forum. *PLOS ONE*, *17*(12), e0278402.

Lanjouw, J. O. (1998). Patent protection in the shadow of infringement: Simulation estimations of patent value. *The Review of Economic Studies*, *65*(4), 671–710.

Lanzolla, G., Lorenz, A., Miron-Spektor, E., Schilling, M., Solinas, G., & Tucci, C. (2020). Digital transformation: What is new of anything? Emerging patterns and management research. *Academy of Management Discoveries*, *6*(3), 341–650.

Lee, S. Y., Florida, R., & Acs, Z. (2004). Creativity and entrepreneurship: A regional analysis of new firm formation. *Regional studies*, *38*(8), 879–891.

Lehnert, P., Pfister, C., Harhoff, D., & Backes-Gellner, U. (2022). Innovation effects and knowledge complementarities in a diverse research landscape. *Swiss Leading House VPET-ECON. Leading House "Economics of Education" Working Paper*, (164).

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., & Griliches, Z. (1987). Appropriating the returns from industrial research and development. *Brookings papers on economic activity*, *1987*(3), 783–831.

Lorents, A., Colin, M.-E., Bjerke, I. E., Nougaret, S., Montelisciani, L., Diaz, M., Verschure, P., & Vezoli, J. (2023). Human brain project partnering projects meeting: Status quo and outlook. *eNeuro*, *10*(9).

Lou, B., & Wu, L. (2021). AI on drugs: Can artificial intelligence accelerate drug development? evidence from a large-scale examination of bio-pharma firms. *MIS Quarterly*, *45*(3).

Loukas, L., Stogiannidis, I., Malakasiotis, P., & Vassos, S. (2023). Breaking the bank with ChatGPT: Few-shot text classification for finance. https://arxiv.org/abs/2308.14634

Lu, S. F., Rui, H., & Seidmann, A. (2018). Does technology substitute for nurses? staffing decisions in nursing homes. *Management Science*, *64*(4), 1842–1859.

Ludwig, T., Stickel, O., Boden, A., & Pipek, V. (2014). Towards sociable technologies: An empirical study on designing appropriation infrastructures for 3D printing. *Proceedings of the 2014 conference on Designing interactive systems*, 835–844.

Maker Media GmbH. (n.d.). Das Make-Magazin [Accessed: 2025-01-08].

Mansfield, E. (1968). *Industrial research and technological innovation: An econometric analysis*. Norton.

Maravilhas, S., & Martins, J. (2019). Strategic knowledge management in a digital environment: Tacit and explicit knowledge in Fab Labs. *Journal of Business Research*, *94*, 353–359.

Markram, H., Meier, K., Lippert, T., Grillner, S., Frackowiak, R., Dehaene, S., Knoll, A., Sompolinsky, H., Verstreken, K., DeFelipe, J., Grant, S., Changeux, J.-P., & Saria, A. (2011). Introducing the human brain project [Proceedings of the 2nd European Future Technologies Conference and Exhibition 2011 (FET 11)]. *Procedia Computer Science*, *7*, 39–42.

Mauroner, O. (2017). Makers, hackers, diy-innovation, and the strive for entrepreneurial opportunities. *International Journal of Entrepreneurship and Small Business*, *31*(1), 32–46.

May, A., Wachs, J., & Hannák, A. (2019). Gender differences in participation and reward on stack overflow. *Empirical Software Engineering*, *24*(4), 1997–2019.

McCullough, J. S., Casey, M., Moscovice, I., & Prasad, S. (2010). The effect of health information technology on quality in us hospitals. *Health affairs*, *29*(4), 647–654.

McCullough, J. S., Parente, S. T., & Town, R. (2016). Health information technology and patient outcomes: The role of information and labor coordination. *The RAND Journal of Economics*, *47*(1), 207–236.

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., et al. (2016). How open science helps researchers succeed. *elife*, *5*.

Melero, E., & Palomeras, N. (2015). The renaissance man is not dead! the role of generalists in teams of inventors. *Research Policy*, *44*(1), 154–167.

Mersand, S. (2021). The state of Makerspace research: A review of the literature. *TechTrends*, *65*, 174–186.

Metitieri, T., & Mele, S. (2020). Women in neuroscience: A short time travel. *Reference Module in Neuroscience and Biobehavioral Psychology*.

Mikhak, B., Lyon, C., Gorton, T., Gershenfeld, N. A., McEnnis, C., & Taylor, J. (2002). Fab lab: An alternate model of ict for development. *2nd international conference on open collaborative design for sustainable innovation*.

Miller, A. R., & Tucker, C. E. (2011). Can health care information technology save babies? *Journal of Political Economy*, *119*(2), 289–324.

Moilanen, J. (2012). Emerging Hackerspaces – peer-production generation. In I. Hammouda, B. Lundell, T. Mikkonen, & W. Scacchi (Eds.), *Open source systems: Long-term sustainability* (pp. 94–111). Springer Berlin Heidelberg.

Mokyr, J. (2002). *The gifts of athena: Historical origins of the knowledge economy*. Princeton University Press.

Mondal, S., Saifullah, C. K., Bhattacharjee, A., Rahman, M. M., & Roy, C. K. (2021). Early detection and guidelines to improve unanswered questions on stack overflow. *14th Innovations in software engineering conference (formerly known as India software engineering conference)*, 1–11.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of web of science and scopus: A comparative analysis. *Scientometrics*, *106*, 213–229.

Moreno, R., Paci, R., & Usai, S. (2005). Geographical and sectoral clusters of innovation in europe. *The annals of regional science*, *39*, 715–739.

Moritz, M., Redlich, T., Buxbaum-Conradi, S., & Wulfsberg, J. P. (Eds.). (2024). *Global collaboration, local production: Fab City als Modell für Kreislaufwirtschaft und nachhaltige Entwicklung*. Springer Gabler.

Morrison, M. (2017). “a good collaboration is based on unique contributions from each side”: Assessing the dynamics of collaboration in stem cell science. *Life Sciences, Society and Policy*, *13*(1), 1–20.

Mortara, L., & Parisot, N. (2018). How do fab-spaces enable entrepreneurship? Case studies of ’makers’ - entrepreneurs [Number Of Volumes: 89465]. *International Journal of Manufacturing Technology and Management*, *32*(1), 16.

Moser, P. (2005). How do patent laws influence innovation? evidence from nineteenth-century world’s fairs. *American economic review*, *95*(4), 1214–1236.

Murray, F., Aghion, P., Dewatripont, M., Kolev, J., & Stern, S. (2016). Of mice and academics: Examining the effect of openness on innovation. *American Economic Journal: Economic Policy*, *8*(1), 212–52.

Nager, A., & Atkinson, R. D. (2016). A trillion-dollar opportunity: How brain research can drive health and prosperity. *Information Technology and Innovation Foundation*, *11*.

Nelson, R. R. (1959). The simple economics of basic scientific research. *Journal of political economy*, *67*(3), 297–306.

Nelson, R. R., & Winter, S. G. (1985). *An evolutionary theory of economic change*. harvard university press.

Nonaka, I., Takeuchi, H., & Umemoto, K. (1996). A theory of organizational knowledge creation. *International journal of technology Management*, *11*(7-8), 833–845.

Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment & Evaluation in Higher Education*, *33*(3), 301–314.

OECD. (2024). OECD FUAs [Accessed February 2024]. https://www.oecd.org/regional/regional_statistics/functional_urban_areas.htm

Oppedisano, R. (2024). *Collaborative spaces and fablabs - a global organisational and collaborative network*. Springer Nature Switzerland AG.

Pakes, A. (1985). On patents, r & d, and the stock market rate of return. *Journal of Political Economy*, *93*(2), 390–409.

Pakes, A., & Schankerman, M. (1984). The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources. In *R&d, patents, and productivity* (pp. 73–88). National Bureau of Economic Research, Inc.

Peschl, M. F., & Fundneider, T. (2012). Spaces enabling game-changing and sustaining innovations: Why space matters for knowledge creation and innovation. *Journal of Organisational Transformation & Social Change*, *9*(1), 41–61.

Pfister, C., Koomen, M., Harhoff, D., & Backes-Gellner, U. (2021). Regional innovation effects of applied research institutions. *Research Policy*, *50*(4), 104197.

Pham, C. M., Hoyle, A., Sun, S., & Iyyer, M. (2023). TopicGPT: A prompt-based topic modeling framework.

Popadiuk, S., & Choo, C. W. (2006). Innovation and knowledge creation: How are these concepts related? *International Journal of Information Management*, *26*(4), 302–312.

Preißner, S., Raasch, C., & Schweisfurth, T. (2024). When necessity is the mother of disruption: Users versus producers as sources of disruptive innovation. *Journal of Product Innovation Management*, *41*(1), 62–85.

Randhawa, K., Josserand, E., Schweitzer, J., & Logue, D. (2017). Knowledge collaboration between organizations and online communities: The role of open innovation intermediaries. *Journal of Knowledge Management*.

Rayna, T., & Striukova, L. (2021). Fostering skills for the 21st century: The role of Fab labs and makerspaces. *Technological Forecasting & Social Change*, *164*(120391).

Reagans, R., Zuckerman, E., & McEvily, B. (2004). How to make the team: Social networks vs. demography as criteria for designing effective teams. *Administrative science quarterly*, *49*(1), 101–133.

Roche, M. P., Oettl, A., & Catalini, C. (2024). Proximate (Co-)Working: Knowledge Spillovers and Social Interactions. *Management Science*, *70*(12), 8217–9119.

Rose, M., & Baruffaldi, S. (2020). Finding doppelgängers in scopus: How to build scientists control groups using sosia. *SSRN Electronic Journal*.

Rose, M. E., & Kitchin, J. R. (2019). Pybliometrics: Scriptable bibliometrics using a python interface to scopus. *SoftwareX*, *10*, 100263.

Rose, N. (2014). The human brain project: Social and ethical challenges. *Neuron*, *82*(6), 1212–1215.

Royston, P., & Parmar, M. K. (2002). Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in medicine*, *21*(15), 2175–2197.

Sandvik, J. J., Saouma, R. E., Seegert, N. T., & Stanton, C. T. (2020). Workplace knowledge flows. *The Quarterly Journal of Economics*, *135*(3), 1635–1680.

Savastano, M., Bellini, F., D'Ascenzo, F., & Scornavacca, E. (2017). *FabLabs as platforms for digital fabrication services: A literature analysis: 8th International Conference, IESS 2017, Rome, Italy, May 24-26, 2017, Proceedings* (M. Cavallari, M. Drăgoicea, & S. Za, Eds.; 1st ed. 2017, Vol. 279). Springer International Publishing; Imprint: Springer.

Scherer, F. M. (1965). Firm size, market structure, opportunity, and the output of patented inventions. *The American economic review*, *55*(5), 1097–1125.

Scherer, F. M., & Harhoff, D. (2000). Technology policy for a world of skew-distributed outcomes. *Research policy*, *29*(4-5), 559–566.

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

Schrape, J.-F. (2020). Kollaborative Labs und offene Werkstätten. *ökologisches Wirtschaften - Fachzeitschrift*, *33*(1), 22.

Schwierzy, J. (2025). *The economics of additive manufacturing: From technology adoption to market disruption* [Doctoral dissertation, Technical University Munich].

Shah, S. K., & Tripsas, M. (2007). The accidental entrepreneur: The emergent and collective process of user entrepreneurship. *Strategic entrepreneurship journal*, *1*(1-2), 123–140.

Singh, J., & Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality? *Management Science*, *56*(1), 41–56.

Siva, N. (2023). What happened to the human brain project? *Lancet (London, England)*, *402*(10411), 1408–1409.

Smith, M. R., & Marx, L. (Eds.). (1994). *Does technology drive history?: The dilemma of technological determinism*. MIT Press.

Soete, L., & Freeman, C. (1997). *The economics of industrial innovation* (1st). Routledge.

Stacey, M. (2014). The FAB LAB Network: A Global Platform for Digital Invention, Education and Entrepreneurship. *Innovations: Technology, Governance, Globalization*, *9*(1-2), 221–238.

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

Svensson, P., & Hartmann, R. K. (2018). Policies to promote user innovation: The case of makerspaces and clinician innovation. *Research Policy*, *47*, 277–288.

Technische Universität Ilmenau. (2024). Ilmkubator [Accessed: 2025-01-08].

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research policy*, *15*(6), 285–305.

Toivanen, O., & Väänänen, L. (2016). Education and invention. *Review of Economics and Statistics*, *98*(2), 382–396.

Torrisi, S., Gambardella, A., Giuri, P., Harhoff, D., Hoisl, K., & Mariani, M. (2016). Used, blocking and sleeping patents: Empirical evidence from a large-scale inventor survey. *Research Policy*, *45*(7), 1374–1385.

Tortoriello, M., McEvily, B., & Krackhardt, D. (2015). Being a catalyst of innovation: The role of knowledge diversity and network closure. *Organization Science*, *26*(2), 423–438.

Troxler, P., Kocher, P.-Y., Harboe, J., & Gaudenz, U. (2020). Sharing is sparing: Open knowledge sharing in FabLabs. *Journal of Peer Production*, (5).

Troxler, P., & Schweikert, S. (2016). Developing a business model for concurrent enterprising at the Fab Lab. In *Institute of Electrical and Electronics Engineers 2016 – 2010 IEEE International Technology Management* (pp. 1–8). IEEE.

Troxler, P., & Wolf, P. (2010). Bending the rules: The FabLab innovation ecology. *11th International CINet Conference Zurich, 5-7 September*.

van Holm, E. J. (2014). What are Makerspaces, Hackerspaces, and Fab Labs? *SSRN Electronic Journal*.

van Holm, E. J. (2017). Makerspaces and local economic development. *Economic Development Quarterly*, *31*(2), 164–173.

Viswanathan, V., Gashteovski, K., Lawrence, C., Wu, T., & Neubig, G. (2023). Large language models enable few-shot clustering. http://arxiv.org/pdf/2307.00524.pdf

von Hippel, E. (2005). *Democratizing Innovation*. The MIT Press.

von Hippel, E., & Krogh, G. (2003). Open source software and the "Private-Collective" innovation model: Issues for organization science. *Organization Science*, *14*(2), 209–223.

von Krogh, G. (1998). Care in knowledge creation. *California Management Review*, *40*(3), 133–153.

Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PloS one*, *10*(5), e0127298.

Wang, L. X. (2021). The complementarity of drug monitoring programs and health it for reducing opioid-related mortality and morbidity. *Health Economics*, *30*(9), 2026–2046.

Wang, L. X. (2022). Global drug diffusion and innovation with the medicines patent pool. *Journal of Health Economics*, *85*, 102671.

Wang, L. X., & Bloch, C. (2023). Digital interventions in the health sector: Country cases and policy discussions. *IMF Notes*, *2023*(004).

Wang, W. (2023). *Valuation of patents using stock market responses* [Doctoral dissertation, University of California, Berkeley].

Watson, C. (2013). The brain race: Can giant computers map the mind? [Cited: January 9, 2022. In: The Conversation].

Weller, C., Kleer, R., & Piller, F. T. (2015). Economic implications of 3D printing: Market structure models in light of additive manufacturing revisited. *International Journal of Production Economics*, *164*, 43–56.

West, J., & Kuk, G. (2016). The complementarity of openness: How MakerBot leveraged Thingiverse in 3D printing. *Technological Forecasting and Social Change*, *102*, 169–181.

Williams, H. L. (2013). Intellectual property rights and innovation: Evidence from the human genome. *Journal of Political Economy*, *121*(1), 1–27.

Winickoff, D., Kreiling, L., Borowiecki, M., Gardeni, H., & Philpi, J. (2021). *Collaborative platforms for emerging technology—creating convergence spaces* (OECD Science, Technology and Industry Policy Papers No. 109). OECD.

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *SSRN Electronic Journal*.

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*(7744), 378–382.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science (New York, N.Y.)*, *316*(5827), 1036–9.

Zakoth, D., Mauroner, O., & Emes, J. (2023). The role of makerspaces in innovation processes: An exploratory study. *R&D Management*, *54*(2), 398–428.

Ziedonis, A. A. (2007). Real options in technology licensing. *Management Science*, *53*(10), 1618–1633.

Zolduoarrati, E., & Licorish, S. A. (2021). On the value of encouraging gender tolerance and inclusiveness in software engineering communities. *Information and Software Technology*, *139*, 106667.