Higher Education Institutions and Their Impact on Employment and Innovation: Regional Identification and Empirical Analyses

Dissertation

submitted to the

Faculty of Business, Economics and Informatics

of the University of Zurich

to obtain the degree of

Doktor der Wirtschaftswissenschaften, Dr. oec.

(corresponds to Doctor of Philosophy, PhD)

presented by

Patrick Lehnert

from Germany

approved in July 2020 at the request of

Prof. Dr. Uschi Backes-Gellner

Prof. Dietmar Harhoff, PhD

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 15.07.2020

Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgments

First and foremost, I thank my doctoral advisor, Prof. Dr. Uschi Backes-Gellner, for her outstanding supervision and for sharing her excellent scientific knowledge and expertise. Throughout my graduate studies, she offered invaluable advice both within and outside academia and thus contributed substantially to my academic and personal advancement.

I am grateful to my doctoral co-advisor, Prof. Dietmar Harhoff, PhD, for his helpful support and insightful advice.

I thank my current and former colleagues at the University of Zurich for the many fruitful discussions and for creating a supportive work atmosphere. In particular, I thank Dr. Curdin Pfister for his academic guidance and the numerous conversations on many aspects of life. Furthermore, I thank Christa Pálffy and Dr. Simone Tuor Sartore for their constant support, Natalie Reid for her editorial guidance, and Marc Egli, Michael Niederberger, Marie-Lena Pfaff, and Vanessa Sticher for their excellent research assistance.

I gratefully acknowledge financial support from the Swiss State Secretariat for Education, Research and Innovation (SERI) through its Leading House on the Economics of Education, Firm Behaviour and Training Policies and from the German Commission of Experts on Research and Innovation (EFI).

I thank my parents and grandparents for their moral, emotional, and financial support throughout all those years I spent on my education. Being able to pursue a university education is a privilege that I do not take for granted.

Finally, my deepest gratitude goes to Natalie. Her unconditional love, constant encouragement, and endless patience are the foundation of everything I have ever achieved.

Contents

List of Tables

List of Figures

List of Abbreviations

AdV	Working Committee of the Surveying Authorities of the Laender of the Federal Republic of Germany
ASTER	Advanced Spaceborne Thermal Emissions and Reflection Radiometer
BKG	Federal Agency for Cartography and Geodesy
BMBF	Federal Ministry of Education and Research
BUR	Business and Enterprise Register
CLC	CORINE Land Cover
DHS	Demographic and Health Survey
DiD	Difference-in-Differences
DMSP OLS	Defense Meteorological Satellite Program Operational Linescan System
EEA	European Environment Agency
EFI	Commission of Experts for Research and Innovation
EPO	European Patent Office
ESDAC	European Soil Data Centre
ESS	Swiss Earnings Structure Survey
FE	Fixed Effects
GADM	Database of Global Administrative Areas
GDP	Gross Domestic Product
GEE	Google Earth Engine
GFSO	German Federal Statistical Office
HEI	Higher Education Institution
ISCED	International Standard Classification of Education
microm	Micromarketing-Systeme und Consult GmbH
MS	Mobilité Spatiale

NDBI	Normalized Difference Built-up Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NOGA	General Classification of Economic Activities
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Squares
PET	Professional Education and Training
PRO	Public Research Organization
R&D	Research & Development
RF	Random Forest
RWI	Leibniz Institute for Economic Research
SCCRE	Swiss Coordination Centre for Research in Education
SERI	State Secretariate for Education, Research and Innovation
SFSO	Swiss Federal Statistical Office
STEM	Science, Technology, Engineering, and Mathematics
UAS	University of Applied Sciences
UNI	Academic University
U.S.	United States
USGS	U.S. Geological Survey
VET	Vocational Education and Training
VIIRS	Visible Infrared Imaging Radiometer Suite

Chapter 1

Introduction

Higher Education Institutions (HEIs) are a key driver of regional economic activity. They contribute significantly to regional economies by increasing the level of human capital and by fostering innovation activities (Valero and Van Reenen, 2019). Numerous empirical studies investigating how the human capital composition of a region's workforce affects economic activity find a positive relationship between economic growth and the percentage of the population with a degree from an HEI (i.e., people who acquired human capital at an HEI) (e.g., Dettori et al., 2012; Peiró-Palomino, 2016; Sterlacchini, 2008). This growth effect occurs because the level of regionally available human capital increases the innovation output of economic actors, in turn leading to economic growth (e.g., Leten et al., 2014; Vandenbussche et al., 2006; Wang, 2010). Moreover, HEIs positively influence regional economic growth through (a) their own research and innovation activities and (b) direct linkages (e.g., cooperative research projects) with other economic actors (e.g., Abramovsky and Simpson, 2011; Adams et al., 2001; Jaffe, 1989). Therefore, the presence of an HEI within a region yields benefits for the region's economy both directly through the institution's innovation activities and indirectly through the institution's effect on human capital formation.

However, HEIs differ in the type of research they conduct and thus in the type of knowledge they impart. Traditional Academic Universities (UNIs) perform basic research and provide their students with basic scientific knowledge, that is, prepare their students for an academic career (Bentley et al., 2015; Mowery and Sampat, 2005). Since the 1960s, however, many European countries have introduced new types of HEIs—often called Universities of Applied Sciences (UASs)—that focus on conducting *applied* research and imparting vocational (rather than academic) knowledge (Lepori and Kyvik, 2010). Despite this development, the vast majority of studies investigating the effects of HEIs on innovation and economic growth analyze only UNIs (e.g., Andersson et al., 2009; Bianchi and Giorcelli, 2020; Toivanen and Väänänen, 2016). UASs, however—with the exception of Pfister et al.'s (2021) initial evidence of a substantial UAS effect on innovation—remain widely unexplored, particularly for their impact on regional economies.

In this dissertation, I investigate how UASs—HEIs with a focus on applied research and vocational knowledge—affect employment and innovation in the regions where they are located. More specifically, I study UASs in Switzerland and Germany, two countries with strong vocational tiers in their education systems. Due to this particular feature of their education systems, Switzerland and Germany (unlike other European countries such as Norway or the United Kingdom) maintain a clear distinction between academic higher education and vocational higher education (Lepori and Kyvik, 2010). Consequently, UASs in Switzerland or Germany might affect regional economies differently than UNIs, because the two types of HEIs have different foci in their research and teaching. As I investigate outcomes related to technological innovation, in my empirical analyses I follow Pfister et al. (2021) by restricting these analyses to UASs that are active in Science, Technology, Engineering, and Mathematics (STEM).[1]

UASs in Switzerland and Germany share both their applied research profile and their focus on vocational knowledge as core legal mandates (Commission of Experts for Research and Innovation (EFI), 2018; Federal Ministry of Education and Research (BMBF), 2004; State Secretariate for Education, Research and Innovation (SERI), 2019). Because of their applied research focus, UASs are more likely to directly cooperate with local firms (Arvanitis et al., 2008; Hachmeister et al., 2015; Warnecke, 2019) and, in so doing, might accelerate and improve innovation processes. Furthermore, because of the UAS focus on

[1] Innovation outcomes from other disciplines, such as the social sciences or the arts, are reflected in other indicators than the ones I use in this dissertation.

vocational knowledge, UAS graduates possess different human capital than UNI graduates (Kiener, 2013; Lackner, 2019; Lepori and Müller, 2016). This vocational knowledge can be a key element for firms' innovation activities (Backes-Gellner and Pfister, 2019). In both Switzerland and Germany, UASs thus possess great potential for contributing to regional innovation activities.

Nonetheless, UASs in the two countries differ in their student selection. On one hand, Swiss UASs admit only students who (a) have completed a dual apprenticeship program with a professional baccalaureate (the Swiss "Berufsmaturität") and thus (b) have already acquired substantial and formal vocational knowledge in a specific occupation (Nikolai and Ebner, 2013; Swiss Coordination Centre for Research in Education (SCCRE), 2018).[2] On the other hand, German UASs admit not only students who have previously completed a dual apprenticeship program but also students with a general baccalaureate (the German "Allgemeine Hochschulreife") or a specialized baccalaureate (the German "Fachhochschulreife") (Lackner, 2019; Nikolai and Ebner, 2013). Those students without prior apprenticeship training usually acquire vocational knowledge during their studies at German UASs through one-semester internships (BMBF, 2004; Lackner, 2019; Nikolai and Ebner, 2013). Given these differences in acquired vocational knowledge, graduates from Swiss UASs possess on average deeper practical knowledge than graduates from German UASs. As this deeper practical knowledge can be crucial for UAS graduates to successfully promote innovation in firms (Backes-Gellner and Pfister, 2019), graduates from Swiss UASs might perform differently in promoting innovation than graduates from German UASs.

This dissertation contributes to a more thorough understanding of how UASs increase regional innovation by studying two aspects of UASs that potentially determine this increase. First, UASs might affect regional innovation through their human capital formation. The combination of vocational and applied research knowledge that UASs give their students might be important for helping local firms increase their innovation

[2] Although studying at a UAS in Switzerland is possible upon fulfillment of certain additional requirements for students with only a general baccalaureate (the Swiss "Maturität"), very few students choose this option (SCCRE, 2018).

outputs. Analyzing UASs in Switzerland, where the student selection process ensures that UAS graduates possess precisely this combination of knowledge, thus can yield important insights into whether local firms use the type of human capital that UASs produce for innovation activities. Second, the research infrastructure available within a region might influence the innovation effect of UASs. Complementarities between UASs and research institutions producing different types of research knowledge might lead to an additional increase in regional innovation. Analyzing UASs in Germany, where a variety of public research institutions coexist, thus can shed light on the extent to which UASs increase innovation in different regional research infrastructures.

In the first part of this dissertation (chapter 2), I study the impact of Swiss UASs at the firm level. More specifically, I examine whether firms increase employment in Research & Development (R&D) as a reaction to the introduction of UASs in Switzerland. If firms consider the skills of UAS graduates a valuable input to their innovation processes, I expect an increase in personnel with R&D tasks and, likewise, an increase in the wage sum paid to R&D personnel. The analysis in Chapter 2 thus contributes to the literature on how the skills available in local labor markets affect firms' innovation activities. While the literature on vocational education and innovation already shows the importance of workers with secondary-level vocational skills for firms' innovation activities (e.g., Rupietta and Backes-Gellner, 2019; Toner, 2010), I analyze whether augmenting the skill base of these workers with applied research skills taught at UASs further boosts innovation activities. Furthermore, I extend Pfister et al.'s (2021) analysis of the development of patents after the introduction of UASs in Switzerland by providing evidence on a particular mechanism—changes in R&D employment—that might lead to the strong increase in patents that they find.

To empirically analyze how the new availability of the skills of Swiss UAS graduates in local labor markets affects R&D employment, I use repeated cross-sectional data from the Swiss Earnings Structure Survey (ESS) and apply Pfister et al.'s (2021) identification strategy. With the ESS data, I measure at the establishment level what percentage of a firm's employees performs R&D tasks and what percentage of the total wage sum a firm spends on these employees. Furthermore, I assign establishments to treated labor market regions (i.e., regions located within 25 kilometers of a UAS campus) and untreated labor market regions (i.e., regions located farther away than 25 kilometers from a UAS campus). As the location and timing of UAS campus openings in Switzerland were the result of a quasi-random process involving complex political negotiations, and thus unrelated to employment or innovation, these openings serve as a quasi-natural experiment that allows the identification of causal effects (Pfister et al., 2021). Consequently, to investigate how treated establishments change their R&D employment, I follow Pfister et al. (2021) by applying a Difference-in-Differences (DiD) framework that exploits the spatial and temporal variation in the openings of Swiss UAS campuses.

My analysis of Swiss UASs shows that establishments treated by a UAS campus employ more R&D personnel and pay a larger percentage of their total wage sum to R&D personnel. This result implies that firms value the skills of UAS graduates and use these skills for innovation activities. Assessments of heterogeneous effects across firms suggest that both very small firms (including potential start-ups) and very large firms profit most. The introduction of UASs in Switzerland thus stimulates both new R&D activities and the R&D activities of firms that had conducted R&D before it. In sum, the findings in Chapter 2 imply that the regional availability of human capital acquired at UASs increases firms' innovation activities.

In the second part of this dissertation (chapters 3 and 4), I study the impact of German UASs at the regional level. More specifically, I examine the evolution of patents in regions where a UAS has opened in comparison to regions without a UAS. Moreover, to determine whether the innovation effect of UASs depends on the regional research infrastructures that UASs can draw upon, I investigate complementarities between UASs and other types of research institutions. If UASs and such institutions produce complementary research knowledge, their coexistence within a region can lead to a further increase in innovation. In addition to UASs (which perform applied research) and UNIs (which perform basic research), the German landscape of research institutions comprises the large sector of Public Research Organizations (PROs) such as the Max Planck Society or the Fraunhofer

Society. The research activities of these PROs cover the entire spectrum from basic to applied research. However, despite evidence on the separate innovation effects of different types of research institutions (e.g., Comin et al., 2019; Intarakumnerd and Goto, 2018; Popp, 2017), the question of whether their coexistence within a region leads to knowledge complementarities remains unanswered.

As in the analysis of Swiss UASs in Chapter 2, I exploit variation in the location and timing of UAS campus openings in Germany to identify their effect on the evolution of patents. However, unlike in Switzerland, this variation did not quasi-randomly occur in Germany. To account for the resulting endogeneity in the location and timing of UAS campus openings, the empirical literature investigating the effects of the introduction of HEIs offers diverse strategies. For example, in analyzing the openings of United States (U.S.) colleges, Andrews (2020) utilizes the circumstance that their locations were competitively determined. He then identifies runner-up locations as a comparison group. In an analysis of Italian UNIs, Cowan and Zinovyeva (2013) exploit the panel structure of their data to account for time-invariant unobservable sources of endogeneity. Furthermore, to account for time-variant observable sources of endogeneity, they include an extensive set of control variables in their regressions.

Unfortunately, suitable identification strategies or comprehensive control variables do not exist for Germany. Therefore, to apply an identification strategy such as Cowan and Zinovyeva's (2013)—that is, one that uses controls for sources of endogeneity—for studying the regional innovation effect of German UASs, I require data on factors determining both innovation and UAS campus locations, such as regional economic activity. More specifically, the analysis of the regional innovation effects of German UASs requires data that goes back at least to 1985, when German UASs began their research activities (Enders, 2010; BMBF, 2004; Kulicke and Stahlecker, 2004; Wissenschaftsrat, 2002). In addition, to be able to precisely define treatment and control regions, the analysis needs detailed regional data at the municipality level (the smallest administrative regional unit in Germany). However, data on economic activity with (a) a sufficiently long time series (at least from 1985) and (b) the necessary level of regional detail (municipalities) does not exist for Germany. Yet, this absence of data occurs for any type of evidence-based analysis on the regional effects of local policy reforms and is thus not unique to the analysis of German HEIs.

Therefore, in Chapter 3, I develop a method for retrieving a proxy measure fulfilling these criteria. Inspired by the literature that uses nighttime satellite imagery for proxying Gross Domestic Product (GDP) (e.g., Chen and Nordhaus, 2011; Henderson et al., 2012), I use daytime imagery from Landsat satellites for proxying regional economic activity. Compared to night lights, the most commonly used satellite-based proxy for economic activity, Landsat data has two advantages for evidence-based policy analyses of local policy reforms. First, Landsat data constitutes the earliest satellite data (Morain, 1998; Williams et al., 2006), making my proxy measure available over a uniquely long time series. Such a long time series allows researchers to study local policy reforms that took place more than 30 years ago (such as the introduction of UASs in Germany). Second, due to its much higher spatial resolution than, for example, night lights data, Landsat data provides information at a more detailed regional level (Donaldson and Storeygard, 2016). Therefore, Landsat data enables researchers to investigate the effects of policy reforms in small localities for which other data sources are often not available at all (such as German municipalities) or not available to non-resident researchers due to access restrictions. Therefore, the proxy for economic activity I derive from Landsat data complements other data sources, such as administrative statistics or night lights, for evidence-based analyses of local policy reforms.

This proxy consists of six different surface groups on the earth, with each group representing different types of land cover (terrestrial surface features such as built-up land, forest, or cropland). To classify every pixel location—Landsat's unit of observation, which encompasses a geographical area of 30×30 square meters at the equator—into one of the six surface groups, I apply supervised machine-learning techniques. In so doing, I closely follow the literature that identifies land cover from Landsat data (e.g., Dewan and Yamaguchi, 2009; Goldblatt et al., 2016). The result of this procedure is an annual dataset indicating the surface group of every pixel location in Germany from 1984 through 2020.

To assess the value of surface groups as a proxy for economic activity, I analyze data on GDP and household income, both covering a limited time series of the years for which I compute the surface groups. To show that surface groups are a valid proxy for regional economic activity, I use these data as indicators of economic activity, performing Ordinary Least Squares (OLS) regressions for these limited time series and at the available levels of regional disaggregation. The GDP and household income data are available at detailed regional levels, with some even smaller than municipalities. For these detailed regional levels and for the limited time series that the GDP and household income data cover, surface groups explain a large percentage of the variation in economic activity. Moreover, I compare the predictive value of surface groups to that of night lights, finding that surface groups more accurately proxy economic activity at the detailed regional levels that I analyze.

In Chapter 4, I empirically analyze the innovation effect of UASs in Germany. To measure innovation, I follow Pfister et al. (2021) by using patent data from the European Patent Office (EPO) to construct indicators of both innovation quantity and quality at the regional level of municipalities. To account for endogeneity in the regional distribution of German UAS campuses, I include the surface groups I derive in Chapter 3 as control variables in the regressions. As these surface groups are a valid proxy for economic activity at small regional units such as municipalities, they account for (at least a part of) the time-variant sources of endogeneity in the distribution of UAS campuses. In addition, to account for the time-invariant unobserved sources of endogeneity, I exploit the panel structure of the patent data and perform Fixed Effects (FE) estimations.

The analysis in Chapter 4 shows that UASs positively influence regional innovation, with this effect becoming much larger when UASs can draw upon the knowledge of other research institutions that coexist within a region. Knowledge complementarities arise between UASs and research institutions focusing on basic research (e.g., Max Planck institutes) and those focusing on applied research (e.g., Fraunhofer institutes). These complementarities increase both the quantity and the quality of regional innovation outputs. While UASs thus increase innovation in regions where no other research institution exists, they possess greater potential for increasing innovation in regions where they can also draw upon the knowledge of coexisting research institutions. The coexistence of institutions producing research knowledge complementary to that of UASs thus further increases innovation beyond the increase from a stand-alone UAS.

Finally, Chapter 5 concludes that through providing evidence on two aspects of the UAS effect on regional innovation, this dissertation provides important evidence on how UASs contribute to regional economies. Thus far, the literature analyzing the contributions of HEIs to regional economies has largely neglected UASs, even though UASs are essential elements of the education systems of many European countries. By studying UASs in Switzerland and Germany, two countries in which UASs have a distinct focus on vocational knowledge and applied research (despite some institutional differences between the Swiss and German ones), this dissertation contributes to a better understanding of how UASs affect regional economies.

I provide evidence on two aspects of the contribution of UASs to regional economies. First, firms value the vocational and applied research skills that UASs teach their students. These skills present new resources for firm-level innovation, as indicated by the increase in R&D employment and R&D wage sum in firms affected by a UAS opening. Second, knowledge complementarities arise between UASs and other research institutions, leading to further increases in the UAS effect on innovation. Taken together, the findings of this dissertation show that UASs—with their unique focus on vocational knowledge and applied research—significantly contribute to regional economies. The type of human capital that UASs produce and the type of research that UASs conduct constitute important drivers of regional innovation.

Chapter 2

Employment of R&D Personnel After an Educational Supply Shock: Effects of the Introduction of Universities of Applied Sciences in Switzerland

This chapter was published in 2020 in *Labour Economics Vol. 66* as “Employment of R&D personnel after an educational supply shock: Effects of the introduction of Universities of Applied Sciences in Switzerland” by Lehnert, Pfister, and Backes-Gellner.

2.1 Introduction

A tertiary-level education expansion creates a supply shock of graduates with R&D-specific skills in the labor market, giving firms easier and more opportunities to hire graduates for their R&D departments. Therefore, governments often expand tertiary-level education institution to increase innovation and competitiveness through providing more future workers with the skills necessary for R&D activities (Organization for Economic Co-operation and Development (OECD), 2010, 2017). However, causal empirical evidence on the employment effects of tertiary education expansions is rare, particularly on how firms respond in terms of employing R&D personnel. Thus investigating whether and, if so, to what extent firms make use of workers with new R&D skills is very important. More specifically, analyzing whether firms employ more R&D personnel and increase overall spending on R&D personnel (as wages) is crucial.

This chapter analyzes how the educational supply shock resulting from the introduction of a particular type of tertiary education institutions[1]—UASs, which recruit their students from vocational apprenticeship graduates and which both teach and conduct applied research—has affected firms' employment in R&D in Switzerland. The government introduced UASs to provide graduates from dual apprenticeship programs (i.e., upper-secondary-level vocational education) with an educational upgrade to the tertiary level (SCCRE, 2018). Therefore, students at UASs are not comparable to students at other tertiary education institutions, because UAS students already have a sound vocational knowledge base through their secondary-level vocational education. By obtaining a UAS degree, these students attain more advanced vocational and professional knowledge combined with applied research skills, often using them for, and in close cooperation with, local firms. Therefore, we expect the introduction of UASs to have different effects on firms' R&D than the introduction of UNIs, which generally recruit their students from high school graduates. Due to the UASs' close combination of a strong practical vocational knowledge base with applied research skills, we expect substantial effects on R&D and innovation, particularly in firms or regions that did not previously have a strong tradition of innovation.

To identify the causal effect of the introduction of UASs, we exploit a quasi-random variation in the location and timing of the openings of UAS campuses in Switzerland in the 1990s. In so doing, we follow a growing literature that uses the openings of new tertiary education institutions as an identification strategy (e.g., Jäger, 2013; Kamhöfer et al., 2019; Kyui, 2016; Pfister et al., 2018; Toivanen and Väänänen, 2016). We apply a DiD design to compare the employment of R&D personnel in treated firms (i.e., firms in labor market regions where a UAS campus has opened) to the employment of R&D personnel in untreated firms (i.e., firms in labor market regions where no UAS campus has opened).

[1] This chapter uses the term "tertiary education institution" as an equivalent to "higher education institution (HEI)".

For our analysis, we draw on repeated cross-sectional data from the ESS. This data allows us (a) to observe firms at the establishment level (i.e., at their different locations), which we need to identify treated and untreated locations, and (b) to precisely measure firms' R&D personnel by providing information on the job tasks of individual workers and, in turn, on firms' direct labor input dedicated to innovation activities. Through these unique ESS data features, we can investigate how the labor supply shock resulting from the introduction of UASs affects the employment of R&D personnel (measured as the percentage of a firm's personnel with R&D tasks as their main job activity). Moreover, we can also investigate the effect on wages paid to R&D employees (measured as the percentage of the total wage sum paid to personnel with R&D tasks as their main job activity) in treated establishments in comparison to untreated establishments.

Our analysis shows that firms affected by the opening of a UAS campus employ more R&D personnel and spend more money on R&D personnel, clearly engaging more intensively in R&D. This finding implies that firms located near a UAS campus use the R&D skills available in the labor market after the opening of the campus. Furthermore, we study whether these effects are heterogeneous across different types of firms and find that both very small firms (with 5–9 employees—and thus possibly start-ups) and very large firms (with 5,000 or more employees) profit from the introduction of UASs. These results show that an education expansion providing individuals with relevant practical skills in applied R&D can stimulate firms' innovation activities. Particularly for small firms, the skills of UAS graduates constitute a valuable resource, enabling these firms to engage in or intensify their R&D activities.

This chapter contributes to three strands of the literature on how the skills available in local labor markets influence firms' innovation activities. First, we contribute to the literature on the innovation effects of tertiary education institutions (which produce skills for local labor markets) by providing evidence on the innovation effect of introducing a new tertiary education institution that—compared to a UNI—teaches different types of skills to graduates of different types of secondary education. Despite numerous studies on the effect of UNIs on innovation (e.g., Audretsch and Feldman, 1996; Cowan and Zinovyeva, 2013; Jaffe, 1989), no evidence exists on the effect of the structurally very different UASs—tertiary education institutions teaching graduates from dual apprenticeship programs and focusing on vocational and applied research knowledge—on innovation activities.

Studies investigating the innovation effect of traditional UNIs, which teach high school graduates and focus on theoretical knowledge and basic research, confirm that firms gain from UNIs and that these spillovers are concentrated in firms close to a UNI (e.g., Andersson et al., 2009; Anselin et al., 1997; Autant-Bernard, 2001). Moreover, studies examining the innovation effect of UNIs focusing on technical knowledge in STEM—universities that still teach graduates from high-school programs and that do not focus on vocational and applied research knowledge—show that these institutions raise, for example, the propensity for graduates to become inventors (Bianchi and Giorcelli, 2020; Toivanen and Väänänen, 2016). In addition, studies on the openings of U.S. colleges, some of which teach and conduct applied research but primarily to high school graduates, also show positive innovation effects (Andrews, 2019; Moretti, 2004).

Nonetheless, all these tertiary education institutions structurally differ from the UASs we analyze in this chapter. Although some of these institutions teach a certain amount of applied or STEM-related skills, they do not focus on students who are graduates from dual apprenticeship programs and who already have a solid practical vocational knowledge base from their three- to four-year education, which includes an apprenticeship in firms. Thus we specifically analyze the effects of the close combination of solid practical vocational skills from a dual apprenticeship program with applied research skills. Such a combination might have particularly sizable effects on R&D intensity and innovation in firms or regions that did not previously have a strong tradition of innovation.

Second, we add to the literature that examines the importance of vocational skills for innovation activities by demonstrating that augmenting the skill base of vocationally trained workers with applied research skills contributes to innovation activities in firms that have access to the new type of graduates. Only few studies show that secondary-level vocational skills (i.e., those of graduates from dual apprenticeship programs) positively affect innovation in firms (e.g., Meuer et al., 2015; Rupietta and Backes-Gellner, 2019; Toner, 2010). Moreover, Cinnirella and Streb (2017), who study 19$^{\text{th}}$-century Prussia, identify the knowledge of "master craftsmen" (a form of advanced skills in a specific occupation) as an important driver of technological development in that century. This chapter extends these findings by providing first evidence that augmenting secondary-level vocational skills (which UAS students already have from their dual apprenticeship training) with tertiary-level applied research skills has additional effects on firms' innovation activities, as measured by an increase in R&D employment and intensity.

Third, we extend the literature that assesses the innovation effect of tertiary education expansions (e.g., Cowan and Zinovyeva, 2013; Leten et al., 2014; Toivanen and Väänänen, 2016) by directly showing that firms' use of the newly available skills for their R&D activities constitutes one mechanism underlying the innovation effect of tertiary education expansions. While Pfister et al. (2018) find that the introduction of UASs in Switzerland increases innovation outcomes (as measured by patenting activities), we show that firms' employing UAS graduates as an input for their R&D activities is a potential driver of the increase in innovation outcomes. Moreover, our results show that the particular skill combination of UAS graduates (sound practical vocational knowledge and applied research skills) forms an important missing link between vocationally trained middle-skilled workers in production and academically trained workers in R&D. Their presence helps increase innovation capabilities through a more effective combination of different types of knowledge and, consequently, more creative solutions to innovation problems (Backes-Gellner and Pfister, 2019; Schultheiss et al., 2019).

The chapter proceeds as follows: Section 2.2 explains the institutional background of the introduction of UASs in Switzerland and hypothesizes how the resulting supply shock of skilled labor may influence firms' R&D. Section 2.3 describes our ESS data and our measures of R&D personnel. Section 2.4 presents and discusses our DiD approach to identify the treatment effect. Section 2.5 reports the main results and further assesses whether the treatment effect is heterogeneous across different types of firms. Section 2.6 concludes.

2.2 Institutional Background

2.2.1 UASs and the Vocational and Applied Research Skills of Their Graduates

In comparison to UNI graduates, UAS graduates possess a very distinct and unique set of vocational and applied research skills. This distinction results from the legal mandates of UASs within the Swiss education system. This system consists of both a vocational and an academic pillar at the tertiary level (according to the International Standard Classification of Education (ISCED)). About 70 percent of students who complete compulsory schooling opt for a vocational track by starting some form of Vocational Education and Training (VET), usually a dual apprenticeship program.[2] An apprenticeship program includes practical on-the-job training in host firms (about 80 percent of the time) and theoretical teaching in vocational schools (about 20 percent), with both parts adhering to well-defined curricula. Graduates receive a nationally recognized VET certificate.[3] Until the 1990s reform, VET students had no direct career path to a university education, as UNIs accepted only students from the academic (i.e., non-VET) education track. The goal of creating the UASs was to provide an equivalent—yet different—educational path for VET graduates.[4]

After the 1990s reform, dual VET graduates who acquired a professional baccalaureate (*Berufsmaturität*)—a degree that supplements the VET degree and requires additional courses and exams—had the new option of entering a UAS. While studying at a UAS, these dual VET graduates accumulate a combination of advanced occupational knowledge and applied research skills. Thus UAS graduates possess two types of skills: First, they have a solid vocational skill set within their occupation (from, e.g., a four-year apprenticeship training as a laboratory technician or a three-year apprenticeship training as a mechanic).

[2] See https://www.bfs.admin.ch/bfs/en/home/statistics/education-science/diploma/upper-secondary.html (last retrieved on April 16, 2019).

[3] See, e.g., Backes-Gellner et al. (2017) or Eggenberger et al. (2018) for a short description of the economically important aspects of the Swiss VET system.

[4] In addition to giving VET graduates a formal education equivalent to that of UNIs, UASs are legally required to apply scientific methods and knowledge in their teaching and research, to provide services to public or private sector firms, and to collaborate with firms and other research institutions. For further information on UASs and their legal mandates, see Projektgruppe Bund–Kantone Hochschullandschaft 2008 (2004) or SERI (2015).

Second, through their UAS education, they have enhanced their original skill set and acquired new applied research knowledge in their occupational field. Their skill set thus differs from that of UNI graduates, who have acquired more theoretical and basic research skills, not practical *vocational* and *applied* research skills.

This difference in the skill profiles of UAS graduates and UNI graduates results from the strict separation of the academic and vocational tracks in the Swiss education system. The selection mechanism determining eligibility for a UAS ensures that the skills of UAS graduates differ from those of UNI graduates (for an overview of transitions from secondary to tertiary education in Switzerland, see SCCRE, 2018). After completing their secondary-level education, students in the academic track can easily proceed to a UNI and students in the vocational track (with a professional baccalaureate) can easily proceed to a UAS. Although theoretically possible (i.e., if students fulfill certain additional requirements after completing secondary education), very few students transfer from the academic to the vocational track, or vice versa. Through this design feature of the Swiss education system, students at UASs have a strong vocational knowledge base to build upon, and UASs add applied research skills.

The field of electrical engineering offers a clear illustration of the differences between a UAS graduate and a UNI graduate. A UNI graduate with a degree in electrical engineering would, for example, acquire the necessary basic research skills to achieve theoretical progress in how to more efficiently convert solar energy into electric power. In contrast, a UAS graduate would learn how to apply the results of that research to the actual development of a photovoltaic system. Thus the UAS graduate can conduct applied research projects on how to apply this knowledge in the production of solar panels for rooftops.

In addition, the UAS mandate requires UASs to conduct applied research projects in cooperation with private-sector firms. Thus UASs frequently collaborate with (local) firms (Arvanitis et al., 2008), giving students opportunities to practice applied research on real-world problems during their studies. Consequently, given UAS goals, we expect UAS graduates to be well suited for integration into firms' R&D activities. We therefore focus on UAS campuses specializing in STEM, because we expect firms' technological R&D activities to concentrate in these fields.[5] As Schultheiss et al. (2019) show, UAS graduates are able to provide the missing link between R&D workers with academic knowledge and workers with vocational knowledge, thereby creating synergies that increase innovation outputs. Case studies of innovative Swiss firms, conducted by Backes-Gellner and Pfister (2019), also support the existence of such synergy effects. Therefore, we expect firms located near a UAS campus specializing in STEM to employ more R&D personnel to achieve efficiency gains. Likewise, we expect these firms to spend a larger percentage of their total wage sum on the wages of R&D employees, to further increase innovation outputs and improve competitiveness.

2.2.2 The Introduction of UASs: A Quasi-Random Process

The introduction of UASs in Switzerland in general and the location decisions in particular created a quasi-random outcome resulting from a very complex political process (Pfister et al., 2018). To analyze the location and timing decisions, and to determine whether these decisions truly created a quasi-random outcome, Pfister et al. (2018) examined a large number of sources, including policy reports and legal documents from the federal, cantonal,[6] and municipal[7] governments; official bulletins and historical records; reports from the UAS commission and the UAS councils;[8] UAS annual reports; federal and cantonal laws and intercantonal agreements; and more than 100 articles in 16 newspapers from the relevant period. Given Switzerland's federalist political system, Pfister et al. (2018) find that both the temporal and spatial variations of UAS campus introductions are quasi-random.

Specifically, Pfister et al.'s (2018) analyses show that a political trench warfare within

[5] A number of studies show the importance of STEM workers for adopting and generating innovation, and for increasing economic productivity and growth (e.g., Griliches, 1992; Jones, 1995; Peri et al., 2015; Winters, 2014). Moreover, firms with higher percentages of creative and STEM workers are more innovative (Brunow et al., 2018).

[6] Switzerland has 26 cantons, which function similarly to states in the U.S.

[7] Municipalities (*Gemeinden*) are similar to U.S. counties.

[8] See Appendix 2.7.5.1 for descriptions of the different institutional actors (including the UAS commission and the UAS councils) involved in the UAS introduction process.

and between cantons primarily drove (a) the introduction of new campuses in a particular region and (b) the closing or relocation of old ones. Although the political authority conferring UAS accreditation was the federal government, the political units bearing the main financial burdens were the cantons. The federal government's legal requirements for UASs (see section 2.2.1)—a solid financial base, campuses large enough to accommodate a sizable number of students, and equally distributed campuses throughout Switzerland—led to a heated political debate between (and even within) cantons. Historical events, personalities, micropolitics, package deals, concessions, and coalition building (related to, e.g., the introduction of a new UAS campus specializing in a field of study other than STEM, such as business or health) became fundamental determinants of the quasi-random location decisions and time delays in the openings of UAS campuses. For example, the Bern campus opened in 1997 and closed again in 2003, due to federal merging and relocation requirements.

The process resulted in 15 STEM campuses belonging to five UASs in the German-speaking area of Switzerland. We focus on this area of the country (not the French- or Italian-speaking parts) because the VET system is most strongly embedded—both culturally and institutionally—in the German-speaking part (Bolli et al., 2018; Freitag and Bühlmann, 2003). The overall variation in the location and the timing of the introduction of UAS STEM campuses was related to various political factors, not to a region's economic strength or level of innovation. As this variation was not foreseeable, it was thus unrelated to local economic characteristics.

Appendix 2.7.5 contains an in-depth description of the process leading to the introduction of UASs and generating the quasi-random variation, including an overview of UAS campus locations, introduction years, and illustrative examples. The qualitative evidence in Appendix 2.7.5.1 confirms the quasi-randomness of the introduction process. Results of a quantitative analysis of pre-treatment trends in Section 2.4.3 and extensive empirical evidence based on economic indicators from both ESS and other data in Appendix 2.7.5.2 strongly support the quasi-random distribution of UAS campuses. This appendix details the economic preconditions of UAS campus regions, the composition of the treatment and control groups, and any unobservable factors potentially determining UAS campus introductions. All these analyses support our argument that the introduction of UASs in Switzerland constitutes a quasi-random process.

2.3 Data

To estimate the effect of the supply shock of skilled labor on firms' employment of R&D personnel, we use the largest representative firm survey available in Switzerland, the ESS. The Swiss Federal Statistical Office (SFSO), which started the ESS in 1994, conducts it biennially, thus covering the relevant period of the UAS openings between 1997 and 2003. Every wave of the ESS comprises a new stratified random sample of all firms in Switzerland, with survey participation mandatory. Our data thus constitutes repeated cross-sections.

The ESS contains information on more than 300,000 firm-year observations between 1994 and 2014, and each firm provides detailed information on its employees and their tasks.[9] To determine the location of an employee's workplace, the ESS indicates the Mobilité Spatiale (MS) regions,[10] that is, homogeneous micro-regions whose boundaries rely on (among other things) regional mobility analyses and structural labor markets (Schuler et al., 2005). Therefore, MS regions are well suited for our analysis. The next higher administrative regions are the cantons; the next lower, the municipalities. Each canton consists of about four MS regions, and an MS region consists on average of about 22 municipalities.

Following Janssen et al. (2016), who exploit variation in workplace locations within firms at the cantonal level, we use the information on employees' workplace locations

[9] The number of employees that each firm reports in the survey depends on firm size: Firms with fewer than 20 employees report information on every employee; firms with 20 to 49 employees randomly report every second employee; and firms with 50 employees or more randomly report every third employee. However, as the data shows that some firms chose to report more observations than necessary, we do not rely on the three firm-size categories when calculating the variables for our estimations.

[10] In 1982, a federal study defined the non-administrative MS regions (Schuler et al., 2005). Since then, the borders of these regions have not changed, except for minor revisions resulting from municipality mergers (Schuler et al., 2005).

to determine the locations of firms' establishments.[11] We split firms by assigning all employees working in the same MS region to one establishment.[12]

Firms surveyed in the ESS also report the main job activity—e.g., construction, secretarial, R&D[13]—of their employees and the exact wage in October of the survey year, thereby allowing us to precisely measure firms' R&D personnel at the establishment level. From this information, we calculate two outcome variables that represent firms' R&D personnel (see appendix 2.7.1 for descriptive statistics).

First, the percentage of R&D personnel (RDP^{pct}) gives the fraction of employees with R&D as their main job activity relative to the total number of employees within an establishment (adjusted to full-time equivalents). We calculate this variable in the following way:

$$RDP_i^{pct} = \frac{\sum_{j=1}^{N_i} rd_j l_j}{\sum_{j=1}^{N_i} l_j} \tag{2.1}$$

with i as the establishment, j as the employee, N as the number of employee observations, rd as the binary indicator of R&D as main job activity, and l as the individual employment level ($0 < l \leq 1$).[14] As Equation 2.1 shows, we adjust the variable to full-time equivalents by weighting each employee observation with the individual employment level, thereby avoiding potential biases caused by part-time employment.

Second, the percentage of R&D wages (RDW^{pct}) indicates the fraction of wages paid to R&D employees relative to the total wage sum. For this variable, we proceed similarly:[15]

$$RDW_i^{pct} = \frac{\sum_{j=1}^{N_i} w_j rd_j}{\sum_{j=1}^{N_i} w_j} \tag{2.2}$$

11 The results in Table 2.2 are robust to restricting the sample to single-establishment firms (see appendix 2.7.6.1).

12 This procedure eliminates any biases that might result from fixing the location of a multi-establishment firm to, for example, the MS region where most of its employees work, even though the firm conducts R&D at an establishment in a different location.

13 R&D is one of 24 categories in the job activity variable.

14 For individuals with $l > 1$ we set $l = 1$.

15 Before aggregating at the establishment level, we deflate every wage observation in the dataset to 2010 prices according to the Consumer Price Index provided by the SFSO. See https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases.assetdetail.cc-d-05.02.08.html (last retrieved on September 6, 2018).

with w as the monthly wage.[16]

With this set of dependent variables, we can investigate whether firms employ more workers performing R&D, relative to other workers, and whether firms also spend more on total R&D personnel. As firms can report only one job activity per employee, our outcome variables can be considered lowered bounds; therefore, they are strong indicators for R&D-active personnel.

2.4 Methodology

2.4.1 Identification Strategy

Using Pfister et al.'s (2018) identification strategy, we exploit the quasi-random variation in the location and the timing of UAS campus openings to estimate the causal effect of the resulting supply shock of skilled labor on firms' employment of R&D personnel. Between 1997 and 2003, 15 UAS STEM campuses opened in the German-speaking part of Switzerland (for an overview, see figure 2.1 and appendix 2.7.5.1), where the VET system is particularly strong. As we argue in Section 2.2.2, the openings of these campuses followed a quasi-random process.

To estimate how firms adjust their R&D personnel after the educational supply shock, we apply the following DiD model:

$$Y_{i,t} = \beta_0 + \beta_1 Treatment_{i,t-3} + \beta_2 TreatmentGroup_i + \beta_3 t + \beta_4 X_{i,t} + \mu_{i,t} \tag{2.3}$$

In this model, Y is the dependent variable (RDP^{pct} or RDW^{pct}) for establishment i in survey year t. The treatment variable $Treatment$ equals one for establishments in the treatment group when the treatment sets in, that is, when the UAS campus has opened. The coefficient β_1 thus identifies the treatment effect we are interested in. Given the

[16] Firms report wages in October of each survey year. To avoid any bias, we use the base salary, with no extra earnings such as overtime or bonus payments.

standard curriculum lenght of six semesters and the average graduation time of slightly more than three years,[17] we allow for a time lag of three years between the opening of a UAS campus and an expected treatment effect. The binary variable $TreatmentGroup$ identifies whether an establishment belongs to the treatment group or to the control group, that is, whether or not it is located within a region that is affected by a STEM UAS campus, thereby capturing time-invariant differences between establishments in the treatment and control groups. To control for time trends that are common to all establishments, we include a set of dummies indicating the survey year t. The vector X contains a set of establishment-level control variables and μ is the error term.[18]

As establishment-level control variables, we include the industry sector at the 2-digit level of the General Classification of Economic Activities (NOGA) 2002,[19] firm size in three categories as indicated in the Business and Enterprise Register (BUR),[20] and the canton[21] of a firm's main location.[22] We do so for two reasons. First, the ESS sampling

[17] For example, the Winterthur campus opened in 1998. According to the standard curriculum of six semesters, the first graduates would have left the UAS and entered the labor market in 2001. As we use biennial data and thus do not observe firms in 2001, in Equation 2.3 we observe that $Treatment_{i,t-3} = 0$ if $t \leq 2000$ and $Treatment_{i,t-3} = 1$ if $t \geq 2002$ for all firms in the Winterthur campus treatment region. Likewise, for firms in the treatment region of the Burgdorf campus, which opened in 1997, $Treatment_{i,t-3} = 1$ if $t \geq 2000$. The three-year time lag, the most conservative lag for our model, might lead to an underestimation of the true treatment effect, because the lag does not cover students who stay at a UAS for more than six semesters (e.g., due to part-time work during studies).

[18] We cluster standard errors at the firm level, because establishments belonging to the same firm share common organizational structures affecting personnel decisions (among other things).

[19] In 1994, sampling was based on the "Allgemeine Systematik der Wirtschaftszweige" (ASWZ) 1985. We convert this classification to the NOGA 2002 according to the correspondence tables provided by the SFSO, see https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases/publications.assetdetail.176072.html (last retrieved on April 16, 2019) and https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases/publications.assetdetail.82758.html (last retrieved on April 16, 2019). Since 2010, sampling in the ESS is based on the NOGA 2008. However, the ESS also provides the NOGA 2002 industry sector for survey year 2010.

[20] The first category comprises firms with fewer than 20 employees; the second, firms with 20 to 49 employees; and the third, firms with 50 or more employees.

[21] Until 2000, sampling in the ESS was not based on firms' locations. Since 2002, it has been based on which of seven greater regions a firm is located in. However, cantons can request that the ESS draws samples representative for the respective cantons in single years. Therefore, cantons are the lowest regional level at which sampling actually takes place in the ESS.

[22] As the ESS does not provide information on a firm's main location, we reconstruct it in line with the SFSO (2016) definition of the main location as the location where the largest number of employees work. In other words, the main location is set to the canton with the most employee observations in the data. If the number of employee observations is equal for two or more cantons, the main location is set to the canton with the largest cumulative wage sum. If both the number of employee observations and the cumulative wage sum are equal for two or more cantons, the main location is set to the canton with the longest tenure of an individual employee. But if the main location cannot be identified according to this stepwise procedure, we drop the respective observations from our sample (20 establishment-year

procedure depends on these three characteristics and varies between survey years, affecting the composition of firms in our sample (e.g., general sample size increase since 2002, overrepresentation of single cantons within survey years). Second, independent of the treatment, the sampling characteristics influence our outcome variables and thus need controlling for (e.g., Lechner, 2010). Therefore, our analysis identifies the average treatment effect on the treated conditional on the three sampling characteristics of industry sector, firm size, and canton.

2.4.2 Definition of Treatment and Control Groups

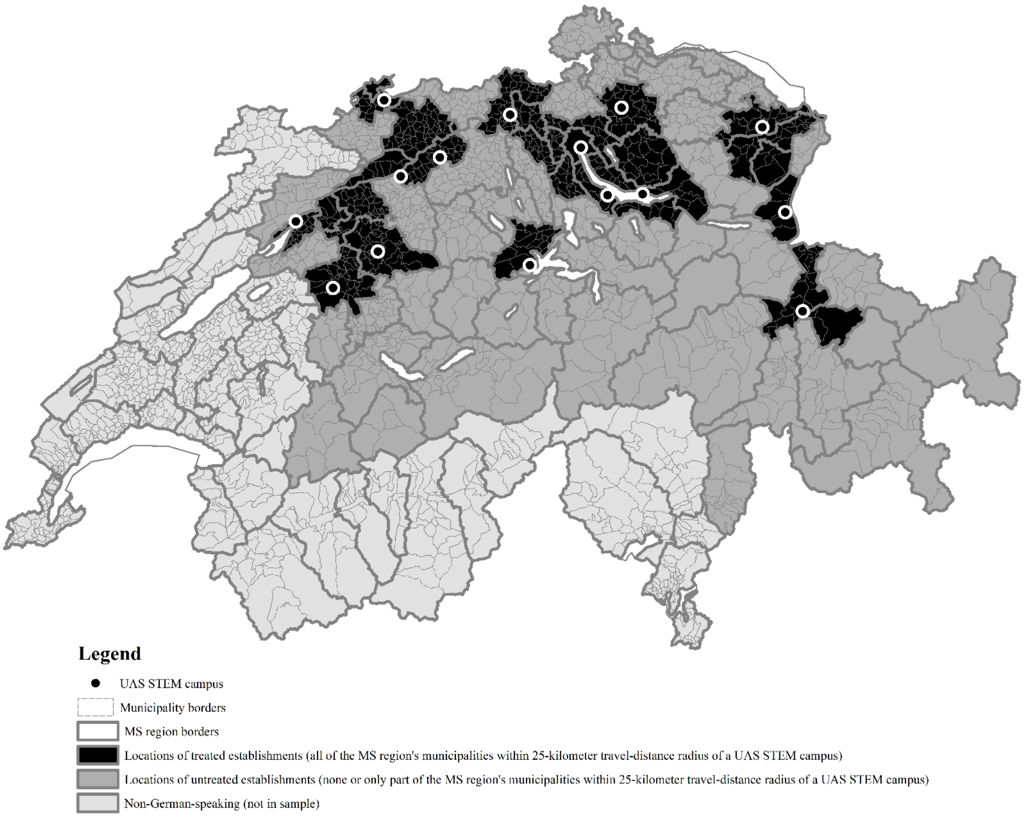
We use MS regions to define whether an establishment is treated or untreated based on its location. As Pfister et al. (2018) demonstrate, Swiss citizens commute only small travel distances, with around 90 percent commuting less than 25 kilometers (15.5 miles) to work. Moreover, Pfister's (2017) analysis and our additional analyses in Appendix 2.7.5.2 show that the net mobility of UAS graduates moving from the treatment to the control group and vice versa is extremely low. Following these arguments, we also assume that the local labor market for UAS graduates is restricted within a 25-kilometer travel-distance[23] radius of a UAS campus.

However, the ESS data contains regional information only at the MS region level, but not at the municipality level as in Pfister et al.'s (2018) analysis. As the MS regions constitute a classification at a geographical level larger than municipalities (i.e., an MS region consists of several municipalities), we are not able to reconstruct the exact borders of the 25-kilometer area around a UAS campus with our data. Instead, we use Pfister et al.'s (2018) municipality-level data to identify how many of the municipalities belonging to an MS region are located within 25 kilometers of a UAS STEM campus. We calculate for each MS region the percentage of treated municipalities (i.e., the number of treated municipalities relative to the total number of municipalities within an MS region). To be conservative, we assign to the treatment group only establishments located in MS regions in

observations).

[23] We use Pfister et al.'s (2018) travel distance measure, calculated with the Google application programming interface.

Figure 2.1: Geographic locations of UAS campuses in STEM, treated establishments, and untreated establishments



Source: Authors' illustration with geodata from SFSO, GEOSTAT (Generalisierte Gemeindegrenzen der Schweiz, Ausgabe 2015) available from https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik/administrative-grenzen/generalisierte-gemeindegrenzen.assetdetail.330759.html (last retrieved on April 16, 2019).

which 100 percent of the MS region's municipalities are treated (i.e., within 25 kilometers of a UAS STEM campus). We then assign to the control group all establishments in MS regions in which less than 100 percent of the MS region's municipalities are treated. In other words, as soon as one municipality within an MS region is untreated, we assign the entire MS region to the control group. This assignment procedure leads to an underestimation of the true treatment effect, because the control group may also contain treated establishments within 25 kilometers of a UAS STEM campus. Our estimation results thus provide a lower bound of the true treatment effect.[24] Figure 2.1 graphically illustrates the geographic locations of UAS STEM campuses and those of treated and untreated establishments.

2.4.3 Assessment of Parallel Trends Assumption

The crucial assumption of a DiD model is the parallel trends assumption—that both treatment and control groups have to show the same time trend in the dependent variable in absence of the treatment. In our case, the first UAS STEM campuses opened in 1997. With the three-year time lag, the first year in which we expect UAS graduates to enter local labor markets is three years later, in 2000.

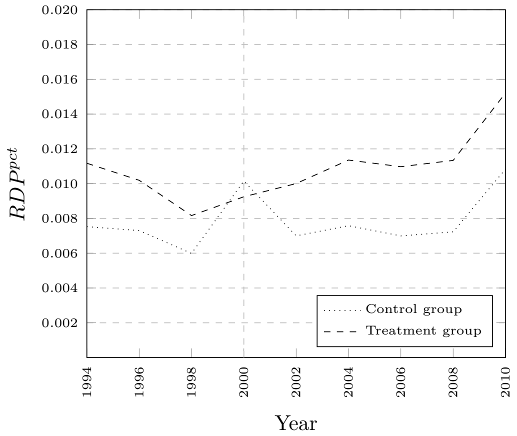
To investigate whether the parallel trends assumption holds in our sample, we plot the trends in the dependent variables for both the treatment and control groups. Before the treatment sets in, the two curves should follow the same time trend, that is, run parallel. First, we plot the raw means. Second, as we condition on the three sampling control variables—industry sector, firm size, and canton—in our DiD model (see section 2.4.1), we plot the predicted means from an OLS regression of the respective dependent variable on the three sampling control variables. For our DiD estimator to identify the causal effect of the introduction of UASs, parallel time trends in the predicted means suffice.

Figure 2.2 shows the time trends of the two dependent variables, percentage of R&D personnel (RDP^{pct}) and percentage of R&D wages (RDW^{pct}), throughout the observed period for the treatment and control groups. A look at the raw means in Figures 2.2a and

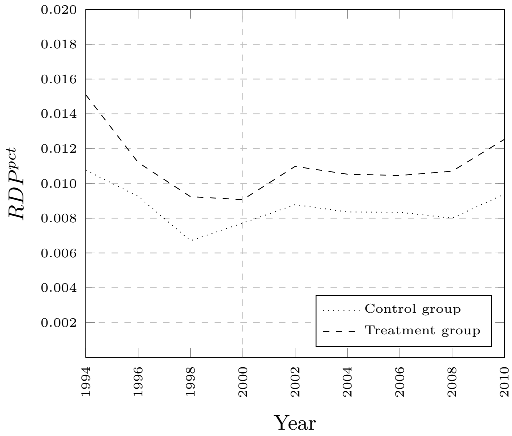
[24] The results in Table 2.2 are robust to increasing or decreasing the treatment radius between 20 and 40 kilometers (see appendix 2.7.6.1).

Figure 2.2: Trends of dependent variables in treatment and control groups

Panel (1): Percentage of R&D personnel (RDP^{pct})

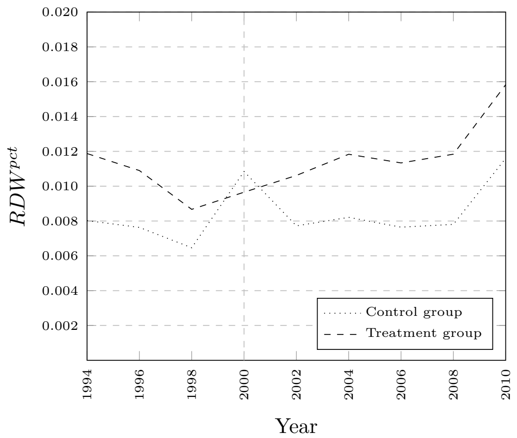


(a) Raw means

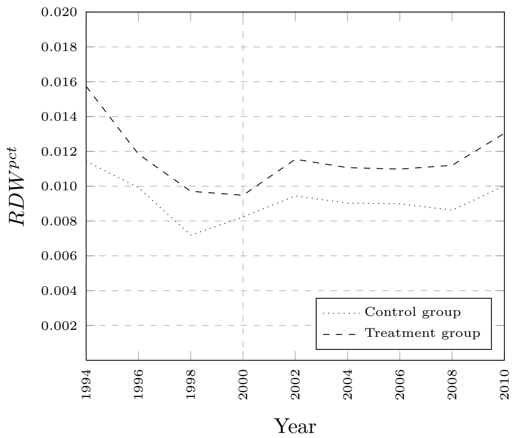


(b) Predicted means

Panel (2): Percentage of R&D wages (RDW^{pct})



(c) Raw means



(d) Predicted means

Source: Authors' illustrations based on ESS data.
Note: Predicted means stem from a linear regression of the respective dependent variable on the three sampling characteristics.

2.2c (i.e., without adjusting the values for differences that arise from changes in sampling between survey years) reveals a declining trend for both the treatment and control groups, with the decline slightly stronger for the treatment group. Furthermore, the peak in the control group in 2000, the first treatment year in our model specification, appears odd at first glance. However, the predicted means in Figures 2.2b and 2.2d show that the pre-treatment trends run parallel and that the peak in the control group vanishes completely. Thus the sampling characteristics fully explain the slight differences in the trends of the treatment and control groups, implying that these differences are indeed unrelated to the treatment.

To further investigate whether the parallel trends assumption holds and whether the effect of the control variables is stable over time, we perform an OLS regression analysis for the pre-treatment period. We regress the dependent variables on the treatment group dummy, the three sampling control variables, and a set of survey year dummies, as well as their interactions with the treatment group dummy and the sampling control variables, as follows:

$$\begin{aligned} Y_{i,t} = & \gamma_0 + \gamma_1 TreatmentGroup_i + \gamma_2 t + \gamma_3 TreatmentGroup_i \times t + \\ & \gamma_4 X_{i,t} + \gamma_5 X_{i,t} \times t + \nu_{i,t} \end{aligned} \tag{2.4}$$

where ν is the error term.

Table 2.1 provides the results of these regressions. These results strengthen our interpretation of the graphs in Figure 2.2—that the pre-treatment time trends do not differ between the treatment and the control groups. None of the estimations show jointly significant interactions between treatment group and survey year dummies, even when we do not include the sampling control variables. Moreover, while the set of sampling control variables is jointly significant in all estimations, their interactions with the survey year dummies are not. Thus the effect of the control variables is stable over time. These results clearly show that assuming parallel pre-treatment trends for the two dependent variables, RDP^{pct} and RDW^{pct}, is reasonable.

In addition to assessing the parallel trends assumption, the empirical evidence on the

Table 2.1: Pre-treatment trends tests

	RDP^{pct}			RDW^{pct}		
	(1)	(2)	(3)	(4)	(5)	(6)
$TreatmentGroup_i$	0.365**	-0.023	-0.504	0.385**	-0.018	-0.479
	(0.159)	(0.161)	(0.615)	(0.162)	(0.164)	(0.620)
Year (t): 1996	-0.023	0.072	-1.049	-0.039	0.055	-0.974
	(0.198)	(0.201)	(0.646)	(0.199)	(0.202)	(0.657)
Year (t): 1998	-0.154	0.166	-2.184**	-0.156	0.181	-2.209*
	(0.166)	(0.161)	(1.113)	(0.170)	(0.165)	(1.128)
$TreatmentGroup_i$	-0.076	0.143	1.145*	-0.058	0.166	1.087
$\times 1996$	(0.214)	(0.193)	(0.685)	(0.217)	(0.195)	(0.696)
$TreatmentGroup_i$	-0.147	0.024	1.864	-0.164	0.010	1.901
$\times 1998$	(0.207)	(0.197)	(1.189)	(0.212)	(0.201)	(1.207)
Controls	No	Yes***	Yes**	No	Yes***	Yes***
Controls $\times$ Year	No	No	Yes	No	No	Yes
F-test	0.254	0.348	1.798	0.323	0.515	1.687
$TreatmentGroup_i \times t$						
N	28,524	28,524	28,524	28,524	28,524	28,524
R^2	0.001	0.184	0.208	0.001	0.186	0.212

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients and standard errors are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

quasi-randomness of UAS campus introductions in Appendix 2.7.5.2 further supports our identification strategy (see also section 2.2.2). More specifically, economic preconditions and other unobservable factors did not influence the spatial or temporal distribution of UAS campuses. Therefore, we argue that the DiD estimator identifies the causal effect of UAS campus introductions on R&D employment.

2.5 Results

2.5.1 Main Results

Table 2.2 provides the OLS estimation results of the DiD approach specified in Equation 2.3, with coefficients representing percentage point changes. We find that treated establishments employ a significantly higher percentage of workers who perform R&D as their main job activity after the opening of a UAS STEM campus (columns 1 and 2). Given the sample mean of 1.02 percent, this treatment effect of 0.16 percentage points is also economically significant. The sampling control variables explain the time-invariant differences between the treatment and control groups, but do not substantially change the treatment effect.

Furthermore, we find that treated establishments spend a significantly larger percentage of their total wage sum on R&D personnel after the introduction of UASs (table 2.2, columns 3 and 4). Given the magnitude of 0.14 percentage points and the sample mean of 1.07, this wage effect is statistically significant at the five percent level, as well as economically significant. Again, conditioning on the sampling control variables leads to the treatment group dummy turning insignificant, while the treatment effect even slightly increases.

A look at the individual employee level reveals that the increase in the employment of R&D personnel does not coincide with a decrease in the average wage of an individual R&D employee. After the introduction of UASs, the average wage of an individual R&D worker employed at the establishments in our estimation sample increases steadily in both the treatment and control groups, with a larger increase in the treatment group (see

Table 2.2: Main estimation results

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.166***	0.158***	0.143**	0.145**
	(0.060)	(0.055)	(0.061)	(0.056)
$TreatmentGroup_i$	0.227***	0.027	0.233***	0.035
	(0.055)	(0.054)	(0.056)	(0.056)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.016	1.016	1.073	1.073
N	232,228	232,228	232,228	232,228
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

appendix 2.7.2). This descriptive evidence suggests that firms do not employ more R&D personnel in the form of UAS graduates as cheap replacements for UNI graduates, because in such a case the average wage of an individual R&D employee would have decreased.

Our estimation results indicate that firms' employing more R&D personnel is one channel through which tertiary education institutions influence innovation. Firms that experience an education-driven labor supply shock after the openings of UAS STEM campuses yield a significantly higher R&D intensity as measured by the percentage of R&D personnel and the percentage of wages paid to R&D personnel.

2.5.2 Robustness Checks

To test the robustness of our DiD estimation results, we estimate a large number of alternative specifications of our model. In this section, we briefly summarize these robustness checks. We provide the detailed analyses, including the corresponding regression tables, in Appendices 2.7.5.2 and 2.7.6.1.

We perform two analyses that account for unobserved characteristics, thereby helping us detect any violations of the parallel trends assumption that our assessments in Section 2.4.3 did not. First, in our empirical assessment of the quasi-randomness of UAS campus introductions in Appendix 2.7.5.2, we consider unobservable campus region-specific characteristics (e.g., level and growth of the economy) by including campus region-specific FE and trends in our estimations. In addition to determining UAS campus locations, such characteristics could affect the outcome of the treatment. We find that the treatment effects of our main specification persist even when we include unobservable campus region-specific characteristics. This finding represents strong evidence for the quasi-randomness of UAS campus introductions and for the treatment effect's independence from campus region-specific characteristics.

Second, to account for unobserved time-invariant establishment characteristics, we perform an FE estimation in Appendix 2.7.6.1. As we cannot track firms over time in the ESS data due to its repeated cross-sections, we group the establishment-level observations according to observable characteristics (MS region, industry sector, and firm size) to obtain an unbalanced quasi-panel. Using these establishment groups as our unit of observation, we perform a quasi-panel FE estimation. The results of this estimation confirm the treatment effects of our DiD analysis in significance and size, indicating that unobserved time-invariant establishment characteristics do not explain the treatment effect.

Moreover, Appendix 2.7.6.1 shows that our main results are robust to a variety of alternative specifications of the model parameters and the sample composition. We demonstrate that econometrically reasonable decreases or increases of the treatment radius do not alter the main result. To ensure a correct assignment of employees to establishments and thus to test the robustness of our assignment procedure for multi-establishment firms (see section 2.3), we restrict the sample to single-establishment firms and find similar effects. We also show that our results are robust to excluding very large firms (5,000 or more employees) from the sample and to clustering standard errors at the regional (instead of at the firm) level.

In sum, all robustness checks confirm the correct specification of our DiD model. Neither unobservable characteristics nor the model parameters explain our findings. Therefore, we argue that our main results in Section 2.5.1 indicate the causal treatment effect of UAS introductions on R&D employment. While these results show that firms experiencing a supply shock of skilled labor engage more intensively in R&D, we do not yet know precisely which types of firms drive the effects. Therefore, in Section 2.5.3 we further assess the mechanisms underlying this finding.

2.5.3 Effect Heterogeneity

After analyzing the average effect of the introduction of UASs on firms' R&D personnel, we investigate whether this effect is heterogeneous. The finding that, on average, establishments treated by a UAS STEM campus (a) have a larger percentage of employees with R&D as their main job activity and (b) spend a larger percentage of their total wage sum on R&D employees either could result from firms just starting to engage in R&D (potential start-ups) or from firms engaging more intensively in R&D while having conducted R&D before the introduction of a UAS. Moreover, certain types of firms (e.g., small firms or firms in specific industry sectors) might profit more than other types from the UAS graduates' skills in R&D.

To identify which firms drive our findings, we examine whether the effect of the introduction of UASs is heterogeneous across different types of firms. We do so in three steps. First, by estimating Equation 2.3 with an alternative outcome variable, we investigate whether establishments just starting to engage in R&D determine our findings. Second, we assess whether the effect varies with the size of the firm to which an establishment belongs. Third, we investigate whether the effect differs with regard to a firm's industry sector.

To shed light on the question of whether establishments that start to engage in R&D only after the supply shock of skilled labor contribute to the effects we find, we construct a binary dependent variable RD^{bin} that indicates whether a firm conducts R&D or not. The variable RD^{bin} equals one if a firm has at least one R&D employee (i.e., if $RDP^{pct} > 0$). Then we estimate Equation 2.3 with RD^{bin} as the dependent variable.

Table 2.3 shows the result of this estimation. We find that, after the introduction of UASs, treated establishments have a substantially higher probability of conducting R&D than before. Thus firms recently starting to engage in R&D at least partly explain the treatment effects we find. Unfortunately, as our data comprises repeated cross-sections, we cannot track firms over time and thus cannot definitively determine whether those firms recently starting to engage in R&D are pre-existing firms or start-ups. However, by using employees' tenure as proxies for firm and establishment age to separate existing firms and start-ups, we conduct further analyses and present detailed results in Appendix 2.7.6.2. While the construction of these proxies calls for some caution in interpreting the results across age classes, the analyses still strongly indicate that not only pre-existing firms start engaging in R&D but also start-ups and younger firms (i.e., whose employees have only a very short maximum tenure as proxy for firm and establishment age).

While firms that start to engage in R&D contribute to the main effects we find, these effects might still be heterogeneous across firm-size classes. Therefore, to assess effect heterogeneity, we split our sample into subsamples of establishments belonging to firms

Table 2.3: Estimation results on the probability to conduct R&D

	RD^{bin}	
	(1)	(2)
$Treatment_{i,t-3}$	0.037	0.580***
	(0.168)	(0.158)
$TreatmentGroup_i$	0.549***	0.178
	(0.157)	(0.152)
Year dummies	Yes***	Yes***
Controls	No	Yes***
Sample mean of dep. var.	4.608	4.608
N	232,228	232,228
R^2	0.001	0.157

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

with different firm sizes. To determine a firm's size, we use the self-reported information on the total number of employees[25] (i.e., the total number in all of a firm's establishments taken together) available in the ESS and group firms according to the size classes proposed by the OECD (2015, p. 206). We then estimate Equation 2.3 (at the establishment level) separately for each subsample, with RDP^{pct}, RDW^{pct}, and RD^{bin} as the dependent variables.

Table 2.4 provides the estimated treatment effects for the different firm-size classes. These results suggest that establishments belonging to very small firms (from 5–9 employees) and those belonging to very large firms (5,000 or more employees) profit most from the introduction of UASs. For very small firms, the increase in R&D-specific labor supply might have facilitated the recruitment of R&D personnel. Moreover, for UAS graduates, a job position at a very small firm might be a suitable start for a career in R&D. Slightly larger firms (10–20 employees) intensify their engagement in R&D due to the improved recruitment options in the form of UAS graduates. The marginally significant negative effect on the percentage of R&D wages indicates that medium-sized firms (250–499 employees) might replace their previous R&D personnel with UAS graduates, because these firms might have a competitive advantage in the local labor markets but otherwise do not change their engagement in R&D. In contrast, larger firms (1,000–4,999 employees) appear to establish R&D in regions where they were not previously active. Very large firms[26] (5,000 or more employees) also open new R&D establishments and intensify their engagement in R&D as well.

These results are robust to alternative specifications of firm-size classes with more equal class sizes. Appendix 2.7.6.2 shows the results for dividing firms into quintiles and deciles of the firm-size distribution. While these specifications do not allow the assessment of very large firms, the results are structurally very similar, with small firms experiencing the strongest effects.

[25] The self-reported firm size contains the total number of employees at the time of the survey. In comparison, the firm-size category we use as a control variable (see section 2.4.1) is a stratification criterion in the ESS and contains information from the BUR. In the effect heterogeneity analyses, we still control for the firm-size category from the BUR.

[26] The 3,027 establishments in the subsample of firms with 5,000 or more employees belong to 128 firms. The main results in Table 2.2 are robust to excluding these establishments (see appendix 2.7.6.1).

Table 2.4: Estimation results by firm size class

	Firm size (S_i)	N	RDP^{pct} (1)	RDW^{pct} (2)	RD^{bin} (3)
	$S_i \leq 4$	53,029	0.076	0.064	0.181
			(0.122)	(0.124)	(0.170)
	$5 \leq S_i \leq 9$	35,134	0.336**	0.345**	0.615**
			(0.138)	(0.142)	(0.292)
	$10 \leq S_i \leq 19$	29,300	0.355*	0.300	0.592
			(0.183)	(0.188)	(0.450)
	$20 \leq S_i \leq 49$	36,799	0.253	0.226	0.548
			(0.156)	(0.160)	(0.460)
	$50 \leq S_i \leq 99$	22,644	0.287*	0.259	0.515
			(0.165)	(0.172)	(0.602)
$Treatment_{i,t-3}$	$100 \leq S_i \leq 249$	22,190	-0.125	-0.136	-0.216
			(0.173)	(0.176)	(0.586)
	$250 \leq S_i \leq 499$	12,062	-0.319	-0.356*	-0.910
			(0.202)	(0.213)	(0.743)
	$500 \leq S_i \leq 999$	8,348	-0.100	-0.108	-0.147
			(0.235)	(0.235)	(0.894)
	$1,000 \leq S_i \leq 4,999$	9,695	0.159	0.142	1.120*
			(0.193)	(0.197)	(0.677)
	$5,000 \leq S_i$	3,027	0.523***	0.490**	3.928***
			(0.194)	(0.189)	(1.302)

Source: Authors' calculations based on ESS data.
Notes: Results from separate regressions. Robust standard errors clustered at the firm level in parentheses. All models include intercept, treatment group dummy, year dummies, and sampling control variables. Coefficients and standard errors are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Another source for heterogeneous effects might be that, as one would expect, only firms in industry sectors related to STEM profit from the openings of UAS STEM campuses. To examine effect heterogeneity across industry sectors, we perform subsample estimations for establishments belonging to firms in different industry sectors. We use the main categories of the NOGA 2002 (see also section 2.4.1) as the industry sectors. Then, as we did for the subsamples of different firm-size classes, we estimate Equation 2.3 (at the establishment level) for each industry-sector subsample, with RDP^{pct}, RDW^{pct}, and RD^{bin} as the dependent variables.

Our assessments of effect heterogeneity across industry sectors, shown in Table 2.5, indicate that establishments belonging to firms in the "manufacture of goods" and "other community, social, and personal service activities" sectors drive our main results. Establishments belonging to firms in the "manufacture of goods" sector started to engage in R&D and intensified their engagement in R&D after the introduction of UASs. We find similar effects for the "other community, social, and personal service activities" sector, which comprises, among other things, television and entertainment companies. As expected, establishments belonging to firms in sectors unrelated to STEM remain unaffected when treated by a UAS STEM campus or its graduates.[27] One exception is establishments in the "hotels and restaurants" sector, which has a slightly higher probability of conducting R&D, possibly because the availability of UAS graduates enabled these establishments to perform R&D in the first place.

In sum, our assessments of effect heterogeneity suggest that establishments belonging to firms with five to nine employees and those belonging to firms with 5,000 employees or more, as well as establishments belonging to firms in the "manufacture of goods" and "other community, social, and personal service activities" sectors drive our main results. Furthermore, firms' starting to engage in R&D contributes to these results. However, as the SFSO changed the sampling procedure for the ESS with the 2002 survey, and as firms in some of the subsamples might be underrepresented before 2002, our assessments of

[27] When examining UAS STEM campuses, we do not expect any treatment effects for industry sectors not involved in STEM activities. For an analysis of the effect of UASs on establishments in sectors such as "education" or "health, veterinary, and social work," campuses in other fields would have to be considered. Our main results in Table 2.2 are robust to excluding these establishments.

Table 2.5: Estimation results by industry sector

	Industry sector ($NOGA2002_i$)		N	RDP^{pct} (1)	RDW^{pct} (2)	RD^{bin} (3)
$Treatment_{i,t-3}$	A	Agriculture and forestry	5,106	-0.164 (0.150)	-0.172 (0.158)	-0.063 (0.213)
	C	Mining and quarrying	824	0.070 (0.370)	-0.098 (0.383)	0.995 (2.590)
	D	Manufacture of goods	53,852	0.366^{**} (0.156)	0.352^{**} (0.163)	1.519^{***} (0.532)
	E	Electricity, gas, and water supply	1,240	0.440 (0.270)	0.439 (0.267)	-0.759 (1.878)
	F	Construction	12,965	0.007 (0.021)	0.005 (0.025)	-0.345 (0.272)
	G	Wholesale and retail trade	45,028	-0.001 (0.045)	0.002 (0.045)	0.085 (0.199)
	H	Hotels and restaurants	9,626	0.004 (0.003)	0.005 (0.003)	0.113^{*} (0.064)
	I	Transport, storage, and communication	13,879	-0.058 (0.043)	-0.062 (0.047)	-0.546 (0.451)
	J	Financial interme- diation; insurance	16,158	0.035 (0.064)	0.031 (0.062)	0.286 (0.267)
	K	Real estate, renting; other business activities	36,779	0.174 (0.241)	0.120 (0.244)	0.238 (0.445)
	M	Education	6,164	-0.127 (0.192)	-0.178 (0.206)	-0.149 (0.669)
	N	Health, veterinary, and social work	13,295	-0.167 (0.126)	-0.174 (0.130)	-0.599 (0.439)
	O	Other community, social, and personal service activities	17,312	0.540^{**} (0.214)	0.540^{**} (0.215)	1.171^{***} (0.404)

Source: Authors' calculations based on ESS data.
Notes: Results from separate regressions. Robust standard errors clustered at the firm level in parentheses. All models include intercept, treatment group dummy, year dummies, and sampling control variables. Coefficients and standard errors are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Nomenclature of NOGA 2002 categories provided by SFSO (2002), titles shortened for illustration (original titles provided in appendix 2.7.3).

effect heterogeneity should be interpreted with some caution. Nonetheless, they contribute to a better understanding of where the overall treatment effect might originate.

2.5.4 Variation of Treatment Effects Over Time

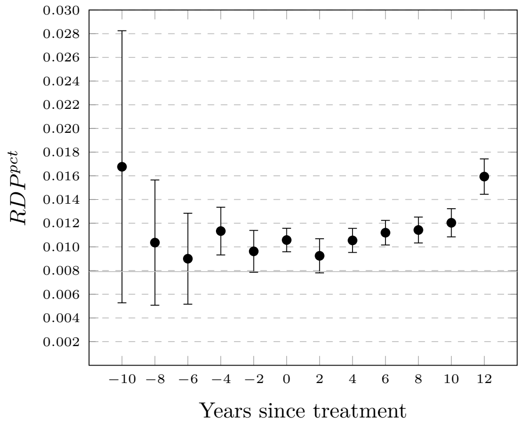
To analyze whether the treatment effect remains stable over time, we estimate an alternative specification of our DiD model considering the years since treatment (y). In so doing, we follow Autor's (2003) standard procedure and now include y as an indicator for the temporal distance of an observation to the treatment year (instead of including a binary treatment variable), with the control group observations serving as the reference. As the ESS has a biennial structure, we set y to the next lower value if y is an uneven number. For example, for firms affected by the opening of the Rapperswil campus in 2001, $y = 0$ for firms observed in 2002, $y = 2$ for firms observed in 2004, etc.

Figure 2.3 shows the treatment effects over time. Figures 2.3a and 2.3c plot the raw means and the corresponding 95-percent confidence intervals of the dependent variables percentage of R&D personnel and percentage of R&D wages, respectively. In the pre-treatment period, the number of observations in the data is considerably lower, leading to larger confidence intervals. Although, as in Figure 2.2, the variable means suggest a declining pre-treatment trend, the confidence intervals include the control group mean (with the exception of $y = -4$), indicating that this declining trend is not statistically significant. The post-treatment period shows a clear upward trend in the dependent variables, indicating an increase of the treatment effect over time.

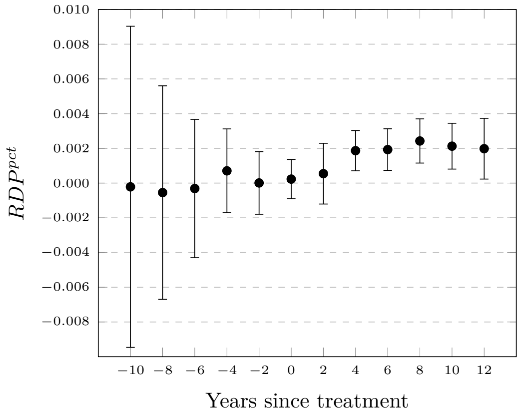
Figures 2.3b and 2.3d show that none of the pre-treatment coefficients is statistically different from zero, whereas after at least four years, all coefficients are positive and statistically significant at the one percent level. The panels plot the coefficients and their 95-percent confidence intervals from OLS regressions of the respective dependent variable on years since treatment, survey year dummies, and sampling control variables (appendix 2.7.4 includes the detailed regression results). The results of these regressions also confirm our analyses in Section 2.4.3, suggesting parallel pre-treatment trends conditional on the sampling control variables. Furthermore, the treatment effect persists over time.

Figure 2.3: Variation of treatment effects over time

Panel (1): Percentage of R&D personnel (RDP^{pct})

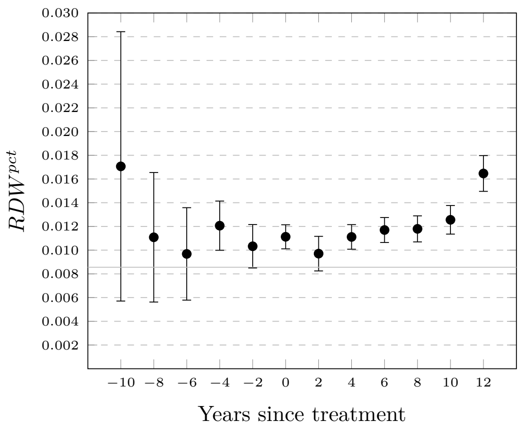


(a) Raw means

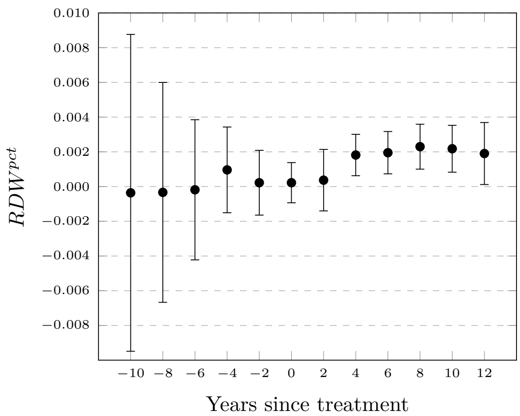


(b) Estimation coefficients

Panel (2): Percentage of R&D wages (RDW^{pct})



(c) Raw means



(d) Estimation coefficients

Source: Authors' illustrations based on ESS data.
Notes: Estimation coefficients stem from the regressions in Table 2.10. In addition to raw means or estimation coefficients, graphs show the respective 95-percent confidence intervals. The solid horizontal lines in panels (a) and (c) show the raw mean for the control group as the reference.

2.6 Conclusion

In this chapter, we investigate how an exogenous increase in skilled labor supply resulting from the introduction of UASs in Switzerland influences the employment of R&D personnel in treated firms, in terms of both employment and wages paid to R&D personnel. Firms located near a UAS STEM campus experience a supply shock of skilled labor in the form of UAS STEM graduates entering the regional labor markets.

Applying a DiD design that exploits the quasi-random variation in the location and the timing of UAS campus openings, we find that the percentage of R&D personnel relative to total personnel (i.e., the percentage of employees with R&D as their main job activity) in treated establishments increases significantly in comparison to untreated establishments. Moreover, we find that treated establishments spend a higher percentage of their total wage sum for R&D personnel, measured as the percentage of R&D wages (the sum of wages paid to R&D personnel relative to the total wage sum).

We conduct further analyses to examine which types of firms are responsible for the effects we find. First, these analyses suggest that firms' just starting to engage in R&D is one determinant of these effects. Second, both very small firms (particularly those with five to nine employees and thus possibly start-ups) and very large firms (with 5,000 or more employees) profit from the introduction of UASs. Third, firms in the "manufacture of goods" and "other community, social, and personal service activities" sectors conduct more R&D, thereby positively contributing to the overall effects.

Our findings suggest that the policy reform of introducing UASs in the German-speaking part of Switzerland provided firms with R&D personnel, thereby laying the foundation for increasing innovation activities of treated firms. The easier availability of labor with R&D skills to firms in the treated regions and the resulting incentives for a stronger engagement of those firms in R&D foster the employment of—and increase the budget spent on—R&D personnel, thereby providing one potential channel for an increase in innovation activities in the treated regions. Therefore, firms' using the R&D-specific skill resources of UAS graduates for their R&D purposes constitutes one of the underlying mechanisms leading to the increase in patenting activities that Pfister et al. (2018) found.

Our results provide important insights for policymakers designing reforms aimed at increasing firms' innovation activities through a tertiary education expansion (along the lines of UASs in Switzerland). The particular combination of sound professional knowledge and applied research skills that STEM UASs provide is an effective means of stimulating R&D even in firms that may not have yet conducted it. With UAS graduates possessing skills tailored specifically for the R&D requirements of both small firms and firms starting to engage in R&D for the first time, a tertiary education expansion leading to the availability of such skills in regional labor markets both promotes and increases firms' innovation activities.

2.7 Appendix

2.7.1 Descriptive Statistics for Dependent Variables

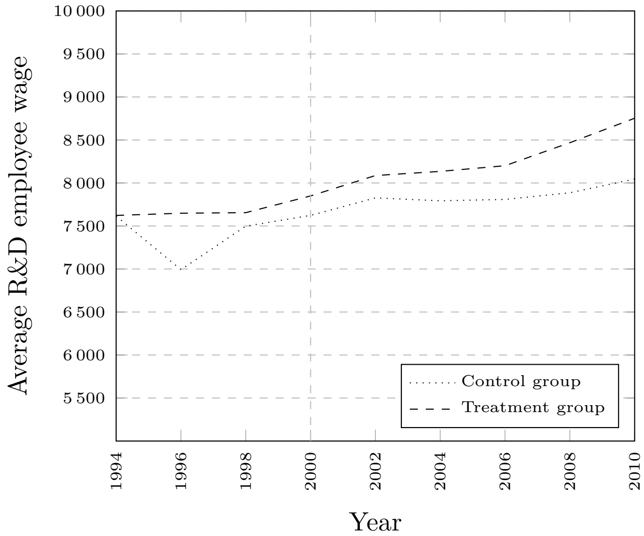
Table 2.6: Descriptive statistics for percentage of R&D personnel (RDP^{pct})

	Control group			Treatment group		
Year	N	Mean	SD	N	Mean	SD
1994	3,633	0.753	6.597	6,736	1.117	8.083
1996	2,992	0.730	6.371	5,694	1.019	7.666
1998	3,433	0.599	5.068	6,036	0.816	6.429
2000	3,362	1.015	7.629	6,062	0.925	7.334
2002	14,155	0.700	5.489	22,710	1.001	7.286
2004	14,300	0.758	6.114	24,471	1.136	8.066
2006	14,415	0.699	5.690	24,790	1.097	7.775
2008	14,606	0.723	5.921	24,505	1.134	7.961
2010	14,815	1.084	7.643	25,513	1.521	9.472
Total	85,711	0.791	6.271	146,517	1.148	8.049

Source: Authors' calculations based on ESS data.
Note: Variable values multiplied by 100 to represent percentages.

2.7.2 R&D Employee Wage at the Individual Level

Figure 2.4: Development of average R&D employee wage at the individual level (in Swiss Francs)



Source: Authors' illustrations based on ESS data. Calculations at the individual employee level.

Table 2.7: Descriptive statistics for percentage of R&D wages (RDW^{pct})

	Control group			Treatment group		
Year	N	Mean	SD	N	Mean	SD
1994	3,633	0.802	6.663	6,736	1.187	8.303
1996	2,992	0.763	6.396	5,694	1.090	7.898
1998	3,433	0.646	5.302	6,036	0.867	6.617
2000	3,362	1.084	7.842	6,062	0.966	7.431
2002	14,155	0.772	5.805	22,710	1.062	7.472
2004	14,300	0.821	6.330	24,471	1.183	8.159
2006	14,415	0.765	5.944	24,790	1.134	7.826
2008	14,606	0.781	6.147	24,505	1.184	8.059
2010	14,815	1.160	7.886	25,513	1.583	9.602
Total	85,711	0.856	6.500	146,517	1.200	8.171

Source: Authors' calculations based on ESS data.
Note: Variable values multiplied by 100 to represent percentages.

Table 2.8: Descriptive statistics for probability to conduct R&D (RD^{bin})

	Control group			Treatment group		
Year	N	Mean	SD	N	Mean	SD
1994	3,633	3.799	19.119	6,736	4.765	21.305
1996	2,992	3.643	18.739	5,694	4.724	21.218
1998	3,433	3.292	17.844	6,036	4.092	19.812
2000	3,362	4.610	20.974	6,062	4.223	20.113
2002	14,155	4.076	19.775	22,710	4.716	21.199
2004	14,300	4.231	20.130	24,471	4.704	21.172
2006	14,415	4.176	20.005	24,790	4.627	21.007
2008	14,606	4.005	19.609	24,505	4.705	21.175
2010	14,815	5.089	21.979	25,513	5.679	23.145
Total	85,711	4.244	20.160	146,517	4.821	21.422

Source: Authors' calculations based on ESS data.
Note: Variable values multiplied by 100 to represent percentages.

2.7.3 NOGA 2002 Categories of Industry Sector

Table 2.9: NOGA 2002 categories of industry sector

Category	Title
A	Agriculture and forestry
B	Fishing and fish farming
C	Mining and quarrying
D	Manufacture of goods
E	Electricity, gas, and water supply
F	Construction
G	Wholesale and retail trade; repair of motor vehicles and consumer durables
H	Hotels and restaurants
I	Transport, storage, and communication
J	Financial intermediation; insurance (excluding compulsory social security)
K	Real estate, renting, and related activities; other business activities
L	Public administration and defence; compulsory social security
M	Education
N	Health, veterinary, and social work
O	Other community, social, and personal service activities
P	Private household
Q	Extra-territorial organizations and bodies

Source: SFSO (2002)
Notes: Categories B, P, and Q not included in the ESS sample we use for our estimations. Category L excluded because it does not contain private firms.

2.7.4 Variation of Treatment Effects Over Time

Table 2.10: Estimation results over time

		RDP^{pct}		RDW^{pct}	
		(1)	(2)	(3)	(4)
Years since treatment	-10	0.953	-0.022	0.934	-0.036
		(0.595)	(0.472)	(0.589)	(0.466)
	-8	0.302	-0.055	0.327	-0.034
		(0.322)	(0.314)	(0.330)	(0.323)
	-6	0.202	-0.031	0.222	-0.019
		(0.216)	(0.203)	(0.219)	(0.206)
	-4	0.414***	0.071	0.438***	0.096
		(0.135)	(0.123)	(0.138)	(0.126)
	-2	0.233**	0.001	0.253**	0.022
		(0.100)	(0.092)	(0.104)	(0.095)
	0	0.295***	0.023	0.287***	0.022
		(0.058)	(0.058)	(0.060)	(0.059)
	2	0.134	0.054	0.117	0.037
		(0.094)	(0.089)	(0.095)	(0.090)
	4	0.345***	0.187***	0.333***	0.182***
		(0.064)	(0.059)	(0.066)	(0.061)
	6	0.337***	0.193***	0.327***	0.195***
		(0.063)	(0.061)	(0.064)	(0.062)
	8	0.379***	0.243***	0.354***	0.230***
		(0.067)	(0.065)	(0.068)	(0.066)
	10	0.419***	0.213***	0.408***	0.218***
		(0.072)	(0.067)	(0.073)	(0.069)
	12	0.528***	0.198**	0.499***	0.190**
		(0.094)	(0.089)	(0.096)	(0.091)
Year dummies		Yes***	Yes***	Yes***	Yes***
Controls		No	Yes***	No	Yes***
Sample mean of dep. var.		1.016	1.016	1.073	1.073
N		232,228	232,228	232,228	232,228
R^2		0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

2.7.5 The Introduction of UASs in Switzerland as a Quasi-Random Process: Institutional Background, Governmental Requirements and Political Process, Location and Timing Decisions, and Empirical Evidence

In this appendix, we provide qualitative and quantitative empirical evidence on the quasi-random process of the introduction of UASs in Switzerland. In the first part, we describe the institutional background of the introductions of UASs to show that the 1995 UAS law (*Fachhochschulgesetz*) and the subsequent introductions of UASs constituted a fundamental change in the Swiss tertiary education sector. This change resulted in a new type of tertiary education institution, one considered equivalent to—yet different from—("gleichwertig aber andersartig") a UNI. Relying on Pfister (pp. 12–22 2017), we provide qualitative descriptions of the process that led to the quasi-random location and timing of the introductions of UAS campuses specializing in STEM fields across different regions in Switzerland. We focus on the political system, the institutional actors involved, and their relationship to one another.

In the second part, to support our qualitative argumentation on the quasi-randomness of the location and timing of UAS introductions with quantitative empirical evidence, we provide ample statistical analyses comparing regions with and without UASs. For these analyses, we use both data from the ESS and other data.[28]

Taken together, the qualitative and quantitative empirical evidence strongly support our identifying assumption that the variation in the location and timing of UAS campus openings can be considered quasi-random. Thus this variation allows a causal interpretation of the estimation results in this chapter.

[28] We thank the SFSO for data provision of the Swiss Earning Structure Survey, the Business Census, and the Survey of Higher Education Graduates.

2.7.5.1 Institutional Background of the Swiss Higher Education Sector, the 1995 UAS Law, and the Subsequent Introductions of UASs Across Regions and Time

Institutional Actors. The Swiss federal government was the institutional actor defining the legal basis—and thus setting the legal mandates—of the UASs.[29] Most importantly, these mandates required the UASs to recruit apprenticeship graduates as their students (as opposed to academic universities, which recruit high school graduates as their students), thereby providing educational careers for VET graduates. Moreover, the mandates required UASs to teach their students applied R&D, to conduct applied R&D, and to provide services to and collaborate with public and private sector firms.

In imposing these legal mandates, the federal government aimed at fundamentally restructuring Switzerland's higher education sector. Before the reform, Switzerland's tertiary level consisted of traditional UNIs (tertiary A-type in the Swiss nomenclature) and professional education institutions (tertiary B-type, i.e., not university level), particularly Professional Education and Training (PET) colleges.[30] The 1995 UAS law introduced a second tertiary A-type institution, the UASs. Although some PET colleges were transformed into UASs, the legal mandates—with the consequent funding for conducting applied R&D—applied to all UASs, and only to them. Therefore, the UASs differ substantially from both UNIs and PET colleges.[31]

In addition to legally setting the mandates for the UASs, the federal government was the institutional actor with the power to confer accreditation. It did so at three hierarchical levels: in ascending order, UAS study programs, UAS campuses (locations at which study programs are offered), and UASs (entire institutions consisting of one or more UAS campuses). The Swiss UASs—much like the University of California, with its

[29] For further information on the reform and the introduction process, see Bundesgesetz über die Fachhochschulen (Fachhochschulgesetz) (1995).

[30] PET colleges—ISCED level 5B institutions offering education in business administration, STEM, and design—allow individuals with a VET degree to continue their educational career at the tertiary level, thereby enhancing technical and managerial expertise in their occupational fields. Nonetheless, they are college-level, not university-level, institutions.

[31] By exploiting exogenous variation in the creation of new education institutions that rely on predecessor institutions, we proceed similarly to Kamhöfer et al. (2019), who exploit the introduction of UASs in Germany, and Eyles and Machin (2019), who exploit the conversion of existing state schools into academy schools in England.

more than 10 campuses across the state (e.g., UCLA in the south and UC Berkeley in the north)—are composed of a system of campuses distributed across different regions of Switzerland. In conferring accreditation, the federal government required not only the fulfillment of legal mandates, but also a sufficiently large size and a solid financial base for each campus. Furthermore, to ensure easy access for all VET graduates to UASs, the federal government required and equalized the geographical distribution of UASs and their campuses throughout Switzerland.

The 26 Swiss cantons (which function similarly to U.S. states but tend to be much smaller in geographical size) were the institutional actors responsible for implementing the federal law and bearing the main financial burden of the UASs and their campuses. Due to Switzerland's federalist political system, the cantons[32] have a high degree of autonomy in policy-making. However, the legal requirements that the federal government imposed—particularly the equal distribution of campuses across Switzerland—substantially restricted cantonal discretionary leeway. The cantons thus had to establish inter-cantonal coordination and collaboration, provoking a heated inter-cantonal debate (and one between the cantons and the federal government).[33]

The quasi-randomness of UAS introductions thus resulted from Switzerland's federalist political system and from the geographical distribution requirements of the federal government. As outlined in the following, the political system thus required complex negotiations and coalition building with (constantly changing) institutional actors, including, for example, package deals related to historical coincidences and personalities, but unrelated to local economic factors.

[32] Switzerland has a federalist political system to accommodate cultural diversity—because the country consists of different religious and linguistic regions—while creating national unity. The different political levels (federal, cantonal, and municipal) all have legislative and executive institutions. In addition, while Swiss federalism gives cantons autonomy in policy-making in all areas not explicitly regulated by the federal constitution (e.g., education), this high degree of autonomy requires extensive coordination among cantons. For further information on Switzerland's political system, see, e.g, https://www.eda.admin.ch/aboutswitzerland/en/home/politik/uebersicht/foederalismus.html (last retrieved on November 28, 2019).

[33] For example, the citizens of the canton of Geneva, an economically very strong canton, voted down a popular petition requesting a UAS for their canton. In contrast, Basel (consisting of the two cantons Baselland and Baselstadt), submitted two requests proposing more than one UAS for this northwestern part of Switzerland to prevent a "destructive competition" among potential UAS campuses.

Location and Timing Decisions. The institutional actors involved in the UAS reform engaged in a complex process that determined the location and timing of UAS campus introductions. In total, 15 UAS STEM campuses existed throughout the observation period of this chapter (1994–2010). Table 2.11 summarizes the locations and years of introduction of these campuses.

After enacting the UAS law, the federal government started an open call for applications, in which institutions (including PET colleges) could apply for UAS accreditation. To guarantee a legitimate application and introduction process with equal opportunities for all applicants and to evaluate the applications, the federal government created the UAS commission (*Eidgenössische Fachhochschulkommission*).[34] In 1996, the federal government received 10 applications for UASs, all of which would include several campuses. One year later, this number had increased to 14 UASs because some cantons[35] could not establish inter-cantonal collaborations and therefore decided to submit their own proposals.

The federal government reacted to this first round of applications by restricting to seven the allowed maximum number of UASs. The cantons, which at first opposed this decision and started angry negotiations on the number and the locations of UASs, eventually had to collaborate with other cantons and reach agreements with them. Nonetheless, disagreements among the cantons (e.g., on financial distributions and the geographical borders of a UAS region) disrupted negotiations and led to new collaboration agreements, and constantly changing applications. Moreover, the cantons promoted their own potential UAS campuses, resulting in heated political debates between—and even within—cantons. Therefore, the UAS commission and the federal government repeatedly admonished the applicants to overcome their provincialism and to start cooperating across cantons.

For example, in 1996 eight cantons in eastern Switzerland applied for accreditation of a single "University of Applied Sciences Zurich-East". As these cantons could not agree on the distribution of power, the number of UAS applications in eastern Switzerland increased to four. The federal UAS commission rejected the four applications and suggested

[34] In addition, the UAS commission controlled the implementation of the UAS introduction, that is, the commission decided on additional requirements for UASs and UAS campuses having received conditional accreditation.

[35] All Swiss cantons are accustomed to a high level of political autonomy in educational matters.

Table 2.11: UAS campus locations and years of introduction

UAS	Year of introduction	Location of campus
Bern University of Applied Sciences	1997–2003	Bern
	1997	Biel
	1997	Burgdorf
University of Applied Sciences of Central Switzerland	1997	Horw
University of Applied Sciences of Eastern Switzerland	2001	Buchs
	2000	Chur
	2001	Rapperswil
	2000	St. Gallen
University of Applied Sciences of Northwestern Switzerland	1998	Brugg-Windisch
	1997	Muttenz
	1998–2003	Oensingen
	2003–2006	Olten
University of Applied Sciences of Zurich	1998	Wädenswil
	1998	Winterthur
	1998	Zurich

Source: Illustration based on Pfister et al. (2018).

establishing a "University of Applied Sciences East" (consisting of campuses in the cantons of Zurich and St. Gallen) and a "University of Applied Sciences South-East" (consisting of campuses in the canton of Grisons and in the principality of Liechtenstein). The federal government, however, accepted two UASs—one in Zurich and one in eastern Switzerland—with a geographical distribution different from all previous applications. In response to this decision, the eastern cantons started to plan the "University of Applied Sciences of Eastern Switzerland", with STEM campuses in Buchs, Chur, Rapperswil, and St. Gallen. The canton of Zurich created the "University of Applied Sciences of Zurich", with STEM campuses in Wädenswil, Winterthur, and Zurich.

Finally, the federal government gave conditional accreditation to seven UASs (and their campuses): five in the German-speaking part of Switzerland, one in the French-speaking part, and one in the Italian-speaking part. However, to guarantee that the campuses were equally distributed across the country, sufficiently large, and financially solid, the federal government required the merger of campuses and study programs within UASs regions. For example, after having received conditional accreditation for three STEM campuses in the cities of Bern, Biel, and Burgdorf, the federal government forced the "Bern University of Applied Sciences" to combine three campuses into two.[36] The one in Bern—located between the campuses in Biel and Burgdorf—had to close its doors. Local Bern politicians, firm representatives, and campus administrators resisted to these mergers, emphasizing the economic importance of the city of Bern and the surrounding region, but with no success.[37] Consequently, the engineering and IT departments left Bern, while the arts, social work, and business administration departments remained. As a result of the interactions and conflicts among the federal, cantonal, and regional actors, the spatial and temporal distribution of UAS campuses was clearly independent of regional economic

36 Similar to the application process in eastern Switzerland, the application process in the region of Bern included different plans for several UASs whose geographical distribution, selection of STEM campuses, and cantons involved in bearing the financial burden varied substantially. For example, a potential UAS labelled "BEJUNE" was planned for the region involving the cantons of Bern, Jura, and Neuchâtel. Likewise, the role of the STEM campus in St. Imier for a potential "Bern University of Applied Sciences"—and the consequences for the Biel STEM campus located nearby—was unclear for a long time.

37 The representatives argued that the city of Bern produced more than half of the cantonal GDP and that the city accommodated 72 percent of the medical and 85 percent of the telematics industries.

strength, innovation strength, or any other factors relevant for the analyses in this chapter.

Similar consolidation processes took place in the "University of Applied Sciences of Northwestern Switzerland" region. Although granting conditional accreditation to three STEM campuses in Brugg-Windisch, Muttenz, and Oensingen in 1997 and 1998, the federal government required the merger of STEM campuses, allowing only one UAS in northwestern Switzerland. Therefore, even though the four affected cantons of Aargau, Baselland, Baselstadt, and Solothurn—all of which had applied for this one UAS in northwestern Switzerland—had to start collaborating, each continued to form political alliances to promote the location of campuses within its borders. The western part of Aargau planned a campus in Brugg-Windisch with study programs in engineering, IT, chemistry, and life sciences; the eastern part of Aargau collaborated with the canton of Solothurn and planned a UAS region geographically restricted to the cities of Aarau and Olten. The Oensingen campus with study programs in engineering and IT therefore relocated to Olten in 2003 to support the Aarau-Olten project.

Threatened by the proximity of a potential northwestern UAS in Aarau-Olten, Basel argued for a UAS applying for two campuses (one specializing in IT and one in the arts). While giving conditional consent for the arts campus, the federal government rejected Basel's application for an IT campus. However, the decision for the arts campus in Basel—located closer to Aarau and Olten than to Brugg-Windisch—and the heated political debate within the canton of Aargau led to the abandonment of the Aarau-Olten UAS project. In addition, both the federal government and the UAS commission repeatedly admonished the cantons for their provincialism, finally leading to Aargau, Baselland, Baselstadt, and Solothurn signing an inter-cantonal agreement on the goal of establishing one regional UAS in northwestern Switzerland. In the end, chemistry and life sciences went to Muttenz, and engineering and IT went to Brugg-Windisch. Olten—specializing in social work, business admininstration, and applied psychology—moved its engineering and IT campus to Brugg-Windisch in 2006. Again, the interactions and conflicts among the different actors involved generated a quasi-random distribution of UAS campuses across regions and time.

In addition to these complex processes on campus locations, a variety of factors led to time delays during—and even after—the implementation process. These factors included repeated negotiations between the institutional actors (regarding, e.g., which canton would pay for what), constantly changing coalitions, heated political debates between and within cantons, and the federal government's strict requirement of an equal distribution of UAS campuses. In addition, the geographical distribution of a UAS region—particularly the number and the locations of UAS campuses constituting a UAS region—was ex ante unclear, because the applications constantly changed during the application period and even later.

In sum, the process that led to the introduction of UASs generated a quasi-random distribution of UAS campuses across regions and time. The interactions of the political actors involved and their conflicting interests resulted in UAS campus locations and timing that were independent of regional economic or innovation strength. The next section presents quantitative empirical analyses that compare treated and untreated regions—and the results of which support the qualitative argumentation on the quasi-random introduction in this Appendix 2.7.5.

2.7.5.2 Additional Empirical Evidence on the Quasi-Randomness of UAS Campus Introductions

To additionally support the quasi-randomness assumption and the qualitative argumentation of the previous part of this Appendix 2.7.5, we empirically analyze a number of factors that could potentially prohibit the causal interpretation of our DiD results in this chapter. In other words, we assess those factors that a critical reader could claim might have determined the spatial and temporal distribution of UAS campuses and other factors contradicting our identification. First, we consider UAS campus regions' unobservable regional characteristics (e.g., economic strength). Second, we assess indicators for observable regional characteristics (e.g., labor market conditions, predecessor institutions). Third, we assess the composition of the treatment and control groups, which needs to be stable for the DiD estimation to identify causal effects.

Unobservable Factors. To empirically demonstrate that unobservable factors, such as economic strength or political power, have no impact on the location and timing decisions of UAS campuses within a UAS region, we repeat our main analysis in Section 2.5 with campus region-specific FE. These campus region-specific FE control for the level of any unobservable characteristics within a campus region. While the campus region-specific FE account only for time-invariant unobservable characteristics, we estimate two additional specifications to control for time-variant unobservable characteristics, such as economic growth. First, we include campus region-specific linear trends (i.e., interactions of the campus region-specific FE and the year). Second, we include campus region-specific non-linear trends (i.e., interactions of the campus region-specific FE and year dummies). If unobservable factors have an influence on the treatment effect (and on the location decisions of UAS campuses within UAS regions), the estimation models would yield a substantially lower treatment effect.

The estimation results accounting for unobservable factors in Table 2.12 show that including campus region-specific FE and campus region-specific trends even increases the treatment effect. Therefore, constant, linear, and non-linear unobservable characteristics do not influence the spatial and temporal distribution of UAS campuses. We do not choose the specification accounting for unobservable characteristics as our main specification in this chapter, because given the qualitative and quantitative empirical evidence in this Appendix 2.7.5, the DiD estimator already identifies the causal effect of UAS introductions on R&D employment. By excluding all factors not necessary for a causal identification, we avoid potential threats to the analysis resulting from overidentification.

Economic Preconditions of UAS Campus Regions. After showing that UAS campus region-specific unobservable characteristics did not influence the distribution of UAS campuses, we conduct a more detailed analysis of factors that we can directly or indirectly observe with ESS and other data to further support the quasi-randomness of UAS introductions. First, we compare the labor market strength of those UAS campus regions that relocated to show that these relocations were independent of the regional economy.

Table 2.12: Estimation results with campus region-specific FE and trends

	RDP^{pct}			RDW^{pct}		
	(1)	(2)	(3)	(4)	(5)	(6)
$Treatment_{i,t-3}$	0.208***	0.183***	0.263***	0.200***	0.173***	0.253**
	(0.050)	(0.057)	(0.081)	(0.051)	(0.059)	(0.082)
Campus region-specific FE	Yes***	Yes	Yes**	Yes***	Yes	Yes**
Campus region-specific linear trends	No	Yes	No	No	Yes	No
Campus region-specific non-linear trends	No	No	Yes	No	No	Yes
Year dummies	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Controls	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Sample mean of dep. var.	1.016	1.016	1.016	1.073	1.073	1.073
N	232,228	232,228	232,228	232,228	232,228	232,228
R^2	0.171	0.172	0.172	0.169	0.169	0.170

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Second, to emphasize that UASs were not deliberately introduced into regions with the largest potential effects (i.e., economically strong regions), we repeat our main analysis without the 10 largest cities in the German-speaking part of Switzerland. Third, to assess whether UAS locations and their resulting effects depend on the existence of potential predecessor institutions (i.e., PET colleges), we conduct separate analyses for regions without such predecessor institutions. Fourth, we study whether other economic factors influence R&D employment in UAS campus regions by studying campus closings. Fifth, we analyze whether structural differences between treated and untreated regions determined the distribution of UAS campuses by excluding untreated observations and identifying the treatment effect only from the temporal variation in UAS campus introductions. We describe these five analyses in detail in the following.

First, we assess the labor market strength of the two previously described UASs that experienced campus relocations, the "Bern University of Applied Sciences" and the "University of Applied Sciences of Northwestern Switzerland." We use data from the SFSO's Business Census to compare the labor market strength of the UAS campus regions belonging to these two UASs. This comparison confirms that economic preconditions did not influence the location and timing decisions of UAS campuses. Table 2.13 shows that the relative numbers of employees and firms in the Bern campus region were higher than those in the Biel or Burgdorf campus regions.[38] The location decisions for the UAS campuses belonging to the "Bern University of Applied Sciences" (i.e., the decisions on campus locations within the UAS region) therefore appear unrelated to economic factors and primarily driven by the obligation of an equal distribution of campuses. For the "University of Applied Sciences of Northwestern Switzerland", the campuses in the regions of Olten and Oensingen relocated to the economically slightly weaker region of Brugg-Windisch.[39]

[38] We use relative measures—full-time equivalents per hectare (about 2.47 acres), number of firms per hectare—excluding uninhabitable areas (e.g., lakes, rivers, and mountains, thereby considering only settlement and urban areas according to the Swiss land use statistics) given the Swiss topography. Biel, for example, is located on Lake Biel, whereas the city of Bern does not have a lake. Comparing absolute numbers of full-time equivalents or firms would distort the results, because Lake Biel makes the inhabited Biel region substantially smaller than that of Bern.

[39] We do not include the Muttenz campus region belonging to the "University of Applied Sciences of Northwestern Switzerland" in Table 2.13, because this campus region was not subject to any relocations. The corresponding numbers for the Muttenz campus region are 20.358 full-time equivalent employees per hectare and 1.978 firms per hectare.

Table 2.13: Number of employees and firms for closed UAS campus regions and for those to which campuses relocated

UAS campus region	Full-time equivalent number of employees per hectare (1)	Firms per hectare (2)
Bern University of Applied Sciences		
Campus region Bern (relocated in 2003)	17.237	1.732
Campus region Biel	9.435	1.195
Campus region Burgdorf	6.195	0.924
University of Applied Sciences of Northwestern Switzerland		
Campus region Brugg-Windisch	9.146	1.128
Campus region Oensingen (relocated in 2006)	9.251	1.142
Campus region Olten (relocated in 2003)	9.972	1.180

Source: Authors' calculations based on data from the SFSO's Business Census, waves 1995 and 1998.

Second, to demonstrate that the distribution of UASs and the resulting outcome are independent from regional economic strength and other unobservable factors, we exclude establishments located in one of the nine MS regions that contain the 10 largest cities in the German-speaking part of Switzerland. Then, we re-estimate our model in Section 2.5 with this restricted ESS sample. Through this sample restriction, we can show the relevance of regional economic strength for the location decisions of UASs and UAS campuses. If the institutional actors who decided on UAS locations had assigned UASs only to economically strong regions (where the effects expected a priori are largest), the estimation excluding the 10 largest cities would not yield a treatment effect. However, the corresponding estimation results in Table 2.14 show a positive and significant treatment effect. Therefore, large cities are not more important for the effect of UASs on R&D employment and for UAS location decisions than other treated areas.

Third, we assess whether UAS locations depend on the existence of potential predecessor institutions (i.e., PET colleges) by analyzing the UAS effect in regions without such predecessor institutions. If such predecessor institutions were necessary for UASs to increase R&D employment, and if, consequently, the institutional actors had deliberately assigned UASs only to regions with predecessor institutions, we would not find any treatment effects in regions without predecessor institutions.

Therefore, in this analysis we focus on the UAS campuses in Olten and Oensingen, which did not (or only shortly) have a predecessor institution. The Olten campus region did not have a PET college specializing in STEM before the reform. The Oensingen campus region had a PET college that was founded in 1994, thus existing for only a few years before the UAS introduction in 1997. We create dummy variables for these two UAS regions and re-estimate our model including interactions of the treatment variable with these dummies. If the treatment effect in a UAS campus region without a former PET college (or with an only shortly existent former PET college) differs from the treatment effect in UAS campus regions with a former PET college, the interaction terms would yield a negative (or positive) and statistically significant coefficient.

However, the results in Table 2.15 show that the treatment effect in UAS campus

Table 2.14: Estimation results without MS regions comprising the 10 largest cities

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.249***	0.218***	0.227**	0.201**
	(0.076)	(0.068)	(0.077)	(0.070)
$TreatmentGroup_i$	0.218***	-0.022	0.233***	-0.013
	(0.065)	(0.068)	(0.067)	(0.069)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	0.981	0.981	1.046	1.046
N	149,146	149,146	149,146	149,146
R^2	0.001	0.181	0.001	0.178

Source: Authors' calculations based on ESS data.
Notes: MS regions Basel city, Bern (including the cities of Bern and Köniz), Biel/Bienne, Lucerne, Schaffhausen, St. Gallen, Thun, Winterthur, and Zurich excluded. Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

regions without a former PET college does not differ from the treatment effect in UAS campus regions with a former PET college. The coefficients of the interaction terms are statistically insignificant, implying that the treatment effect is comparable across the different UAS campus regions, irrespective of former PET colleges. Thus the existence of a predecessor institution was clearly no necessary precondition for determining UAS campus locations. This finding also corroborates the argument that the introduction of UASs constitutes a fundamental change for all education institutions and for the entire education sector.

Fourth, we study UAS campus closings resulting from campus mergers to determine whether regional economic factors influence UAS location decisions and R&D employment in UAS campus regions. If economic factors had determined the locations of these campuses in the first place, the closings would have no effect on R&D employment. In studying campus closings, we estimate an alternative DiD model in which we consider campus closings as the treatment.

For the analysis of campus closings, we create two variables necessary for the regressions. First, for each establishment in an MS region affected by a campus closing, we create a dummy variable that equals one for all years after the closing of the campus ($Closing_{i,t}$). The variable $ClosingGroup_i$ equals one for all establishments in MS regions affected by a campus closing. Second, we exclude observations in the control group of the original model, allowing us to investigate the effect of campus closings relative to the effect of persisting campuses. In other words, in the alternative model the establishments affected by a closing constitute the treatment group and the unaffected establishments constitute the control group.

The results of the alternative model in Table 2.16 show economically significant decreases in both the percentage of R&D personnel and the percentage of R&D wages. This finding indicates that the closing of a UAS campus leads to a decrease in R&D employment and confirms that other economic factors of UAS campus regions did not determine the locations of these campuses in the first place.

Fifth, we re-estimate our main DiD model without the control group observations

Table 2.15: Estimation results for UAS campus regions Oensingen and Olten

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.081	0.154**	0.063	0.140**
	(0.078)	(0.064)	(0.080)	(0.065)
$Treatment_{i,t-3} \times Oensingen_i$	0.132		0.188	
	(0.247)		(0.281)	
$Treatment_{i,t-3} \times Olten_i$		-0.094		-0.069
		(0.194)		(0.209)
$TreatmentGroup_i$	0.070	0.047	0.076	0.054
	(0.069)	(0.061)	(0.070)	(0.062)
$TreatmentGroup_i \times Oensingen_i$	-1.004***		-1.096***	
	(0.208)		(0.221)	
$TreatmentGroup_i \times Olten_i$		-0.402***		-0.407***
		(0.140)		(0.144)
Year dummies	Yes**	Yes**	Yes**	Yes**
Controls	Yes***	Yes***	Yes***	Yes***
Sample mean of dep. var.	0.894	0.935	0.952	0.990
N	74,813	152,789	74,813	152,789
R^2	0.204	0.186	0.202	0.184

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. Models (1) and (3) include observation years through 2002 (relocation of Oensingen campus in 2003). Model (2) and (4) include observation years through 2006 (relocation of Olten campus in 2006). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.16: Estimation results for consolidation of UAS campuses

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Closing_{i,t-3}$	0.030	-0.050	0.016	-0.062
	(0.120)	(0.111)	(0.122)	(0.113)
$ClosingGroup_i$	-0.106	-0.135	-0.107	-0.145*
	(0.075)	(0.085)	(0.077)	(0.086)
Year dummies	Yes	Yes***	Yes	Yes*
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.148	1.148	1.200	1.200
N	146,517	146,517	146,517	146,517
R^2	0.001	0.189	0.001	0.186

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

to cancel out any potential differences in economic preconditions between the treatment and control groups. In other words, we estimate the treatment effect exploiting only the variation in the timing of UAS campus introductions. If the treated regions were structurally different from untreated regions due to the underlying assignment procedure, we would not find any indication of a treatment effect in this analysis.

Table 2.17 presents the results using only the treated observations, showing a positive effect. However, due to a lack of power resulting from the small temporal variation of UAS introductions when using the biennial ESS data, this effect is not significant when we exclude the control group observations (columns 2 and 4 that include the sampling control variables). The insignificance probably results from the ESS's biennial data structure, which significantly reduces the temporal variation necessary to identify the treatment effect. As an estimation that relies only on temporal variation indicates a lower bound of the true treatment effects due to overweighted short-term effects (Borusyak and Jaravel, 2017), we argue that the insignificance using ESS data results from the insufficient data structure. This argument is supported by an analysis of Pfister et al. (2018), who use indicators of regional patenting as outcome variables and find highly significant positive effects in a similar analysis of treated observations only, but with a more detailed temporal structure.

In sum, our analyses of directly or indirectly observable economic preconditions support the argument that UAS campus locations were quasi-randomly determined. Furthermore, Pfister (2017) and Pfister et al. (2018) use patent data to provide evidence that points strongly in the same direction. The insignificance of the effects we find in some of the analyses likely results from the ESS data structure. None of the analyses on the economic preconditions potentially determining UAS introductions provides any evidence contradicting the quasi-randomness of the spatial and temporal variation in UAS campus openings.

Table 2.17: Estimation results without control group

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.156**	0.108	0.138*	0.099
	(0.075)	(0.071)	(0.077)	(0.073)
Year dummies	Yes	Yes*	Yes	Yes
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.148	1.148	1.200	1.200
N	146,517	146,517	146,517	146,517
R^2	0.001	0.189	0.001	0.186

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Composition of Treatment and Control Group. A further concern about our DiD strategy is whether the composition of firms and workers in both treatment and control groups remains stable after the treatment. For our identification strategy to be valid, such a stable composition of firms and workers—the economic actors (or mechanisms) that may cause the estimated effects—needs to be independent of the treatment, other than UAS graduates entering the labor markets in UAS regions. In other words, firms must not relocate from untreated regions to treated regions as a response to the treatment and UAS graduates must not move from treated regions to untreated regions after graduation. Therefore, we consider these two actors in the following.

Unfortunately, the ESS does not allow us to track firms over time and thus to assess relocations of firms. However, to assess the composition of firms in the treatment and control groups, we consult the results of Pfister (2017), who uses patent data to retrieve a proxy for relocations of firms. He calculates the number of patent applicants—most of which are legal entities (i.e., firms)—moving from untreated regions to treated regions before and after the introduction of UASs. Before this introduction, the percentage of applicants moving into a (future) treated region from an untreated region equals 0.6 percent. After the introduction of UASs, the same percentage equals 0.3 percent. Thus the firm composition remains stable within the treatment and control groups, with very few firms moving between the two both before and after the treatment. Pfister's (2017) analyses suggest that relocations of firms (i.e., movements of firms that hold patents) from the control group to the treatment group do not occur more frequently after the introduction of UASs than before. This finding also supports our assumption that firm relocations do not drive the UAS effect on R&D employment.

Regarding the mobility of UAS graduates, Pfister (2017) shows that UAS graduates exhibit particularly low mobility patterns and thus do not move from treated to untreated regions (or vice versa). To further analyze the mobility of UAS graduates, we perform more detailed calculations (extending Pfister, 2017). We use data from the SFSO's Survey of Higher Education Graduates and analyze moving mobility of UAS graduates (i.e., their change of residence after graduation in comparison to their residence during studies) and their working mobility (i.e., their change of workplace after graduation). In addition, we calculate moving mobility and working mobility of UNI graduates. We calculate these indicators as net fluctuations (i.e., the net value of inflows and outflows) in each region. The results show that mobility patterns of both UAS and UNI graduates are very low in Switzerland and thus do not affect the composition of the treatment and control groups.

Table 2.18 shows that the net moving mobility of UAS graduates is very low (column 1) and not different from that of UNI graduates (column 2).[40] Almost all UAS campus regions exhibit mobility shares below one percentage point (i.e., net fluctuations between treated and untreated regions are negligibly small). The working mobility of UAS graduates is also very low and not different from that of UNI graduates. Table 2.19 shows that the net working mobility is very low (mostly around two percent) for both UAS graduates (column 1) and UNI graduates (column 2). One noticeable exception is Zurich, which—as the major business hub in Switzerland—has a somewhat higher, but still low net inflow of 7.79 percent. In sum, these generally low mobility patterns of the Swiss population thus allow us to assume that the composition of UAS graduates remains stable across treatment and control groups.

2.7.6 Further Robustness Checks

In this appendix, we present the results from alternative specifications of the models we estimate in this chapter. First, we conduct robustness checks on our main estimation results in Section 2.5.1, including analyses on group-specific heterogeneity, the treatment radius, the assignment of employees to establishments, the impact of very large firms, the clustering of standard errors, and the levels of the dependent variables. Second, we conduct robustness checks on our results on effect heterogeneity in Section 2.5.3, including assessments on the role of start-ups and on the specification of firm-size classes. All these robustness checks confirm the findings of this chapter.

[40] The share of UNI graduates with STEM degrees moving to a foreign country equals 6.9 percent. The respective share of UAS graduates equals 2.6 percent.

Table 2.18: Moving mobility of UAS graduates and UNI graduates (net fluctuations)

UAS campus region	UAS graduates (1)	UNI graduates (2)
Control group	0.48%	1.39%
Bern	0.70%	0.00%
Biel	-0.94%	-0.70%
Brugg-Windisch	-0.57%	0.08%
Buchs	-0.26%	-0.15%
Burgdorf	-0.14%	-0.17%
Chur	0.26%	0.01%
Horw	0.58%	0.08%
Muttenz	0.11%	-1.03%
Oensingen	0.05%	0.09%
Olten	-0.28%	0.82%
Rapperswil	-0.06%	-0.26%
St. Gallen	-0.68%	-0.37%
Wädenswil	-0.81%	0.67%
Winterthur	0.36%	0.08%
Zurich	1.20%	-0.55%

Source: Calculations by Pfister (2017), Pfister et al. (2018), and authors' calculations with data from the SFSO's Survey of Higher Education Graduates.
Note: Positive values indicate net inflows and negative values indicate net outflows.

Table 2.19: Working mobility of UAS graduates and UNI graduates (net fluctuations)

UAS campus region	UAS graduates (1)	UNI graduates (2)
Control group	-2.14%	0.93%
Bern	3.87%	1.81%
Biel	-0.84%	-1.33%
Brugg-Windisch	0.29%	1.68%
Buchs	-0.28%	-0.09%
Burgdorf	-0.12%	-0.32%
Chur	-0.42%	-0.07%
Horw	-0.53%	-0.14%
Muttenz	-1.14%	0.32%
Oensingen	-0.25%	-0.42%
Olten	-0.99%	-1.00%
Rapperswil	-1.23%	-1.37%
St. Gallen	-0.89%	-0.70%
Wädenswil	0.25%	0.32%
Winterthur	-3.40%	-1.79%
Zurich	7.79%	2.15%

Source: Calculations by Pfister (2017), Pfister et al. (2018), and authors' calculations with data from the SFSO's Survey of Higher Education Graduates.
Note: Positive values indicate net inflows and negative values indicate net outflows.

2.7.6.1 Robustness of Main Results

Group-Specific Heterogeneity. As the ESS data has the structure of repeated cross-sections (i.e., firms and establishments cannot be tracked over time), we cannot apply FE estimation to account for unobservable time-invariant establishment characteristics. However, grouping establishments based on observable characteristics at least allows us to cancel out unobserved group-specific heterogeneity. Therefore, to obtain an unbalanced quasi-panel consisting of observations of each establishment group q at different points in time, we group[41] the establishment-level observations by MS region, industry sector (on 2-digit level of the NOGA 2002), and self-reported firm size (grouped into the 10 categories we use for our effect heterogeneity analysis in section 2.5.3). We then calculate the yearly averages of the dependent variables percentage of R&D personnel and percentage of R&D wages for each establishment group. Using this quasi-panel of establishment groups for conducting an FE estimation helps overcome violations of the parallel trends assumption that our assessments in Section 2.4.3 did not detect.

The quasi-panel FE estimation yields a treatment coefficient of 0.18 on percentage of R&D personnel (table 2.20, column 1). This effect ist statistically significant at the five percent level. Considering the slightly larger quasi-panel sample mean of 1.20, the size of the effect equals the effect we obtain from our main DiD specification. For the percentage of R&D wages, the quasi-panel treatment effect also equals the DiD treatment effect in size with a 0.16 percentage point increase and a sample mean of 1.28. We conclude from the FE estimation results that unobserved group-specific heterogeneity does not explain the DiD results, implying that the DiD results identify the true treatment effect.

[41] On average, a group consists of 3.05 establishments per year.

Table 2.20: Quasi-panel estimation results

	RDP^{pct}	RDW^{pct}
	(1)	(2)
$Treatment_{q,t-3}$	0.183**	0.163**
	(0.077)	(0.080)
Year dummies	Yes***	Yes***
Sample mean of dep. var.	1.199	1.279
N	76,193	76,193
R^2 (within)	0.001	0.001
R^2 (between)	0.001	0.001
R^2 (overall)	0.001	0.001

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors in parentheses. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Variation of Treatment Radius. Pfister et al. (2018) demonstrate that 25 kilometers is the most appropriate radius to choose for the analysis of Swiss UASs, because this radius best represents the commuting distance of about 90 percent of the Swiss population. Nonetheless, we check the robustness of our results to alternative radii, applying radii of 10, 20, 30, 40, and 50 kilometers, respectively.

The results of the specifications with alternative treatment radii (tables 2.21–2.25) confirm the results of our main specification using the 25-kilometer radius, with only some minor differences. Most importantly, when applying radii of 20, 30, and 40 kilometers, respectively, we always find positive and significant treatment effects as in our preferred specification with a radius of 25 kilometers. Only in the specifications using 10 or 50 kilometers, the treatment effects turn insignificant. However, both results are not surprising given our conservative assignment procedure of MS regions to the treatment group. In the 10-kilometers specification, this assignment procedure restricts the treatment group for the estimation to only two (of 73) MS regions and thus the treatment effects of 13 STEM UAS campuses vanish in the estimations. Vice versa, in the 50-kilometers specification, our assignment procedure puts 60 MS regions into the treatment group for the estimation, although the introduction of UASs has no effect on these regions. But apart from these two mechanically explainable insignificant results, all reasonable decreases or increases in the treatment radius strongly support our main results in Section 2.5.1 with the 25-kilometer radius.

Assignment of Employees to Establishments. As our procedure of assigning employees to different establishments within a firm is based on the respective employee's workplace location (following Janssen et al., 2016) instead of on direct information on different establishment locations, we check whether this assignment procedure affects our main results in Section 2.5.1. In so doing, we re-estimate our model with a restricted sample of single-establishment firms (i.e., all employees of a firm are reported to work in the same MS region). This restriction ensures the correct assignment of employees for all firms in the sample. We find that the results are similar if we restrict the sample

Table 2.21: Estimation results with 10-kilometer treatment radius

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	-0.019	-0.137	-0.056	-0.165
	(0.178)	(0.159)	(0.182)	(0.162)
$TreatmentGroup_i$	0.295*	0.291*	0.289*	0.313**
	(0.170)	(0.156)	(0.174)	(0.159)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.016	1.016	1.073	1.073
N	232,228	232,228	232,228	232,228
MS regions in treatment group (%)	2.740	2.740	2.740	2.740
Observations in treatment group (%)	14.383	14.383	14.383	14.383
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.22: Estimation results with 20-kilometer treatment radius

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.183***	0.194***	0.172**	0.188***
	(0.069)	(0.068)	(0.070)	(0.070)
$TreatmentGroup_i$	0.235***	0.039	0.228***	0.044
	(0.064)	(0.065)	(0.066)	(0.066)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.016	1.016	1.073	1.073
N	232,228	232,228	232,228	232,228
MS regions in treatment group (%)	24.658	24.658	24.658	24.658
Observations in treatment group (%)	44.283	44.283	44.283	44.283
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.23: Estimation results with 30-kilometer treatment radius

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.145**	0.144**	0.138**	0.140**
	(0.061)	(0.058)	(0.062)	(0.059)
$TreatmentGroup_i$	0.245***	0.021	0.247***	0.032
	(0.061)	(0.059)	(0.062)	(0.060)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.016	1.016	1.073	1.073
N	232,228	232,228	232,228	232,228
MS regions in treatment group (%)	56.164	56.164	56.164	56.164
Observations in treatment group (%)	74.967	74.967	74.967	74.967
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.24: Estimation results with 40-kilometer treatment radius

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.323***	0.173***	0.318***	0.175***
	(0.058)	(0.058)	(0.060)	(0.059)
$TreatmentGroup_i$	0.200***	0.100	0.217***	0.111
	(0.063)	(0.066)	(0.065)	(0.068)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.016	1.016	1.073	1.073
N	232,228	232,228	232,228	232,228
MS regions in treatment group (%)	75.342	75.342	75.342	75.342
Observations in treatment group (%)	90.529	90.529	90.529	90.529
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.25: Estimation results with 50-kilometer treatment radius

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.313***	0.130	0.298***	0.117
	(0.097)	(0.091)	(0.099)	(0.093)
$TreatmentGroup_i$	0.349***	0.133	0.389***	0.164*
	(0.100)	(0.095)	(0.101)	(0.097)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.016	1.016	1.073	1.073
N	232,228	232,228	232,228	232,228
MS regions in treatment group (%)	82.192	82.192	82.192	82.192
Observations in treatment group (%)	93.806	93.806	93.806	93.806
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

to single-establishment firms (table 2.26). This finding shows that our main results are robust to the location assignments of employees in multi-establishment firms.

Impact of Very Large Firms. Our results on effect heterogeneity by firm-size class in Section 2.5.3 indicate that very large firms (with 5,000 or more employees) are one of the major profiteers of the introduction of UASs. To check whether this particular group of firms also drives our main results in Section 2.5.1, we repeat our estimations without establishments belonging to this group of firms. This analysis thus yields insights into whether very large firms are the main driver of the effect. Table 2.27 shows that excluding very large firms does not alter the treatment effects. Therefore, very large firms are not the main driver of the treatment effects.

Clustering of Standard Errors. While we cluster standard errors at the firm level in our main specification, because establishments belonging to the same firm share common organizational structures, the observations in our sample might also be subject to common regional structures that the canton FE included in the vector of control variables do not yet capture. To assess how the clustering of the standard errors affects our main results, we re-estimate our model with regionally clustered standard errors. Tables 2.28 and 2.29 report the main results with standard errors clustered at the MS region level (the smallest regional level available in our data and also the regional level at which we define the treatment and control groups) and the canton level (the next higher administrative regional level), respectively. As the significance levels of the regression coefficients in Tables 2.28 and 2.29 do not change in comparison to our main specification, our results are robust to clustering standard errors at regional levels.

Table 2.26: Estimation results with single-establishment firms only

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.218***	0.205***	0.201**	0.197**
	(0.079)	(0.073)	(0.080)	(0.075)
$TreatmentGroup_i$	0.226***	-0.071	0.275***	-0.066
	(0.072)	(0.072)	(0.074)	(0.075)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.108	1.108	1.175	1.175
N	143,825	143,825	143,825	143,825
R^2	0.002	0.192	0.001	0.189

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.27: Estimation results without firms with 5,000 or more employees

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.163***	0.157***	0.140**	0.143***
	(0.058)	(0.054)	(0.060)	(0.055)
$TreatmentGroup_i$	0.220***	0.021	0.225***	0.029
	(0.054)	(0.052)	(0.055)	(0.054)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.025	1.025	1.082	1.082
N	229,201	229,201	229,201	229,201
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the firm level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.28: Estimation results with standard errors clustered at the MS region level

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.166	0.158***	0.143	0.145**
	(0.111)	(0.060)	(0.112)	(0.064)
$TreatmentGroup_i$	0.227*	0.027	0.233*	0.035
	(0.127)	(0.067)	(0.130)	(0.071)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.016	1.016	1.073	1.073
N	232,228	232,228	232,228	232,228
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the MS region level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.29: Estimation results with standard errors clustered at the canton level

	RDP^{pct}		RDW^{pct}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.166*	0.158***	0.143	0.145**
	(0.086)	(0.054)	(0.086)	(0.058)
$TreatmentGroup_i$	0.227	0.027	0.233	0.035
	(0.150)	(0.080)	(0.152)	(0.085)
Year dummies	Yes***	Yes***	Yes***	Yes***
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	1.016	1.016	1.073	1.073
N	232,228	232,228	232,228	232,228
R^2	0.001	0.171	0.001	0.169

Source: Authors' calculations based on ESS data.
Notes: Robust standard errors clustered at the canton level in parentheses. All models include intercept. Coefficients, standard errors, and sample mean of dep. var. are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Levels of Dependent Variables. In a further robustness check, we analyze the levels (instead of the percentages) of our two main dependent variables R&D personnel and R&D wages. Unfortunately, the ESS yields inconsistencies in the number of reported employee observations. While these inconsistencies do not affect the percentage of our outcome variables, they do affect the levels.[42] To overcome the inconsistencies in the number of reported employee observations, we again restrict the sample to single-establishment firms for the analysis of levels.

Analyzing the sample of single-establishment firms, for which we can consistently determine the outcome variables' levels, yields positive but statistically insignificant effects (table 2.30). For single-establishment firms, the levels of R&D employment and R&D wages thus at least remain constant and do not fall. The insignificance of the effects potentially results from the far lower average firm size in this restricted sample (32.66 employees) than in the full sample (318.11 employees).

2.7.6.2 Robustness of Results on Effect Heterogeneity

The Role of Start-Ups. To complement our heterogeneity analysis on the role of start-ups as one determinant for the treatment effects (section 2.5.3), we conduct an additional analysis that uses proxy measures for firm and establishment age. As the ESS does not directly indicate a firm's or an establishment's age, we use the maximum tenure of an employee within a firm or an establishment to proxy the age of the firm or the establishment.[43] Put differently, at the firm level and at the establishment level, we calculate age as the maximum tenure reported in the ESS.[44] We define age classes

[42] Sampling in the ESS takes place at the firm level, that is, the SFSO sends the questionnaires to the sampled firms. The SFSO instructs these firms to report information on a certain percentage of their employees, depending on the size of the firm. In the data, we observe that firms very frequently report far more employees than they were instructed to. Unfortunately, whether the additionally reported employees are equally distributed across all establishments or whether they concentrate in a few establishments is unknown. The resulting bias affects the levels, but not the percentage of our outcome variables.

[43] Firm age indicates how old firms are; establishment age indicates the age of a particular establishment of an already existing firm.

[44] For around 0.7 percent of the establishment-level observations in our estimation sample, the age variables contain no information. To be conservative, we set age to 50, the maximum value in the sample, for these observations.

Table 2.30: Estimation results for levels of dependent variables

	RDP^{abs}		RDW^{abs}	
	(1)	(2)	(3)	(4)
$Treatment_{i,t-3}$	0.307*	0.049	2,666.005**	362.129
	(0.178)	(0.151)	(1,261.505)	(980.321)
$TreatmentGroup_i$	0.126	-0.030	795.508*	-81.142
	(0.127)	(0.125)	(774.815)	(799.455)
Year dummies	Yes***	Yes	Yes	Yes
Controls	No	Yes***	No	Yes***
Sample mean of dep. var.	0.830	0.830	4,654.197	4,654.197
N	143,825	143,825	143,825	143,825
R^2	0.000	0.016	0.000	0.011

Source: Authors' calculations based on ESS data.
Notes: Regressions include only single-establishment firms. Robust standard errors clustered at the firm level in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

by dividing the respective age variable into deciles of its distribution. We then conduct heterogeneity analyses by firm and establishment age, similar to the analyses in Section 2.5.3, in which we analyze heterogeneity by firm size or industry sector. The results of the analyses by firm and establishment age strengthen our argument that both start-up firms and new establishments of already existing firms contribute strongly to the R&D employment effects we find in our main analysis.

Table 2.31 shows the estimation results for heterogeneity by firm age. The results indicate that older firms (eighth and ninth firm-age decile) as well as very young firms (first and second firm-age decile) profit most from the introduction of UASs, confirming our results on firms' probability to conduct R&D in Section 2.5.3. Table 2.31 shows a highly significant effect on the probability to conduct R&D for firms in the eighth firm-age decile (26 to 31 years). Furthermore, at the ten percent significance level, firms in the first firm-age decile (zero to three years) have a higher probability to conduct R&D and firms in the second firm-age decile (four to six years) and in the ninth firm-age decile (32 to 38 years) increase the percentage of R&D workers and the percentage of wages paid to these workers.

Regarding establishment age, our results indicate that younger establishments (second establishment-age decile) and old establishments (ninth decile) profit most from the introduction of UASs. Table 2.32 shows that establishments in the second establishment-age decile (four to five years) increase their percentage of R&D workers (significant at the five percent level), the percentage paid to these workers (significant at the ten percent level), and have a higher probability to conduct R&D (significant at the five percent level). At the 10 percent significance level, establishments in the ninth establishment-age decile (31 to 36 years) increase their percentage of R&D workers and the percentage of wages paid to these workers.

In sum, the age heterogeneity analyses with ESS data show that both start-up firms and new establishments of already existing firms contribute most to the effects we find. However, the age variables constitute rough and potentially biased proxies. Employee turnover might be industry-specific and, therefore, the age proxies might yield a downward

Table 2.31: Estimation results by firm-age decile

	Firm age (FA_i)	N	RDP^{pct} (1)	RDW^{pct} (2)	RD^{bin} (3)
	$0 \leq FA_i \leq 3$	18,434	0.381	0.394	0.760*
			(0.243)	(0.247)	(0.404)
	$4 \leq FA_i \leq 6$	21,272	0.398*	0.362*	0.597
			(0.212)	(0.213)	(0.408)
	$7 \leq FA_i \leq 9$	18,783	0.266	0.170	0.406
			(0.235)	(0.239)	(0.473)
	$10 \leq FA_i \leq 12$	20,069	0.084	0.056	0.029
			(0.202)	(0.206)	(0.446)
	$13 \leq FA_i \leq 16$	20,775	0.159	0.145	0.633
			(0.194)	(0.198)	(0.461)
$Treatment_{i,t-3}$	$17 \leq FA_i \leq 20$	20,087	0.136	0.118	0.351
			(0.170)	(0.182)	(0.505)
	$21 \leq FA_i \leq 25$	22,245	0.226	0.217	0.668
			(0.196)	(0.199)	(0.485)
	$26 \leq FA_i \leq 31$	24,233	0.106	0.113	1.343***
			(0.162)	(0.167)	(0.488)
	$32 \leq FA_i \leq 38$	29,342	0.189*	0.191*	0.483*
			(0.109)	(0.113)	(0.471)
	$39 \leq FA_i \leq 50$	36,988	-0.144	-0.165	-0.096
			(0.122)	(0.127)	(0.475)

Source: Authors' calculations based on ESS data.
Notes: Results from separate regressions. Robust standard errors clustered at the firm level in parentheses. All models include intercept, treatment group dummy, year dummies, and sampling control variables. Coefficients and standard errors are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.32: Estimation results by establishment-age decile

	Establishment age (EA_i)	N	RDP^{pct} (1)	RDW^{pct} (2)	RD^{bin} (3)
	$0 \leq EA_i \leq 3$	32,519	0.126	0.125	0.342
			(0.173)	(0.175)	(0.253)
	$4 \leq EA_i \leq 5$	19,908	0.388**	0.364*	0.724**
			(0.190)	(0.193)	(0.338)
	$6 \leq EA_i \leq 7$	25,007	0.168	0.120	0.063
			(0.191)	(0.194)	(0.369)
	$8 \leq EA_i \leq 10$	16,679	0.247	0.195	0.345
			(0.193)	(0.201)	(0.431)
	$11 \leq EA_i \leq 14$	23,468	0.036	-0.000	0.231
			(0.171)	(0.175)	(0.400)
$Treatment_{i,t-3}$	$15 \leq EA_i \leq 18$	25,317	0.111	0.120	0.401
			(0.173)	(0.178)	(0.426)
	$19 \leq EA_i \leq 23$	22,543	-0.025	-0.033	0.237
			(0.155)	(0.162)	(0.505)
	$24 \leq EA_i \leq 30$	23,276	0.138	0.123	0.827
			(0.141)	(0.146)	(0.518)
	$31 \leq EA_i \leq 36$	22,667	0.228*	0.241*	0.408
			(0.130)	(0.135)	(0.617)
	$37 \leq EA_i \leq 50$	20,844	0.101	0.101	0.316
			(0.131)	(0.137)	(0.789)

Source: Authors' calculations based on ESS data.
Notes: Results from separate regressions. Robust standard errors clustered at the firm level in parentheses. All models include intercept, treatment group dummy, year dummies, and sampling control variables. Coefficients and standard errors are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

bias for industries with high employee turnover. This potential bias calls for some caution in interpreting the results of the age heterogeneity analyses. Nevertheless, these results support our findings in Section 2.5.3.

Alternative Specifications of Firm-Size Classes. While we choose the firm-size thresholds as proposed in the OECD's (2015) Frascati Manual in our analyses of effect heterogeneity by firm-size class (section 2.5.3), these thresholds lead to a rather large variation in the number of establishments per firm-size class. Therefore, to check the robustness of our findings on heterogeneity by firm-size class and to gain additional insights to our results, we apply alternative specifications of firm-size classes by splitting up firms into quintiles and deciles of the firm-size distribution.

Tables 2.33 and 2.34 show the estimation results by firm-size quintiles and deciles, respectively. Note that a variation in the number of establishments per class still exists, because we split establishments (two or more of which can belong to the same firm) by firm size. The results of these regressions by firm-size quintile and decile confirm the original finding that very small firms (with five to six employees according to the specification with establishments split into firm-size deciles) drive the treatment effect. However, given that relatively few establishments belong to very large firms (5,000 or more employees), the treatment effects for these establishments are no longer significant as in the original specification due to the different definition of firm-size classes.

Table 2.33: Estimation results by fim-size quintile

			RDP^{pct}	RDW^{pct}	RD^{bin}
	Firm size (S_i)	N	(1)	(2)	(3)
$Treatment_{i,t-3}$	$S_i \leq 3$	40,368	0.069	0.069	0.116
			(0.135)	(0.136)	(0.174)
	$4 \leq S_i \leq 6$	31,070	0.326**	0.313*	0.660**
			(0.159)	(0.164)	(0.283)
	$7 \leq S_i \leq 15$	36,558	0.223	0.204	0.463
			(0.153)	(0.157)	(0.359)
	$16 \leq S_i \leq 43$	42,114	0.255*	0.216	0.391
			(0.147)	(0.151)	(0.422)
	$44 \leq S_i$	82,118	0.026	0.007	0.280
			(0.083)	(0.085)	(0.292)

Source: Authors' calculations based on ESS data.
Notes: Results from separate regressions. Robust standard errors clustered at the firm level in parentheses. All models include intercept, treatment group dummy, year dummies, and sampling control variables. Coefficients and standard errors are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.34: Estimation results by fim-size decile

			RDP^{pct}	RDW^{pct}	RD^{bin}
	Firm size (S_i)	N	(1)	(2)	(3)
$Treatment_{i,t-3}$	$S_i \leq 2$	25,284	0.180	0.181	0.122
			(0.154)	(0.156)	(0.189)
	$S_i = 3$	15,084	-0.121	-0.119	0.083
			(0.256)	(0.257)	(0.346)
	$S_i = 4$	12,661	0.057	0.004	0.352
			(0.273)	(0.283)	(0.451)
	$5 \leq S_i \leq 6$	18,409	0.504***	0.518***	0.872**
			(0.192)	(0.198)	(0.361)
	$7 \leq S_i \leq 9$	16,725	0.126	0.127	0.247
			(0.198)	(0.204)	(0.465)
	$10 \leq S_i \leq 15$	19,833	0.297	0.264	0.579
			(0.225)	(0.232)	(0.528)
	$16 \leq S_i \leq 25$	21,180	0.283	0.215	0.584
			(0.206)	(0.211)	(0.564)
	$26 \leq S_i \leq 43$	20,934	0.273	0.262	0.191
			(0.215)	(0.218)	(0.626)
	$44 \leq S_i \leq 95$	25,724	0.274*	0.240	0.746
			(0.158)	(0.165)	(0.566)
	$96 \leq S_i$	56,394	-0.075	-0.095	-0.015
			(0.097)	(0.100)	(0.344)

Source: Authors' calculations based on ESS data.
Notes: Results from separate regressions. Robust standard errors clustered at the firm level in parentheses. All models include intercept, treatment group dummy, year dummies, and sampling control variables. Coefficients and standard errors are multiplied by 100 to represent percentage point changes. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Chapter 3

Proxying Economic Activity with Daytime Satellite Imagery: Filling Data Gaps Across Time and Space

Part of this chapter is an extended version of early parts of the working paper "Proxying Economic Activity with Daytime Satellite Imagery: Filling Data Gaps Across Time and Space" by Lehnert, Niederberger, Backes-Gellner, and Bettinger (2021).

3.1 Introduction

The lack of credible data hampers our understanding of regional economic development, especially in historical contexts. Most countries lack data at the regional or even municipal levels, and the extant data either focus only on recent years or lack consistency across regions and/or time. To fill these data gaps across time and space, researchers have increasingly used satellite data on night light intensity as a proxy for economic activity (e.g., Dingel et al., 2021; Hodler and Raschky, 2014; Michalopoulos and Papaioannou, 2013; Pinkovskiy and Sala-i-Martin, 2016).

However, night light intensity data have significant weaknesses. They are available only for a limited time series (from 1992) and, due to their spatial resolution (one kilometer at the equator), they are not reliable for disaggregated regional units such as municipalities or suburbs (Chen and Nordhaus, 2011; Kulkarni et al., 2011; Mellander et al., 2015).

Administrative or survey data on economic activity encounter similar problems. They are typically not available for longer historical time series, not regionally disaggregated, or otherwise unreliable or unavailable to the research community.

This chapter demonstrates how daytime satellite imagery solves these key weaknesses and serves as a novel proxy for economic activity across time periods and for highly disaggregated spatial units. We derive this proxy by applying machine-learning techniques to daytime satellite imagery from the Landsat program. The proxy presents valuable information on economic activity over a uniquely long time series (from 1984) at a level of regional disaggregation that is smaller (30-meter resolution) than any alternative.

Daytime satellite imagery from the Landsat program has so far received almost no attention in economics applications. The few existing applications rely on visual interpretation for identifying, for example, agricultural land use, (de)forestation, or urbanization (e.g., Burchfield et al., 2006; Foster and Rosenzweig, 2003). Developing new and more accurate proxies with Landsat data requires novel machine-learning techniques to adapt these data to economic settings. Tools such as the Google Earth Engine (GEE) facilitate the processing and analysis of Landsat's large geographic datasets (Gorelick et al., 2017).

Landsat daytime satellite data have three advantages over other data sources. First, Landsat data have substantially higher disaggregation (30-meter resolution) than regional administrative or other satellite data such as night light intensity (one-kilometer resolution) (Donaldson and Storeygard, 2016). This higher resolution entails more precise information at a much more disaggregated regional level. Our economic proxy can characterize economic development at regional levels and even in much smaller localities such as municipalities or urban districts.

Second, the National Aeronautics and Space Administration (NASA) launched the first satellite of the Landsat program (Landsat-1) in 1972, making Landsat the earliest existing source of regionally highly disaggregated satellite data (Morain, 1998; Williams et al., 2006). While Landsat did not reach its full potential until 1984, the long time horizon of the data allows researchers to construct longer historical data than other regional economic administrative data or other proxies based on satellite data such as night light intensity (which is available from 1992). In comparison to regional economic data, which might be available for some administrative locations, Landsat daytime data pre-date the break-up of the former Soviet Union, German reunification, and other significant changes in regional or even local economic development.

Third, Landsat satellites collect multispectral imagery of the earth, that is, they observe the energy that the earth reflects in different spectral bands (e.g., infrared or ultraviolet). The geographic remote-sensing literature has successfully applied machine-learning techniques that exploit this multispectral information in the identification of different types of land cover from subsets of Landsat data (e.g., Dewan and Yamaguchi, 2009; Goldblatt et al., 2016; Liu et al., 2018; Pekkarinen et al., 2009; Yu et al., 2011). We extend this literature by creating a procedure that combines all Landsat data available from 1984 to map six different types of land cover, which we call "surface groups": built-up surfaces, grassy surfaces, forest-covered surfaces, surfaces with crop fields, surfaces without vegetation, and water surfaces.

As some surface groups are more closely related to economic activity than others (Keola et al., 2015; Sutton and Costanza, 2002), mapping surface groups yields important information on regional economic activity. For example, increases in built-up surfaces, which include agglomerations of cities or transportation networks, coincide with increases in economic activity (Davis et al., 2014; Holl, 2004). Even holding built-up surfaces constant, the other surface groups provide greater predictability of local economic conditions. While Goldblatt et al. (2020) find that the raw spectral values of Landsat-7 imagery can serve as a slightly better proxy than night light intensity in Vietnamese regions, we show that identifying the different surface groups through machine-learning techniques results in a substantially improved proxy for economic activity over time and space. Moreover, compared to Yeh et al.'s (2020) application that uses Landsat imagery to directly predict village asset wealth in Africa, our surface groups can function both as an indicator of land cover and as a more general proxy for economic activity with the potential for worldwide application. Which proxy to choose for empirical research depends on the concrete research question, with other proxies offering advantages through specialization in, for example, asset wealth (Yeh et al., 2020) and our proxy offering advantages through painting an overall picture of regional (or even more subregional) economic activity.

Our procedure for detecting surface groups as a proxy for economic activity produces a metric with high internal and external validity. We lay the foundations for computing and validating the proxy using Germany as an example. In the context of the German reunification, our proxy provides important, previously unavailable, yet reliable information on economic activity in East German regions. As such, the surface groups allow the examination of pre-reunification economic developments at highly disaggregated regional levels and—due to their independence of politically motivated economic statistics produced during the communist era—with very high validity. These analyses are otherwise impossible with other data. Our data and their applications easily extend to other settings and geographies throughout the world.

3.2 The Value of Surface Groups as a Proxy for Economic Activity

3.2.1 Features of Surface Groups

We use a supervised machine-learning algorithm to classify Landsat pixels into one of six surface groups. This classification procedure requires two external data sources. First, the raw imagery of Landsat satellites constitutes the input data to be classified. Before performing the classification, we pre-process this raw imagery to obtain pixel-based annual composites incorporating imagery from multiple Landsat satellites. Second, CORINE Land Cover (CLC) data (which are available only for the five reference years 1990, 2000, 2006, 2012, and 2018) serve as an external source of ground-truth information, that is, they indicate the true surface group for a subset of the input pixels. The training data for the classification algorithm consist of a stratified random sample of Landsat pixels matched to their true surface group from CLC data. The details of the classification procedure are outlined in Section 3.4 and in Appendix 3.5.1.

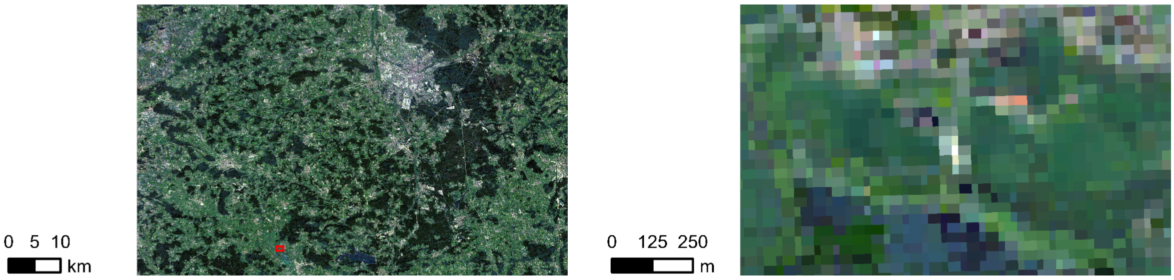
Following prior literature utilizing land cover classifications (e.g., Balzter et al., 2015; Han et al., 2004; Neumann et al., 2007; Pérez-Hoyos et al., 2012; Waser and Schwarz, 2006), we identify and map six different types of land cover—the surface groups—which are similar to Yu et al.'s (2011) work in a Chinese region. These groups include the following: (1) built-up surfaces feature buildings of non-natural materials such as concrete, metal, and glass (e.g., residential buildings, industrial plants, roads); (2) grassy surfaces are covered by green plants or groundcover with similar surface reflectance (e.g., natural grassland); (3) surfaces with crop fields include vegetation for agricultural purposes (e.g., grain fields); (4) forest-covered surfaces contain trees or other plants with similar surface reflectance (e.g., mixed forests); (5) surfaces without vegetation have (almost) no reflective vegetation or buildings (e.g., bare rock); and (6) water surfaces comprise any type of water surface (e.g., lakes). Our algorithm classifies these respective surfaces, which we then combine to form our proxy for economic activity.

The output of our procedure for detecting surface groups is a dataset containing the surface group of every Landsat pixel location in Germany annually from 1984 through 2020.[1] One year comprises more than 630 million Landsat pixels, amounting to more than 23 billion pixel-year observations in the output data. Of these observations, 16.2 percent are classified as built-up; 20.9 percent as grass; 29.5 percent as crops; 25.6 percent as forest; 3.3 percent as no vegetation; and 3.8 percent as water. Only 0.6 percent of observations contain missing values due to, for example, cloud cover that is uninterrupted within a given year for single pixels in the Landsat data. For applications in research projects, researchers can aggregate this pixel-level information to the geographical unit matching their respective research objective (e.g., administrative regional units or ZIP code areas).

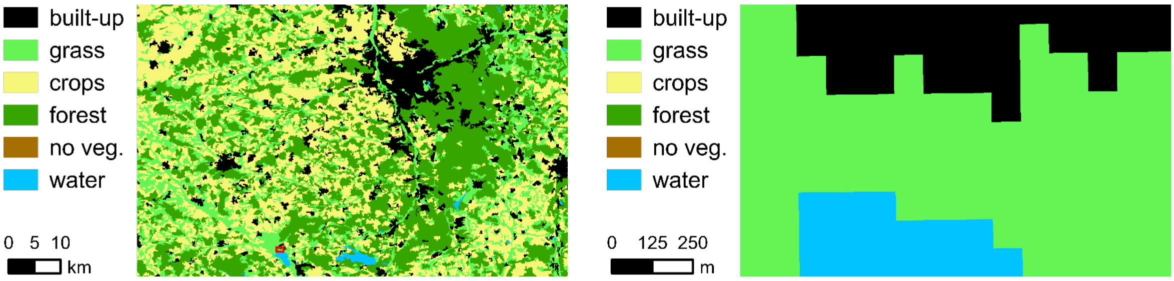
Figure 3.1 illustrates the data sources we use and the output data we produce. As examples, the left column of Figure 3.1 shows a large-scale area with the metropolitan region of Nuremberg (situated in mid-south Germany) in the center of the picture. The right column shows a small-scale area with the village of Muhr am See (situated about 30 miles south-west of Nuremberg) in the upper part of the picture and its accompanying

[1] This chapter uses the most recent surface groups dataset produced in January 2022 for Germany.

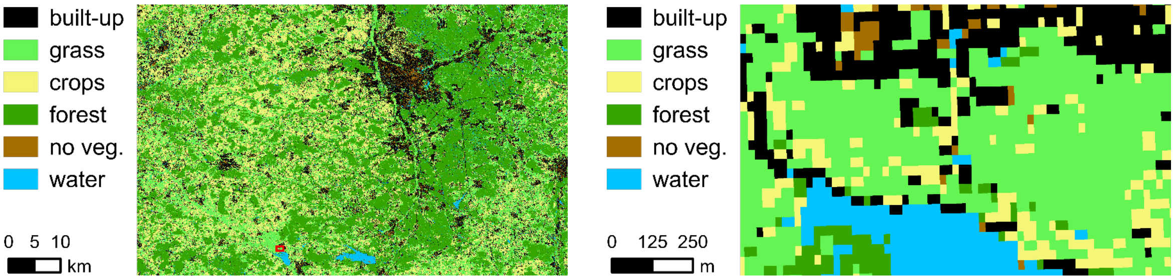
Figure 3.1: Visual comparison of data sources



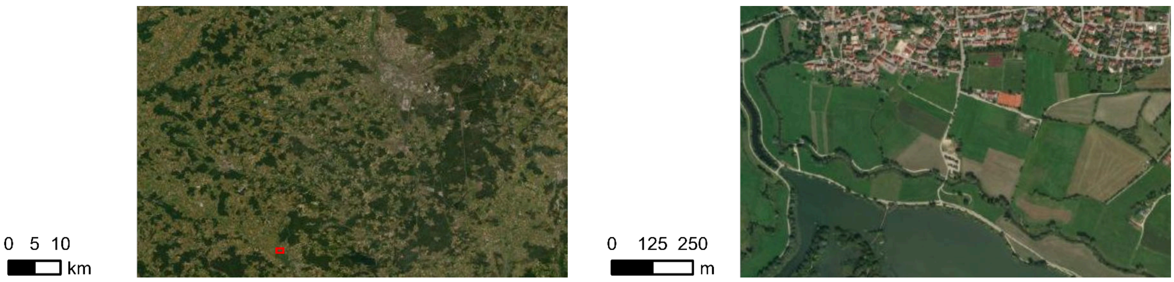
(a) Input: Greenest pixel composite (2018)



(b) Ground-truth: CLC classification (2018)



(c) Output: Surface groups (2018)



(d) Reference: Esri World Imagery (2022)

Source: Authors' illustrations using Esri World Imagery, Landsat data, and CLC data.
Notes: Pictures in the left column show the same approx. 78 × 49 square kilometers area with the metropolitan region of Nuremberg in the center. Pictures in the right column show the same approx. 1.3 × 0.8 square kilometers area with the village of Muhr am See in the upper part and its accompanying lake (Altmühlsee) in the lower part (area framed red in the left column).

lake (Altmühlsee) in the lower part of the picture (the area framed red in the left column). Figure 3.1a, which uses Landsat's visible spectral bands to approximate the perception of the human eye, shows the Landsat composite for 2018 (the input data). Figure 3.1b illustrates the six different types of land cover we identify from the CLC data (the ground-truth data). Figure 3.1c shows the surface group that our classification algorithm produces for every Landsat pixel location in 2018. As a reference, Figure 3.1d shows current high-resolution satellite images from Esri World Imagery (Esri et al., 2009).

3.2.2 Internal Validity

To evaluate whether we achieve an accurate classification of Landsat pixels into the six surface groups (i.e., the measure's internal validity), we assess several indicators of prediction accuracy. In so doing, we follow the standard procedure in the remote-sensing literature that uses supervised machine learning to classify land cover (e.g., Dewan and Yamaguchi, 2009; Goldblatt et al., 2016; Schneider, 2012) and derive these indicators from five-fold cross-validation. This method draws five subsets from the input data and uses these subsets to perform five iterations of pixel classification (see section 3.4 and appendix 3.5.1.5 for more details).

Using the classification output from the five-fold cross-validation, we calculate five common indicators of prediction accuracy with respect to each surface group: overall accuracy, true-positive rate, true-negative rate, balanced accuracy, and user's accuracy. Overall accuracy denotes the percentage of pixels correctly classified, true-positive rate the percentage of pixels correctly classified as belonging to the respective surface group, true-negative rate the percentage of pixels correctly classified as not belonging to the respective surface group, balanced accuracy the average of true-positive rate and true-negative rate, and user's accuracy the percentage of pixels correctly classified as belonging to the respective surface group among all pixels belonging to the respective surface group.

Table 3.1 shows the five-fold cross-validation results with respect to each surface group. With 82.8 percent, overall accuracy for built-up surface areas is similar to that in other studies detecting built-up land with Landsat data (e.g., Dewan and Yamaguchi, 2009;

Goldblatt et al., 2016). The other four indicators are also in line with other studies (e.g., Goldblatt et al., 2018, 2016). Furthermore, we achieve very high overall accuracy for forest (89.5 percent), areas with no vegetation (87.0 percent), and water (90.9 percent).

The five-fold cross-validation results show that our output data constitute an internally valid measure of land cover. All indicators of prediction accuracy reinforce that our classification algorithm accurately identifies the six surface groups, suggesting that we adequately implemented the procedures from the remote-sensing literature. The high internal validity of the surface groups is a prerequisite for their external validity as a proxy for economic activity.

3.2.3 External Validity

To evaluate the external validity of surface groups as a proxy for economic activity, we empirically analyze how much they explain of the variation in direct measures of regional economic activity (which are available for parts of our time series). We draw on two such external direct measures: First, from administrative statistics, we extract a regionally disaggregated direct measure of GDP, the most commonly used economic indicator in the literature evaluating previous satellite-based proxies for economic activity (e.g., Chen and Nordhaus, 2011; Henderson et al., 2012). For Germany, this measure is available at the administrative county (*Kreis*) level from 2000. Second, we use a socioeconomic dataset from Leibniz Institute for Economic Research (RWI) and Micromarketing-Systeme und Consult GmbH (microm) (2019) that provides household income as a further indicator of economic activity with a very high level of regional detail. This indicator is available at the level of grid cells sized one square kilometer (and thus independent of administrative borders), but annually only from 2009. See Section 3.4 and Appendix 3.5.2.2 for more details on the two external validation data sources.

We analyze the external validity of our proxy by comparing the amount of variation in GDP that our proxy and night light intensity generate. We obtain this result from comparing OLS regressions of GDP on the surface groups with OLS regressions of GDP on night light intensity (see Section 3.4 and Appendix 3.5.2.3 for more details on the

Table 3.1: Five-fold cross-validation results

Surface group	Overall accuracy	True-positive rate	True-negative rate	Balanced accuracy	User's accuracy
built-up	0.828	0.606	0.877	0.741	0.514
grass	0.831	0.451	0.910	0.680	0.511
crops	0.832	0.381	0.932	0.657	0.563
forest	0.895	0.685	0.938	0.812	0.708
no veg.	0.870	0.756	0.886	0.821	0.490
water	0.909	0.672	0.958	0.815	0.765

Source: Authors' calculations based on Landsat data and CLC data.
Notes: Indicators calculated with respect to each surface group. Values indicate the average over all five iterations and all five CLC reference years. See Section 3.4 and Appendix 3.5.1.5 for more details (including the results separately for every reference year).

methodology). Our preferred surface-groups specification, which additionally includes year and federal state FE to cancel out any bias due to potential measurement error in the dependent or independent variables, explains 62.3 percent of the variation in GDP. Using night light intensity instead of surface groups in the same specification explains only 47.1 percent of this variation, that is, our proxy achieves 32.3 percent higher precision than previous data at the disaggregated regional level of counties.

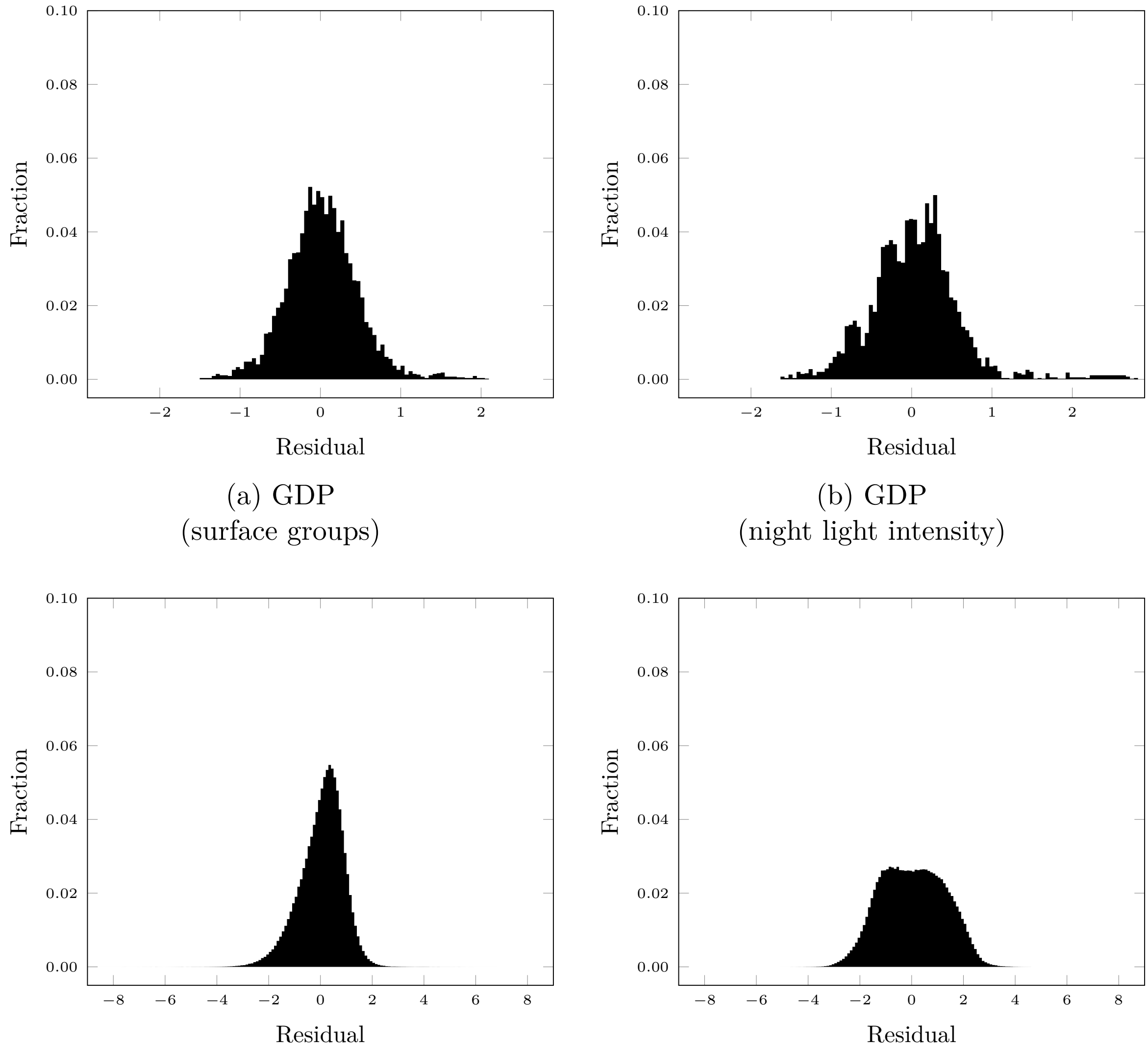
The value of surface groups as a proxy for regional economic activity becomes even more obvious at the very small regional level of grid cells. In a similar OLS analysis, our preferred surface-groups specification explains a much larger percentage of the variation in household income than the corresponding night-lights specification, with 67.5 percent vs. 30.7 percent (i.e., 119.9 percent higher precision).

The value of surface groups in comparison to night light intensity as a proxy for economic activity thus substantially increases with the degree of regional disaggregation. This finding is supported by an additional analysis in Appendix 3.5.2.3 that analyzes the prediction of county-level GDP by county-size category. On average, the surface groups explain a larger percentage of the variation in GDP for smaller counties than for larger counties.

Figure 3.2 underscores and visualizes these findings. It plots the statistical distribution of the OLS regression residuals, which are smaller when the measure is a better proxy for economic activity. The plots show that, for both GDP and household income, this distribution is smoother and narrower for surface groups (figures 3.2a and 3.2c) than for night light intensity (figures 3.2b and 3.2d). For household income, the residual distribution of the night-lights specification even exhibits a plateau—instead of a real peak—around the value zero, whereas the surface groups show a very clear peak and a narrow residual distribution.

Furthermore, we conduct five additional validation analyses in Appendix 3.5.2. First, we find that surface groups are a temporally and spatially less biased proxy for economic activity than night light intensity. This feature is important for the surface groups to serve as a valid proxy for comparisons of economic activity over time and between regions.

Figure 3.2: Statistical distribution of OLS regression residuals



(a) GDP
(surface groups)

(b) GDP
(night light intensity)

(c) Household income
(surface groups)

(d) Household income
(night light intensity)

Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, GFSO data (GENESIS), BKG data, RWI and microm (2019) data, and ESDAC data.
Note: See Appendix 3.5.2.3 for details on the regression specifications. Bin width of histograms is 0.05 in Figures 3.2a and 3.2b and 0.1 in Figures 3.2c and 3.2d.

Temporal bias would occur if the OLS residual is constant for a given region throughout all observation years, and spatial bias would occur if this residual is equal for clusters of regions. Surface groups yield a considerably smaller temporal bias that outweighs their somewhat larger spatial bias in comparison to night light intensity. Second, in line with their smaller bias, surface groups offer more information on within-region changes in economic activity than night light intensity through higher within-region heterogeneity. That is, our surface groups allow for a more precise determination of which subregional units drive the change in a region's economic activity by isolating the change in each subregional unit. Third, surface groups outperform also newer night light intensity data with higher spatial resolution in proxying economic activity. Fourth, we validate surface groups as a proxy for economic conditions in developing countries by comparing their predictive power to that of Yeh et al.'s (2020) prior metric of village asset wealth in Africa, a similar but more specialized outcome variable. This analysis shows that the validity of surface groups is not restricted to developed European countries such as Germany, but that surface groups can also provide valuable insights on economic conditions in developing countries across the world. Fifth, complementing our assessment of our proxy's internal validity, we find that surface groups can serve as valid proxies for their corresponding type of land cover as indicated in administrative statistics.

3.2.4 Surface Groups Economic Proxy

As having one single proxy may be desirable when economic activity is the dependent variable in an analysis, we compute predicted county-level GDP using our OLS model. To assess the external validity of this single-variable proxy, we use one half of the sample to train the coefficients showing the predictive power of our surface groups proxy, and then for the second half of our sample compute predicted GDP. Corroborating the results of the first analysis of external validity, GDP predicted using surface groups explains 63.4 percent of the variation in actual GDP in the second half of the sample, whereas GDP predicted using night light intensity explains only 48.8 percent of this variation (i.e., 29.9 percent higher precision). The corresponding values for household income are 67.4 percent using surface groups vs. 30.8 percent using night light intensity (i.e., 118.8 percent higher precision). However, when using the proxy as an independent variable, we recommend using the full set of proxy variables to minimize the noise and measurement error that might come from the predictive process.

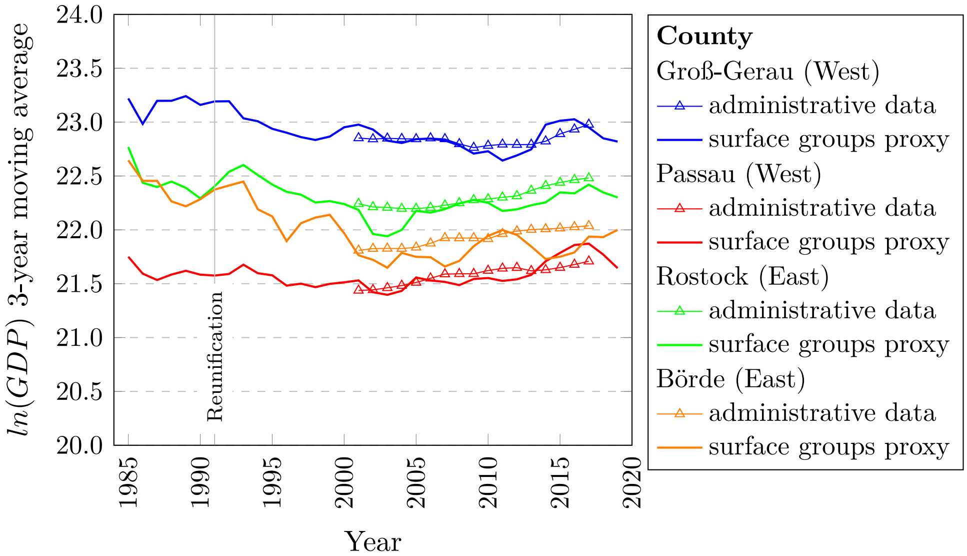
Finally, Figure 3.3 demonstrates the usefulness of our economic proxy. The curves marked with triangles show the extant data for regional economic activity in four regions of Germany—including areas in both East Germany (Rostock, Börde) and West Germany (Groß-Gerau, Passau). The thicker curves without triangles show our single-variable proxy for economic activity (with OLS coefficients trained on the entire sample) for its available years. Starting in 1984, the improved coverage achieved through surface groups almost doubles the number of available years compared to administrative data (which start in 2000 for Germany). Compared to other proxies such as night light intensity (which starts in 1992), surface groups are the only proxy pre-dating the German reunification. To better visualize trends over time, Figure 3.3 plots the three-year moving average of administrative GDP and the surface groups proxy. While the variation between years in the surface groups proxy is larger than in the administrative metric, all curves exhibit identical trends over time. The longer time series and the differences in the developments of the regions over time (e.g., GDP in Börde falls below that in Rostock after reunification) emphasize the proxy's potential for enabling previously impossible analyses.

3.3 Conclusion and Discussion

As Figure 3.3 demonstrates, the proxy we create from daytime satellite imagery is a strong proxy across time periods and across highly disaggregated regional levels, for which other data are unreliable, inaccessible, or entirely inexistent. Moreover, in this particular example, the proxy provides valuable, previously unavailable information on economic activity for East German regions before the fall of the iron curtain.

More generally, our procedure has worldwide relevance. While we apply our procedure to Germany and establish its validity for this country, the procedure is transferable to any

Figure 3.3: Time series of GDP measures in four counties



Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data. Notes: Plots show three-year moving averages. Curves marked by triangles show the natural logarithm of GDP measured in € in administrative data for all years for which county-level administrative GDP data are available. Thick curves without triangles show the surface groups proxy for GDP (predicted from OLS estimates, see appendix 3.5.2.5 for details). Groß-Gerau is situated in mid-west Germany, Passau in south Germany (at the border to Austria), Rostock in north Germany (at the Baltic Sea), and Börde in mid-north Germany.

region or country in the world (as we demonstrate in appendix 3.5.1.6). Our analyses for Germany exemplify that our machine-learning approach using daytime satellite imagery can predict both disaggregated and potentially missing or erroneous economic activity data (e.g., GDP at highly disaggregated levels within a country). However, the methodology and the data it provides for countries across the world can be extended globally to additional contexts where specific economic and developmental markers are needed. Our insight is to demonstrate that our methodology can be helpful for many economic and social science applications where varying degrees of disaggregation are required and where missing or incorrect data are prevalent. Surface groups thus constitute a valuable resource for analyzing historical developments, evaluating local policy reforms, and controlling for economic activity in econometric applications within a country. Although a country's history or industry structure affects the economic importance of different types of land cover (Henderson et al., 2018), the principle that land cover, which the surface groups reflect, relates to economic activity applies to any country in the world. Therefore, surface groups have a potential for economic research that investigates small regions within the same country or within a homogeneous group of countries.

The Landsat daytime satellite data are available for extremely small regional units such as municipalities or urban districts, thus providing new opportunities for urban and regional economic researchers to understand differences in even small regional variation in economic development. The surface groups we derive from these data thus contribute to analyses of the regional impacts of local policy reforms by providing information on economic activity at very detailed regional levels, for which other data sources are entirely unavailable for the necessary observation period, unreliable, less precise, or inaccessible for non-residents of the respective country. With these particular features, the surface groups complement other satellite-based measures for economic activity such as night light intensity.

The use of satellite data is a significant advancement in measuring regional economic activity and over time will generate new opportunities to strengthen our understanding of local economic conditions. For example, as the methods for utilizing daytime satellite data advance, researchers will eventually be able to analyze high-resolution satellite data with image recognition procedures to discern more about the nature of built-up surfaces. Such analyses could, for example, identify stores or industrial buildings, evaluate neighborhood housing quality, or determine when buildings have been renovated. However, retrieving more sophisticated metrics on economic activity requires satellite data with an even finer spatial resolution than Landsat data, such as the Advanced Spaceborne Thermal Emissions and Reflection Radiometer (ASTER) or the Sentinel mission. These or other satellite data offer promising venues for future research, for which this chapter lays first methodological foundations. However, the ASTER and Sentinel data cover substantially shorter time series than Landsat and are thus not as valuable for historical analyses.

3.4 Materials and Methods

3.4.1 Computation of Surface Groups

In developing our procedure for detecting surface groups, we follow the remote-sensing literature that has successfully applied machine-learning techniques to identifying, for example, built-up land cover from subsets of Landsat data (e.g., Liu et al., 2018; Schneider, 2012). Our procedure adds to this literature by combining data from four Landsat satellites to produce a time series of data on different types of land cover starting in 1984. We produce these data in GEE and apply supervised machine-learning techniques with the objective of classifying the annual type of land cover of every Landsat pixel location. We proceed in three steps that we shortly outline here and describe in detail in Appendix 3.5.1.

First, we prepare the Landsat data to retrieve the input data for the classification algorithm. We combine the data of Landsat-4, Landsat-5, Landsat-7, and Landsat-8 to produce composite data containing the qualitatively best observation per pixel location and year. In so doing, we choose those observations that best differentiate between vegetated and unvegetated areas, because we expect economic activity to concentrate in urban or industrial areas. The composite data constitute the input data that we pass on to the classification algorithm.

Second, to be able to classify observations in the input data, we add CLC data as an external source of ground-truth information. This ground-truth dataset comes from a pan-European project commissioned by the European Environment Agency (EEA) and maps land cover in 44 categories. To obtain a classification of land cover types that we can use to train our algorithm, we survey the literature that uses CLC data or Landsat data for classifying land cover (e.g., Balzter et al., 2015; Han et al., 2004; Neumann et al., 2007) and aggregate the 44 categories to larger groups with similar surface characteristics—the six surface groups. The classification algorithm requires this ground-truth information on surface groups for a subset of the input pixels to be able to recognize patterns in the input data and link these patterns to the different surface groups. By using external ground-truth data, we overcome the resource-intensive necessity of visually interpreting (i.e., manually classifying) input pixels to retrieve ground-truth information.

Third, we produce the training data for the classification algorithm. To obtain these training data, we draw a stratified random sample of pixels from the input data and match the ground-truth information on surface groups to the pixels in this sample. We then use the training data to train a Random Forest (RF) algorithm, which classifies every observation in the input data into one of the six surface groups.

Although we apply various filters for excluding invalid Landsat pixels (e.g., cloud shadow) from the composite input data, potentially erroneous pixel classifications might occur in few regions in years with scarce Landsat imagery (particularly in the 1980s). When applying our surface groups proxy in empirical analyses, we recommend removing outlier observations for these particular regions and years from these analyses. We do so in the comparison of county-level GDP developments in Figure 3.3. From 1984 through 2020, we identify 6.2% of all county-year observations as outliers. For more details on this outlier removal, see Appendix 3.5.2.5.

3.4.2 External Validity Analyses

We obtain two indicators of regional economic activity for the external validity analyses, which we shortly outline here and describe in more detail in Appendix 3.5.2.2. First, we use administrative GDP, which the German Federal Statistical Office (GFSO) provides at the county-level from 2000 and which we deflate for our analyses. Second, we use RWI-GEO-GRID (RWI and microm, 2019), a dataset containing socioeconomic indicators collected from a variety of public and private sources but annually only available from 2009. This dataset indicates household income at the level of grid cells sized one square kilometer, an extremely high level of regional detail. Again, we use deflated household income for our analyses.

In addition, to compare the quality of the surface groups as a proxy for economic activity to that of night light intensity, we use night lights data from the U.S. Air Force Defense Meteorological Satellite Program Operational Linescan System (DMSP OLS), available from 1992 through 2013. Similar to previous research (Chen and Nordhaus, 2011; Henderson et al., 2012), we use stable night lights (which are corrected for unusual lighting). To achieve regional correspondence with the administrative GDP data and RWI-GEO-GRID, we calculate average night light intensity at the county and at the grid level. In Appendix 3.5.2.4, we proceed similarly for comparing the surface groups proxy to Visible Infrared Imaging Radiometer Suite (VIIRS) night light intensity.

3.5 Appendix

This appendix provides the technical details of retrieving our surface groups measure as a proxy for economic activity. It presents all procedures and analyses referred to in this chapter and the underlying data. In Appendix 3.5.1, we describe the procedure we develop for retrieving the surface groups measure from daytime satellite imagery and conduct the internal validity analysis. In Appendix 3.5.2, we perform several analyses to demonstrate the value of surface groups as a proxy for economic activity.

3.5.1 Computation of Surface Groups

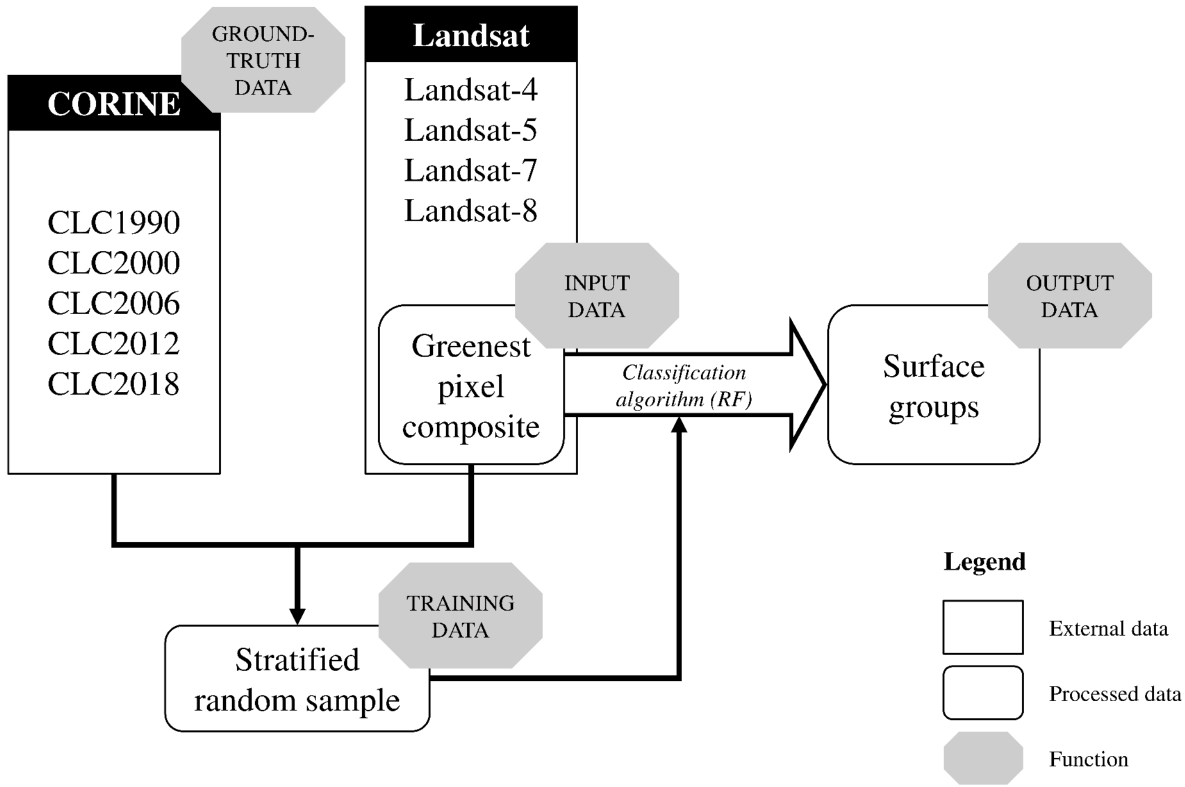
3.5.1.1 Overview

This appendix section describes our procedure for detecting surface groups as a novel proxy for economic activity at detailed regional levels. In developing this procedure, we follow the remote-sensing literature, which has successfully applied machine-learning techniques to identifying, for example, built-up land cover from subsets of Landsat data (e.g., Liu et al., 2018; Schneider, 2012). Our procedure adds to this literature by combining data from four Landsat satellites to produce a time series of data on different types of land cover starting in 1984. To produce these data, we use GEE as a platform and apply supervised machine-learning techniques with the objective of classifying the annual type of land cover of every Landsat pixel location in Germany. We proceed in four steps that Figure 3.4 illustrates.

First, we prepare the Landsat data to retrieve the input data for the classification algorithm. We combine the data of four Landsat satellites (Landsat-4, Landsat-5, Landsat-7, and Landsat-8) to produce composite data containing the qualitatively best observation per pixel location and year.[2] As we choose those observations that best differentiate between vegetated and unvegetated areas for this composite, we call it “greenest” pixel composite. This greenest pixel composite constitutes the input data that we pass on to

[2] We use the Landsat Collections distributed by the U.S. Geological Survey (USGS) (see USGS, 2018) and directly accessible through GEE.

Figure 3.4: Overview of procedure for detecting surface groups



Source: Authors' illustration.

the classification algorithm.

Second, to be able to classify the observations in the greenest pixel composite, we add CLC data as an external source of ground-truth information. These data, which come from a pan-European project commissioned by the EEA,[3] map land cover in European countries for five reference years (1990, 2000, 2006, 2012, 2018). Based on a survey of the literature that applies land cover classifications (e.g., Waser and Schwarz, 2006; Yu et al., 2011), we obtain from the CLC data the six different types of land cover that we call "surface groups": built-up surfaces (*builtup*), grassy surfaces (*grass*), surfaces with crop fields (*crops*), forest-covered surfaces (*forest*), surfaces without vegetation (*noveg*), and water surfaces (*water*). The classification algorithm requires this ground-truth information on surface groups to be able to recognize patterns in the input data and link these patterns to the different surface groups. For example, the spectral values of an input pixel showing a grassy surface differ from those of an input pixel showing a built-up surface. The CLC data provide the classification algorithm with the true surface group for a subset of the input pixels. By using external ground-truth data, we overcome the resource-intensive necessity of visually interpreting (i.e., manually classifying) input pixels to retrieve ground-truth information.

Third, we produce the training data for the classification algorithm. To obtain these training data, we draw a stratified random sample of pixels from the greenest pixel composite and match the CLC ground-truth information on surface groups to the pixels in this sample. We then use the training data to train the classification algorithm, which is a RF algorithm with ten decision trees. After training the algorithm, it classifies every observation in the greenest pixel composite into one of the six surface groups.

Fourth, the classification result is the output data that contain the surface group of every Landsat pixel location annually from 1984 through 2020. To assess the accuracy of the classification (i.e., the internal validity), we perform five-fold cross-validation.

[3] The CLC data are distributed by the EEA (see EEA, 2020) and directly accessible through GEE.

3.5.1.2 Greenest Pixel Composite of Landsat Data as Input Data

Satellite data from the Landsat program serve as input data for the machine-learning procedure for detecting surface groups. Since 1972, Landsat satellites have continuously recorded remotely sensed imagery of the earth, providing a unique basis for various applications in mapping and monitoring land cover (Wulder et al., 2012, 2008). Throughout the history of Landsat, the various operating agencies have launched eight satellites, one of which (Landsat-6) failed to reach orbit (Williams et al., 2006; Wulder et al., 2016). As of 2022, Landsat-7, Landsat-8, and Landsat-9 remain active, with Landsat-9 having launched only in September 2021 (Lulla et al., 2021; Masek et al., 2020).[4]

We gather the input data for our algorithm to detect surface groups from the spectral information that Landsat satellites capture. Every Landsat satellite carries sensors that remotely measure the spectral reflectance of the earth's surface (Markham et al., 2004). The improving technical specifications of these sensors from one satellite generation to the next entail an increase in the number of spectral bands that each satellite captures (Markham and Helder, 2012). Table 3.2 provides the technical specifications of the different sensors that Landsat satellites carry, including their spectral resolution, years of operation, and wavelengths of the spectral bands that the sensors capture.

We use information in the six spectral bands that the sensors on Landsat-4, Landsat-5, Landsat-7, and Landsat-8 have in common (highlighted gray in table 3.2). These bands contain the surface reflectance in the visible blue (BLUE), visible green (GREEN), visible red (RED), short-wave infrared (SWIR1 and SWIR2), and near-infrared (NIR) ranges of the electromagnetic spectrum. Consequently, we begin our observation period with the 1982 launch of Landsat-4. However, due to a series of technical failures throughout the lifetime of Landsat-4 (Rumerman, 1999) and the resulting scarcity of Landsat-4 imagery for Germany, the effective start of our observation period is 1984 (although we include Landsat-4 imagery in later years whenever available).

We exclude imagery from the pre-Landsat-4 period and information in the thermal

[4] The remote-sensing literature and related disciplines have applied Landsat data for numerous purposes, for example, the assessment of water conditions in the Bahamas (Lyzenga, 1981) and the investigation of tree species diversity in the Alps (Torresani et al., 2019).

Table 3.2: Technical specifications of Landsat sensors

Sensor	Multispectral Scanner (MSS)	Thematic Mapper (TM)	Enhanced Thematic Mapper Plus (ETM+)	Operational Land Imager (OLI) / Thermal Infrared Sensor (TIRS)
Spatial resolution	79 meters	30 meters	30 meters	30 meters
Satellites (Operating Years)	Landsat-1 (1972–1978) Landsat-2 (1975–1982) Landsat-3 (1978–1983) Landsat-4 (1982–1993) Landsat-5 (1984–1995)	Landsat-4 (1982–1993) Landsat-5 (1984–2014)	Landsat-7 (1999–present)	Landsat-8 (2013–present)
Band name	Wavelength (in µm)			
Ultra blue				0.43–0.45
Visible blue (BLUE)		0.45–0.52	0.45–0.52	0.45–0.51
Visible green (GREEN)	0.50–0.60	0.52–0.60	0.52–0.60	0.53–0.59
Visible red (RED)	0.60–0.70	0.63–0.69	0.63–0.69	0.64–0.67
Short-wave infrared 1 (SWIR1)		1.55–1.75	1.55–1.75	1.57–1.65
Short-wave infrared 2 (SWIR2)		2.08–2.35	2.08–2.35	2.11–2.29
Near-infrared 1 (NIR)	0.70–0.80	0.76–0.90	0.77–0.90	0.85–0.88
Near-infrared 2	0.80–1.10			
Thermal infrared 1		10.40–12.50 (120-meter resolution)	10.40–12.50 (60-meter resolution)	10.60–11.19 (100-meter resolution)
Thermal infrared 2				11.50–12.51 (100-meter resolution)
Panchromatic			0.52–0.90 (15-meter resolution)	0.50–0.68 (15-meter resolution)
Cirrus				1.36–1.38

Source: Authors' representation based on Goldblatt et al. (2016); Loveland and Dwyer (2012); Markham et al. (2004); Roy et al. (2014); Williams et al. (2006); Wulder et al. (2016). Note: Spectral bands used for detecting surface groups highlighted gray. The table excludes technical details that are beyond the scope of this chapter. For example, the MSS sensor of Landsat-5 was decommissioned in 1995, but the MSS archives contain data only until the sensor became unable to relay data in 1992 (Loveland and Dwyer, 2012; Wulder et al., 2016). OLI and TIRS, the sensors that Landsat-8 carries, are two separate sensors, with the TIRS capturing the two thermal infrared bands and OLI the remaining ones (Roy et al., 2014).

infrared spectral bands for the following reasons. We exclude pre-Landsat-4 satellites because they differ substantially from their successors in captured wavelength and in spatial resolution (Morain, 1998). Therefore, when combining all sensors, we cannot achieve a consistent pixel classification, which is a prerequisite for a valid economic measure. Moreover, due to technological and organizational constraints at the time, imagery in the Landsat archives is scarce for Germany until the 1980s (Wulder et al., 2016). This scarcity of imagery makes the detection of surface groups unfeasible for the pre-Landsat-4 period, regardless of the sensors the satellites carried. Furthermore, we do not use the thermal infrared spectral bands because their technical specifications change over time and differ from the remaining bands (e.g., coarser spatial resolution, different numbers of bands, see table 3.2). In addition, the bands' specifications notwithstanding, temperatures in Germany vary over the seasons so that thermal information would be of little help for detecting surface groups.

As with the night light intensity data that economists commonly use (Donaldson and Storeygard, 2016), we compute the surface groups annually. As Landsat satellites record a geographic location on earth multiple times per year (Wulder et al., 2016), we have to use annual composites of these records. Unfortunately, pre-processed annual composites incorporating imagery from multiple Landsat satellites do not exist, requiring us to produce such composites from the available images and use these composites as input data for our algorithm.

We produce pixel-based annual composites of Landsat images. Among all available observations of a given pixel within a year, we choose the one pixel that best serves the purpose of detecting surface groups. This pixel-based compositing procedure (as compared to scene-based compositing) prevents a loss of information due to, for example, cloud-covered pixels and enables the researcher to choose those pixels best suitable for a specific application (Griffiths et al., 2013)—in our case, the detection of surface groups. Given the long time span that we analyze, the production of annual composites also entails less computational effort than other approaches such as data stacking (Trianni et al., 2015).

For both the compositing and the actual pixel classification (see appendix 3.5.1.4), we follow studies from the remote-sensing literature (e.g., Goldblatt et al., 2016; Liu et al., 2018) and add three indices to the data: First, the Normalized Difference Vegetation Index (NDVI) differentiates vegetated from unvegetated surfaces and is one of the most frequently used indices in the remote-sensing literature (Rouse Jr et al., 1973; Xue and Su, 2017); Second, the Normalized Difference Water Index (NDWI) differentiates open water from other surfaces (McFeeters, 1996);[5] Third, the Normalized Difference Built-up Index (NDBI) differentiates built-up surfaces from other surfaces (Zha et al., 2003). Similar to Liu et al. (2018), we compute these three indices for Landsat data as follows:

$$NDVI_p = \frac{NIR_p - RED_p}{NIR_p + RED_p} \tag{3.1}$$

$$NDWI_p = \frac{GREEN_p - NIR_p}{GREEN_p + NIR_p} \tag{3.2}$$

$$NDBI_p = \frac{SWIR1_p - NIR_p}{SWIR1_p + NIR_p} \tag{3.3}$$

with p denoting pixels as the unit of observation.

For the compositing of Landsat images, we proceed in three steps. First, we collect all images available within a given calendar year for Germany, our study region. We restrict the pool of images to those taken between March and November, that is, we exclude the meteorological winter months in the northern hemisphere. We do so because the potential snow cover and the absence of large parts of the vegetation during winter might confuse the machine-learning algorithm. Second, we drop pixels showing clouds or cloud shadow and pixels with implausible values in one of the spectral bands. Clouds obscure the actual surface we want to observe, and cloud shadow distorts a pixel's actual reflectance, whereas a pixel with clear vision does not (e.g. Zhu and Woodcock, 2012). Implausible values, such as a negative reflectance in one of the spectral bands, might result from erroneous data transmission. Third, among the remaining pixels we choose the best one available. In

[5] Another index exists under the name "NDWI", which was developed to identify liquid water inside plants (Gao, 1996). This other NDWI relies on different spectral bands than the NDWI we use.

so doing, we emphasize the distinction of built-up land from other surfaces, because—as with the logic underlying the use of night light intensity as a proxy for GDP—we expect economic activity to concentrate in urban or industrial areas. Therefore, a clear distinction between built-up surfaces and other surfaces will improve our proxy for economic activity.

Our procedure of compositing Landsat data provides us with a greenest pixel composite that we can use as input data for the machine-learning algorithm. This composite covers the geographical area of Germany and consists of one observation per pixel for every year since 1984. The variables in the dataset are the pixel's values in the six spectral bands we use in this chapter (see table 3.2) and the added indices NDVI, NDWI, and NDBI. If the compositing procedure cannot identify a valid observation for a pixel location within a calendar year (e.g., if all available pixels show clouds), the data contain missing values. Figure 3.1a visualizes the greenest pixel composite with the visible spectral bands BLUE, GREEN, and RED for 2018.

3.5.1.3 CLC Data as Ground-Truth Data

To retrieve ground-truth information for a subset of the greenest pixel composite, we use CLC data. The European Commission began the CORINE[6] program that produces these data in 1985, with the goal of creating a standardized database on land cover to support policymakers in environmental affairs (Büttner et al., 2002; EEA, 2017). Since then, five phases of the program have produced CLC databases for the five reference years 1990, 2000, 2006, 2012, and 2018 (hereafter denoted as CLC1990, CLC2000, CLC2006, CLC2012, and CLC2018) (Büttner and Kosztra, 2017). Each database includes a map for the respective year with a pixel resolution of 100 meters, indicating land cover in a variety of classes (Büttner and Kosztra, 2017; Kosztra et al., 2019).

Although the medium underlying the classification changed over the years from hard-copies to computer-assisted technologies, classification still relies mainly on visual interpretation of satellite imagery by professional experts (Büttner and Kosztra, 2017; Kosztra et al., 2019). This imagery stems from various satellites, including Landsat satellites for

[6] The acronym "CORINE" stands for "coordination of information on the environment" (Büttner et al., 2002).

CLC1990, CLC2000, and CLC2018 (Büttner and Kosztra, 2017). The remote-sensing literature provides successful combinations of CLC and Landsat data in geospatial analyses (e.g., Matejicek and Kopackova, 2010; Pekkarinen et al., 2009).

To train our machine-learning algorithm, we exploit the CLC data as a source of ground-truth information for three reasons. First, the earliest of the CLC data's five reference years (1990) still falls within the operating time of Landsat-4 (1982–1993), the oldest Landsat satellite we use in our computations (see appendix 3.5.1.2). This time overlap improves the prediction of surface groups by providing a better temporal fit of ground-truth data and input data. Second, although with 100 meters the spatial resolution of CLC pixels is lower than that of Landsat pixels, CLC pixels still have a much higher resolution than other external ground-truth data used in the remote-sensing literature (e.g., night light intensity data with a resolution of one kilometer in Goldblatt et al., 2018). This high resolution improves the prediction of surface groups by providing a better spatial fit of ground-truth data and input data. Third, the CLC data provide a detailed classification of surfaces, allowing us to distinguish between various types of surfaces, such as built-up land, forests, or water. In sum, the CLC data constitute an excellent external source of ground-truth information for the purpose of detecting surface groups.

The CLC classification consists of five larger groups (level 1), which are further subdivided into 15 subgroups (level 2) and 44 detailed groups (level 3). However, even at levels 1 and 2, this classification simultaneously indicates types of land cover (the land's directly observable terrestrial features) and land use (the land's socioeconomic purpose) (Cihlar and Jansen, 2001; Comber et al., 2008; Feranec et al., 2007; Fisher et al., 2005). Given that automated analyses of satellite data can detect only land cover and that determining land use requires manual interpretation (Fisher et al., 2005), we cannot directly apply this classification for the training of our algorithm.

To obtain a classification of land cover types that we can use to train our algorithm, we aggregate the CLC level 3 classes to larger groups with similar surface characteristics. We base this aggregation on a survey of the literature that uses CLC data or Landsat data for classifying land cover (e.g., Balzter et al., 2015; Han et al., 2004). However, as this literature does not provide an unambiguous assignment of CLC classes to larger groups with similar surface characteristics, we perform repeated trials of our classification procedure with varying assignments of CLC level 3 classes to larger groups. These trials yield the result that a classification consisting of six surface groups, which correspond to the six types of surfaces Yu et al. (2011) identify from subsets of Landsat data for the Daqing region in China, best represents similar surface characteristics in Germany:

- Built-up surface (*builtup*): The surface group *builtup* contains surfaces with buildings of non-natural materials such as concrete, metal, and glass (e.g., residential buildings, industrial plants, roads). This surface group thus includes all artificial surfaces (CLC class 1) except for green urban areas (CLC class 141), which Neumann et al. (2007) show to have a lower resemblance with artificial surfaces than with vegetated surfaces.
- Grassy surfaces (*grass*): The surface group *grass* contains surfaces covered by grass or other plants with similar surface reflectance (e.g., natural grassland, city parks). This surface group thus includes pastures (CLC class 23) and natural grassland (CLC class 321), which have similar surface characteristics (Neumann et al., 2007; Waser and Schwarz, 2006). In addition, due to the similarities in surface reflectance that we detect in our trials, we add to the surface group *grass* the green urban areas (CLC class 141) that we exclude from the surface group *builtup*.
- Surfaces with crop fields (*crops*): The surface group *crops* contains surfaces with vegetation for agricultural purposes (e.g., hayfields, vineyards). This surface group thus includes all agricultural areas (CLC class 2) except for pastures (CLC class 23), which belong to the surface group *grass*.
- Forest-covered surfaces (*forest*): The surface group *forest* contains surfaces covered by trees or other plants with similar surface reflectance (e.g., mixed forests, moors). This surface group thus includes all forests and semi-natural areas (CLC class 3) except for grassland (CLC class 321), which belongs to the surface group *grass*, and open spaces with little or no vegetation (CLC class 33), which differ in spectral reflectance from the remaining CLC classes in the surface group *forest* (Neumann

et al., 2007; Pérez-Hoyos et al., 2012)

- Surfaces without vegetation (*noveg*): The surface group *noveg* contains surfaces with (almost) no vegetation or buildings (e.g., bare rock, sand plains). This surface group thus includes open spaces with little or no vegetation (CLC class 33), which we exclude from the surface group *forest*.
- Water surfaces (*water*): The surface group *water* contains any type of water surface (e.g., rivers, lakes). This surface group thus includes wetlands (CLC class 4) and water bodies (CLC class 5), which we aggregate following Gallego and Bamps (2008).

Table 3.3 provides a correspondence table of CLC classes for our algorithm. These six surface groups into which the classification algorithm divides the input data constitute the basis for our proxy for economic activity.[7] Figure 3.1b visualizes the ground-truth surface groups that we obtain from the CLC2018 data.

3.5.1.4 Training Data and Classification Algorithm

We apply a machine-learning algorithm that classifies the input data of the greenest pixel composite into the six surface groups *builtup*, *grass*, *crops*, *forest*, *noveg*, and *water*. From the input data, we draw a stratified random sample of pixels to train the algorithm and retrieve the corresponding ground-truth information from CLC data. The classifier we use is a RF algorithm with ten decision trees.

Following Goldblatt et al. (2016), we perform pixel-based classification. For every pixel in our training sample, the machine-learning algorithm predicts the pixel's surface group from the spectral values and the added indices NDVI, NDWI, and NDBI. Compared to object-based classification, which additionally considers information from neighboring pixels, pixel-based classification requires considerably less computational power (Myint et al., 2011; Whiteside et al., 2011). Although the majority of studies in the remote-sensing literature suggest that object-based classification performs better than pixel-based classification (e.g., Whiteside et al., 2011), some studies find no significant performance

[7] Researchers who are interested in investigating more specific types of land cover, such as different types of forests or water bodies, might incorporate the findings from the literature classifying such types of land cover (e.g., Boresjö Bronge and Näslund-Landenmark, 2002; Traustason and Snorrason, 2008).

Table 3.3: CLC classes and assignment for algorithm

CLC class Level 1		Level 2		Level 3		Our algorithm
1	Artificial surfaces	11	Urban fabric	111	Continuous urban fabric	*builtup*
				112	Discontinuous urban fabric	*builtup*
		12	Industrial, commercial, and transport units	121	Industrial or commercial units and public facilities	*builtup*
				122	Road and rail networks and associated land	*builtup*
				123	Port areas	*builtup*
				124	Airports	*builtup*
		13	Mine, dump, and construction sites	131	Mineral extraction sites	*builtup*
				132	Dump sites	*builtup*
				133	Construction sites	*builtup*
		14	Artificial, non-agricultural vegetated areas	141	Green urban areas	*grass*
				142	Sport and leisure facilities	*builtup*
2	Agricultural areas	21	Arable land	211	Non-irrigated arable land	*crops*
				212	Permanently irrigated arable land	*crops*
				213	Rice fields	*crops*
		22	Permanent crops	221	Vineyards	*crops*
				222	Fruit tree and berry plantations	*crops*
				223	Olive groves	*crops*
		23	Pastures	231	Pastures, meadows, and other permanent grasslands under agricultural use	*grass*
		24	Heterogeneous agricultural areas	241	Annual crops associated with permanent crops	*crops*
				242	Complex cultivation patterns	*crops*
				243	Land principally occupied by agriculture, with significant areas of natural vegetation	*crops*
				244	Agro-forestry areas	*crops*
3	Forest and semi-natural areas	31	Forests	311	Broad-leaved forest	*forest*
				312	Coniferous forest	*forest*
				313	Mixed forest	*forest*
		32	Shrubs and/or herbaceous vegetation associations	321	Natural grassland	*grass*
				322	Moors and heathland	*forest*
				323	Sclerophyllous vegetation	*forest*
				324	Transitional woodland/shrub	*forest*
		33	Open spaces with little or no vegetation	331	Beaches, dunes, and sand plains	*noveg*
				332	Bare rock	*noveg*
				333	Sparsely vegetated areas	*noveg*
				334	Burnt areas	*noveg*
				335	Glaciers and perpetual snow	*noveg*
4	Wetlands	41	Inland wetlands	411	Inland marshes	*water*
				412	Peatbogs	*water*
		42	Coastal wetlands	421	Coastal salt marshes	*water*
				422	Salines	*water*
				423	Intertidal flats	*water*
5	Water bodies	51	Inland waters	511	Water courses	*water*
				512	Water bodies	*water*
		52	Marine waters	521	Coastal lagoons	*water*
				522	Estuaries	*water*
				523	Sea and ocean	*water*

Source: Authors' illustration based on Heymann et al. (1994, p. 27). CLC classes listed as in Kosztra et al. (2019).

difference (e.g., Duro et al., 2012), and in particular, Dingle Robertson and King (2011) find no significant difference using Landsat data.[8] Therefore, given the spatial and temporal size of the data we analyze in this chapter, pixel-based classification is the preferable choice. Our assessments of external validity confirm that choosing this classification yields a valid proxy for economic activity.

To classify the pre-processed Landsat data, we use the RF algorithm with ten decision trees.[9] Several studies in the remote-sensing literature find that RF outperforms other algorithms when applied to land cover classification (e.g., Gislason et al., 2006; Rodriguez-Galiano et al., 2012). For example, Goldblatt et al.'s (2016) assessment of the performance of three different algorithms that the remote-sensing literature commonly uses (Classification and Regression Tree, Support Vector Machines, and RF) reveals that RF performs best in predicting built-up land cover in India with Landsat-7 and Landsat-8 data. Furthermore, RF requires less computational power (Gislason et al., 2006). As to the number of decision trees, performance increases with the number of trees, although after ten trees the increase is negligibly small relative to the increase in computational power required (Goldblatt et al., 2016).[10] Therefore, RF with ten decision trees best suits the purpose of this chapter.

We draw a stratified random sample of a total of 30,000 pixels to serve as training data for the classification algorithm. For every year in the CLC data (1990, 2000, 2006, 2012, 2018), we randomly choose 1,000 pixels of each surface group. Generally, the number of pixels in the training data correlates positively with prediction accuracy but negatively with computational effort (Millard and Richardson, 2015; Rodriguez-Galiano et al., 2012). Therefore, we choose a slightly larger number of pixels in the training data than in comparable applications from the remote-sensing literature (e.g., Goldblatt et al., 2016; Schneider, 2012) to achieve an accurate classification, but keep this number low enough to

[8] See, e.g., Ma et al. (2017) for a review of the literature on the advantages and disadvantages of pixel-based vs. object-based classification.

[9] For a description of the RF method's application for land cover classification, see, e.g., Gislason et al. (2006), and for a description of the method's application in economics, see, e.g., Athey and Imbens (2019).

[10] For example, while the prediction's overall accuracy increases by about two percentage points when increasing the number of trees from three to ten, it increases only by about one more percentage point when increasing the number of trees from ten to 100 (Goldblatt et al., 2016).

maintain a reasonable computational effort.

By restricting the pool of pixels from which we draw the stratified random sample we use as training data, we substantially reduce the influence of potentially imprecise ground-truth observations resulting from the difference in spatial resolution between Landsat (30 meters) and CLC data (100 meters). Due to the coarser CLC resolution, the CLC surface groups might not be accurate for Landsat pixels at the boundary of two CLC surface areas. While a CLC pixel might correctly belong to the *builtup* surface group, because more than half of the pixel's area contains built-up surfaces, not all Landsat pixels that fall within the CLC pixel are necessarily *builtup*. Therefore, we do not use Landsat pixels that fall within CLC pixels at the boundary of two CLC surface areas as training data, that is, the CLC pixel and all its neighboring pixels must belong to the same surface group. This restriction reduces the number of imprecise ground-truth observations and thus improves the quality of the training data, which correlates positively with the accuracy of the RF prediction when applied to land-cover classification (Mellor et al., 2015; Rodriguez-Galiano et al., 2012; Rogan et al., 2008).

A comparison of Figures 3.1a and 3.1b (right column) in Section 3.2.1 illustrates the reason for the sampling restriction to inner CLC pixels. For example, we do not use the CLC pixels at the boundary of the *water* and *grass* surface areas in Figure 3.1b. At this boundary, some of the CLC *water* pixels contain parts of the vegetation at the lakeshore, and, vice versa, some of the CLC *grass* pixels contain parts of the lake. Therefore, the boundary CLC pixels are not representative for the true surface groups of the Landsat pixels that fall within these CLC pixels. Excluding the unrepresentative ground-truth information reduces the risk of imprecise ground-truth information in the training data. The resulting benefit of this exclusion is that the algorithm can more accurately classify unrepresentative pixels (e.g., the *forest* and *grass* pixels that belong to the nature reserve (Vogelinsel im Altmühlsee) in the south-west of Figure 3.1c, as well as similar examples throughout Germany).

3.5.1.5 Accuracy Assessment of Output Data

To assess the prediction accuracy of our classification in the output data, we follow Goldblatt et al. (2016) and perform five-fold cross-validation[11] by drawing five subsets from the greenest pixel composite. In drawing the subsets, we apply the same stratification criteria as for the training dataset, with the only difference being that instead of 1,000 pixels per surface group, we now draw only 250. Thus each of the five subsets consists of 7,500 pixels, that is, 250 per surface group and year. For the cross-validation to be valid, the subsets must not overlap. In other words, one pixel can belong to only one subset.

Next, imitating our procedure for generating the output data, we use the five subsets to perform five iterations of pixel classification. During each iteration, we use four of the subsets as a training set. Consequently, every iteration leaves out a different subset, and the training set of four subsets includes precisely the same number of pixels as the training set we actually use for the computations. We train the classification algorithm with the four-subset training set, then classify the left-out subset.

As indicators of prediction accuracy, for every iteration and for each of the six surface groups separately, we calculate overall accuracy, true-positive rate, true-negative rate, balanced accuracy, and user's accuracy (see section 3.2.2). Complementing the five-fold cross-validation results for the entire sample in Table 3.1, Tables 3.4 through 3.9 show the results separately for every CLC year.

The five-fold cross-validation results show that our output data constitute an internally valid measure of land cover. All indicators of prediction accuracy reveal that our classification algorithm accurately identifies the six surface groups, suggesting that we adequately implemented the procedures from the remote-sensing literature. Therefore, the output data of our algorithm is highly suitable for analyzing whether the surface groups are an externally valid proxy for economic activity in Appendix 3.5.2.[12]

[11] For descriptions and discussions of this method, see, e.g., Arlot and Celisse (2010) and Wong (2015)

[12] Additional analyses on the correlation between surface groups and administrative measures of land cover also reveal that surface groups validly indicate their corresponding type of land cover in administrative statistics (see appendix 3.5.2.6).

Table 3.4: Five-fold cross-validation results with respect to built-up surfaces (surface group *builtup*)

	Overall accuracy	True-positive rate	True-negative rate	Balanced accuracy	User's accuracy
1990	0.827	0.664	0.859	0.761	0.481
2000	0.838	0.610	0.886	0.748	0.530
2006	0.838	0.593	0.897	0.745	0.585
2012	0.844	0.606	0.887	0.747	0.493
2018	0.795	0.558	0.853	0.705	0.482
Average	0.828	0.606	0.877	0.741	0.514

Source: Authors' calculations based on Landsat data and CLC data.
Notes: The yearly values indicate the average over all five iterations within the respective year. Average indicates the average over the yearly values as indicated in Table 3.1.

Table 3.5: Five-fold cross-validation results with respect to grassy surfaces (surface group *grass*)

	Overall accuracy	True-positive rate	True-negative rate	Balanced accuracy	User's accuracy
1990	0.839	0.468	0.912	0.690	0.514
2000	0.826	0.513	0.891	0.702	0.495
2006	0.823	0.517	0.888	0.703	0.496
2012	0.852	0.428	0.937	0.682	0.575
2018	0.813	0.327	0.921	0.624	0.477
Average	0.831	0.451	0.910	0.680	0.511

Source: Authors' calculations based on Landsat data and CLC data.
Notes: The yearly values indicate the average over all five iterations within the respective year. Average indicates the average over the yearly values as indicated in Table 3.1.

Table 3.6: Five-fold cross-validation results with respect to surfaces with crop fields (surface group *crops*)

	Overall accuracy	True-positive rate	True-negative rate	Balanced accuracy	User's accuracy
1990	0.816	0.461	0.889	0.675	0.458
2000	0.847	0.416	0.937	0.677	0.583
2006	0.828	0.396	0.931	0.664	0.581
2012	0.852	0.348	0.955	0.652	0.611
2018	0.817	0.281	0.950	0.615	0.583
Average	0.832	0.381	0.932	0.657	0.563

Source: Authors' calculations based on Landsat data and CLC data.
Notes: The yearly values indicate the average over all five iterations within the respective year. Average indicates the average over the yearly values as indicated in Table 3.1.

Table 3.7: Five-fold cross-validation results with respect to forest-covered surfaces (surface group *forest*)

	Overall accuracy	True-positive rate	True-negative rate	Balanced accuracy	User's accuracy
1990	0.892	0.509	0.969	0.739	0.771
2000	0.899	0.725	0.936	0.830	0.701
2006	0.886	0.756	0.914	0.835	0.648
2012	0.901	0.711	0.940	0.825	0.713
2018	0.895	0.726	0.933	0.830	0.709
Average	0.895	0.685	0.938	0.812	0.708

Source: Authors' calculations based on Landsat data and CLC data.
Notes: The yearly values indicate the average over all five iterations within the respective year. Average indicates the average over the yearly values as indicated in Table 3.1.

Table 3.8: Five-fold cross-validation results with respect to surfaces without vegetation (surface group *noveg*)

	Overall accuracy	True-positive rate	True-negative rate	Balanced accuracy	User's accuracy
1990	0.890	0.732	0.921	0.827	0.652
2000	0.891	0.754	0.913	0.834	0.585
2006	0.894	0.644	0.918	0.781	0.434
2012	0.850	0.847	0.850	0.849	0.543
2018	0.827	0.801	0.829	0.815	0.236
Average	0.870	0.756	0.886	0.821	0.490

Source: Authors' calculations based on Landsat data and CLC data.
Notes: The yearly values indicate the average over all five iterations within the respective year. Average indicates the average over the yearly values as indicated in Table 3.1.

Table 3.9: Five-fold cross-validation results with respect to water surfaces (surface group *water*)

	Overall accuracy	True-positive rate	True-negative rate	Balanced accuracy	User's accuracy
1990	0.908	0.683	0.952	0.817	0.740
2000	0.902	0.623	0.959	0.791	0.759
2006	0.915	0.693	0.961	0.827	0.787
2012	0.915	0.692	0.959	0.826	0.768
2018	0.905	0.667	0.957	0.812	0.772
Average	0.909	0.672	0.958	0.815	0.765

Source: Authors' calculations based on Landsat data and CLC data.
Notes: The yearly values indicate the average over all five iterations within the respective year. Average indicates the average over the yearly values as indicated in Table 3.1.

3.5.1.6 Transfer to All Countries Across the World

Producing our surface groups proxy depends on two external datasets—Landsat imagery (to retrieve the greenest pixel composite as input data) and CLC data (ground-truth data). While Landsat data are available for the entire world,[13] consistent ground-truth data are not. As such, we use two different strategies—one covering the European countries included in the CLC data (CLC countries)[14] and another covering the rest of the world (non-CLC countries)—to retrieve ground-truth data.

Our procedure for detecting surface groups is straightforwardly transferable to CLC countries (i.e., most European countries). For these countries, CLC data include comprehensive and consistent ground-truth information. Therefore, producing the surface groups for any given CLC country works exactly as for our German example. As the data do not cover 1990 (the first of the five CLC reference years) for a few CLC countries, we make one adjustment to the training-sample construction for these countries.[15] In the stratified random sample to serve as training data, we randomly draw 1,250 instead of 1,000 pixels per surface group and year. Consequently, as for CLC countries that cover all five reference years, the training data comprise a total of 30,000 pixels (thus a size identical to that for CLC countries with ground-truth data for all five reference years).

We address the challenge in producing our proxy for non-CLC countries—the selection of adequate ground-truth data from which to draw the training sample—through a selection rule based on the Köppen-Geiger climate classification system (Beck et al., 2018; Köppen, 1884). At the highest level of aggregation, this system differentiates between five climate zones of the world: tropical (zone A), arid (zone B), temperate (zone C), cold (zone D), and

[13] For example, Landsat-7 covers any region between the 81.8° north and south latitudes, thus not covering uninhabited places such as Antarctica and the far northern part of Greenland (Bindschadler, 2003)

[14] Countries included in the CLC data are Albania, Andorra, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Gibraltar, Greece, Hungary, Iceland, Ireland, Italy, Kosovo, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Montenegro, The Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, and the United Kingdom (see Copernicus Land Monitoring Service, 2020).

[15] CLC countries without data for the reference year 1990 are Albania, Bosnia and Herzegovina, Cyprus, Finland, Iceland, Kosovo, North Macedonia, Norway, Sweden, Switzerland, and the United Kingdom (see Copernicus Land Monitoring Service, 2020).

polar (zone E) (Beck et al., 2018).[16] When classifying surface groups for non-CLC countries, we calculate which percentage of a country's area falls within each of the five climate zones. We then draw a random sample of 30,000 pixels (same size as in the procedure for CLC countries) from all available CLC data (i.e., from all countries participating in CORINE) and stratify the pixel selection by climate zone, that is, for each climate zone the percentage of pixels in the training sample belonging to that climate zone corresponds to the target country's percentage of pixels belonging to that climate zone. For example, if 30 percent of a country's area belong to climate zone C and the remaining 70 percent to climate zone D, the training sample will consist of 9,000 pixels from climate zone C and 21,000 pixels from climate zone D. All other stratification criteria for CLC countries (e.g., same number of pixels per surface group and CLC year) also apply for non-CLC countries.

As none of the CLC countries features the tropical climate zone A, we assign the percentage of a non-CLC target country's area in climate zone A (if any) to CLC pixels in the temperate climate zone C. We do so, because climate zone C is most similar to climate zone A in terms of vegetation (the main selection criterion in constructing our greenest pixel composite from Landsat data as input data). As we restrict the pool of Landsat images for constructing the greenest pixel composite and, consequently, the training data to those images taken between March and November (thus excluding the meteorological winter months in the northern hemisphere), climate zones A and C also have similar temperature levels during the period we consider.

As in the procedure for CLC countries, we exclude Landsat images taken during meteorological winter months in constructing the greenest pixel composite for non-CLC countries. For non-CLC countries in the northern hemisphere, we exclude images taken between December and February (similar to the exclusion for CLC countries), while we exclude images taken between June and August for non-CLC countries in the southern hemisphere. For countries within the Tropic of Cancer and the Tropic of Capricorn we do not exclude any images, because temperatures (and thus vegetation) in these countries stay almost constant over the seasons.

[16] Beck et al.'s (2018) Köppen-Geiger climate classification data are publicly available at https://doi.org/10.6084/m9.figshare.6396959.

The procedure for producing surface groups for non-CLC countries also offers the flexibility of classifying Landsat pixels only for subregions of a country, with all steps of our classification procedure (i.e., draw of training sample, training of algorithm, and classification of pixels in the Landsat greenest pixel composite) taking place for each subregion separately. Such a separation of subregions can be useful for large countries with differences in vegetation and climate across subregions. For example, splitting the U.S. by states could improve the classification output because the states differ substantially in terms of vegetation and climate. Moreover, the average area size of a U.S. state roughly equals that of a CLC country, so that through splitting the U.S. into states the proportion of training data size and size of the greenest pixel composite would stay constant, thus potentially improving the classification output further. The same reasoning applies to other large countries such as Australia, Canada, or China.

3.5.2 External Validity Analyses

3.5.2.1 Overview

In this appendix section, we investigate the surface groups measure's external validity as a proxy for economic activity. The purpose of the measure is to approximate economic activity over a long time series and at small regional levels. To examine whether the surface groups fulfill this purpose, we require external data on economic activity at small regional levels. With such external data, we can empirically analyze the quality of a surface groups-based prediction of economic activity.

In our main validation analyses, we draw on two external sources of validation data to analyze the surface groups-based prediction of economic activity. First, from administrative statistics, we extract a regionally disaggregated direct measure of GDP, the most commonly used indicator of economic activity in the literature evaluating previous satellite-based proxies for economic activity (e.g., Chen and Nordhaus, 2011; Henderson et al., 2012). The administrative GDP measure is available at the county (*Kreis*) level[17] from 2000.

[17] As of January 1, 2017, Germany comprised 401 counties.

Second, we use the socioeconomic dataset RWI-GEO-GRID (RWI and microm, 2019) that provides household income as a further indicator of economic activity with a very high level of regional detail. This indicator is available at the level of grid cells sized one square kilometer (and thus independent of administrative borders), but annually only from 2009.

To evaluate the surface groups-based prediction of economic activity, we perform OLS regressions of the two indicators of economic activity (GDP and household income) on the surface groups. These regressions allow us to determine how much of the variation in economic activity the surface groups explain. Furthermore, we analyze the distribution of the regression residuals to assess potential biases in the prediction of economic activity. Throughout this evaluation, we compare the surface groups-based prediction of economic activity to the prediction based on DMSP OLS night light intensity data. This commonly used night lights-based prediction thus serves as a benchmark for assessing the quality of our daytime-based prediction using surface groups.

In additional validation analyses, we examine further predictive properties of surface groups. We (a) investigate within-region predictive power, (b) evaluate surface groups against the newer VIIRS night light intensity data with higher spatial resolution, and (c) compare the predictive value of surface groups to Yeh et al.'s (2020) prior approach in Africa.

Furthermore, we investigate the quality of surface groups as proxies for their corresponding types of land cover as indicated in administrative statistics. This analysis complements the assessment of internal validity in Appendix 3.5.1.5 and evaluates whether the surface groups are a useful measure for studies examining different types of land cover (e.g., urbanization, deforestation). The administrative measures of land cover are available at the municipality (*Gemeinde*) level[18] consecutively from 2008 through 2015. We perform separate OLS regressions for each surface group on the corresponding type of administrative land cover and examine the regression residuals.

By using external validation data that are available for limited time series, the analyses in this appendix section provide insight into the quality of the surface groups as a measure

[18] As of January 1, 2017, Germany comprised 11,266 municipalities.

for applications in economic research. After describing the external data we use for these analyses in more detail, we present the analysis of surface groups as a novel six-dimensional proxy for economic activity. Finally, we show how the six surface groups can be combined into a single-variable proxy.

3.5.2.2 Validation Data

To obtain economic indicators at detailed regional levels, we draw on two data sources for our main external validation analyses in Appendix 3.5.2.3. First, we use administrative regional data. We access these data via the "Regionaldatenbank Deutschland",[19] a database belonging to the German Federal Statistical Office's (GFSO) data portal, GENESIS.[20] This database comprises a variety of regional statistics from the GFSO and the statistical offices of the 16 federal states (*Bundesländer*), with varying time series and levels of regional disaggregation. GDP information in the administrative statistics is available at the county level, the next lower administrative regional unit after the federal states, from 2000 through 2018.[21] Following Henderson et al. (2012), we use real (i.e., deflated)[22] GDP measures in euros as a validation measure for our analyses. We denote real GDP as GDP.

Second, we use RWI-GEO-GRID (RWI and microm, 2019), a grid-level dataset containing socioeconomic indicators collected from a variety of public and private sources, but annually available only from 2009 through 2016 (for a more detailed description of this dataset, see Breidenbach and Eilers, 2018). From this dataset, we extract a measure of household income that allows us to analyze economic activity at a regional level even more detailed than the administrative county level. This measure is available at the level of grid cells sized one square kilometer, an extremely high level of regional detail, and indicates the total purchasing power of all households living in a grid cell (Breidenbach and

[19] https://www.regionalstatistik.de/genesis/online/ (last retrieved on July 19, 2021).

[20] The acronym "GENESIS" stands for "Gemeinsames Neues Statistisches Informations-System". See https://www.statistikportal.de/de/datenbanken (last retrieved on July 19, 2021).

[21] We use data table 82111-01-05-4 "Bruttoinlandsprodukt/Bruttowertschöpfung nach Wirtschaftsbereichen – Jahressumme – regionale Tiefe: Kreise und krfr. Städte" available at https://www.regionalstatistik.de/genesis/online?operation=previous&levelindex=1&step=1&titel=Tabellenaufbau&levelid=1626691580813&acceptscookies=false#abreadcrumb (last retrieved on June 29, 2021).

[22] We deflate to 2000 prices according to the consumer price index provided by the GFSO. See https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=61111-0001&startjahr=1991#abreadcrumb (last retrieved on November 4, 2021).

Eilers, 2018). The grid cells in this dataset follow the system of the European Reference Grid distributed by the European Soil Data Centre (ESDAC)[23] (Breidenbach and Eilers, 2018). To evaluate the quality of the surface groups-based prediction at this very detailed regional level, we use real household income measured in euros at the grid level as a further indicator of economic activity. For data protection, the dataset contains missing or zero values for grid cells with a population below five inhabitants or households (Breidenbach and Eilers, 2018). However, we expect economic activity and thus household income in these grid cells to be negligibly small, so that our analysis excludes grid cells essentially without economic activity. Altogether, Germany comprises 381,425 grid cells, between 146,382 and 148,509 of which (depending on the year) contain positive values of household income within our observation period. We denote real household income as HHI.

To compare the quality of the prediction that uses surface groups to the prediction that uses night light intensity in our main validation analyses (appendix 3.5.2.3), we use DMSP OLS night lights data, available from 1992 through 2013.[24] Simply put, these data capture the intensity of light sources on earth at night (Huang et al., 2014). This night light intensity constitutes a valuable proxy for economic activity at the national level and at larger subnational levels such as federal states or metropolitan areas (Chen and Nordhaus, 2011; Elvidge et al., 1997; Pinkovskiy and Sala-i-Martin, 2016). The technological developments of the 21st century have improved both the accessibility of night lights data and the computational capabilities for processing these data (Donaldson and Storeygard, 2016). Consequently, night lights data have become an attractive data source for economists in the last decade. Similar to Henderson et al. (2012), we use the pre-processed version of the DMSP OLS data (i.e., the version corrected for, e.g., clouds or unusual lighting such as forest fires). This version contains one observation per pixel and year, indicating the intensity of light sources on earth at night.[25] The intensity variable

[23] Available from https://esdac.jrc.ec.europa.eu/content/european-reference-grids (last retrieved on August 13, 2019).

[24] We use the Version 4 DMSP-OLS Nighttime Lights Time Series distributed by the National Oceanic and Atmospheric Administration's (NOAA) National Geophysical Data Center, available at https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html#AVSLCFC (last retrieved on October 25, 2021).

[25] For a few observation years, two satellites collected night light intensity. Consequently, the night lights data contain two observations per pixel for these years. Following Henderson et al. (2012), we use the

is a digital number ranging between 0 and 63. To achieve regional correspondence with the administrative GDP data and RWI-GEO-GRID, we calculate the average DMSP OLS night light intensity at the county and at the grid level (denoted as $NL_{DMSPOLS}$).

Furthermore, in Appendix 3.5.2.4 we use two other data sources. First, we use VIIRS night lights data as an alternative to the DMSP OLS benchmark.[26] While VIIRS data offer a higher spatial resolution than DMSP OLS data (500 meters vs. one kilometer at the equator), their available time series is substantially shorter (2012–2020 vs. 1992–2013). As the 2012 and 2013 VIIRS composites differ from later years by not being built from stray-light corrected data (Elvidge et al., 2021), we do not use these two years in our analyses to have a consistent benchmark. Like DMSP OLS data, VIIRS data contain one observation per pixel and year. We denote the regional average of the VIIRS night light intensity variable, which indicates radiance measured in nano Watts per square centimeter per steradian, as NL_{VIIRS}.

Second, we use an index for village-level asset wealth from Yeh et al.'s (2020) work in Africa. They use African Demographic and Health Survey (DHS) data to construct this index, including measures for quality of living (e.g., if households have running water). They then train a neural network to directly predict the index from a combination of Landsat and DMSP OLS night light intensity data. We use both their original DHS-based asset wealth index as an outcome to validate the surface groups against (denoted as AWI) and their predicted asset wealth index as benchmark.[27]

For evaluating the surface groups as proxies for their corresponding types of land cover, we derive such corresponding measures of land cover from administrative land use statistics available from GENESIS.[28] These statistics report how much of a region's area serves a specific land-use purpose. The categories of land use are aggregated versions of the

average of those observations.

[26] We use the annual VIIRS night lights composites version 2 (Elvidge et al., 2021), available from the Colorado School of Mines at https://eogdata.mines.edu/nighttime_light/annual/v20/ (last retrieved on October 27, 2021).

[27] Both the original and the predicted asset wealth index are available as a supplement to Yeh et al. (2020).

[28] We use data table 33111-01-01-5 "Bodenfläche nach Art der tatsächlichen Nutzung – Stichtag 31.12. – regionale Tiefe: Gemeinden (bis 2015)" available at https://www.regionalstatistik.de/genesis//online?operation=table&code=33111-01-01-5&bypass=true&levelindex=0&levelid=1653945321315#abreadcrumb (last retrieved on February 20, 2018).

categories in the official real estate register that indicates land use within administrative regions (see Working Committee of the Surveying Authorities of the Laender of the Federal Republic of Germany (AdV), 1991). As with the CLC classes (see appendix 3.5.1.3), the administrative categories indicate either land use alone or a mixture of land use and land cover. Unfortunately, at their level of aggregation, the administrative data do not allow us to identify all subcategories representing a given surface group in the real estate register. Consequently, the administrative measures we obtain constitute lower bounds of their actual values.[29] Therefore, to obtain corresponding measures of land cover, we sum up those categories that unambiguously indicate land use belonging to the six surface groups. We thus retrieve six administrative measures of land cover, denoted as $builtup_{adm}$, $grass_{adm}$, $crops_{adm}$, $forest_{adm}$, $noveg_{adm}$, and $water_{adm}$ (measured in square kilometers).

The administrative data indicate land use at the municipality level, the smallest administrative regional unit in Germany, consecutively from 2008 through 2015.[30] As Germany experienced a number of territorial reforms over the last two decades and the territorial status underlying the administrative land use information varies between years, we need to update this information to a common territorial status.[31] Because the territorial reforms mainly involved municipal mergers,[32] updating land use primarily entails adding the information from formerly separated municipalities that merged, thereby preventing a loss of measurement precision.[33] For municipalities belonging to the federal state of Lower Saxony, inconsistencies in the region identifier prevent us from reliably assigning the administrative data to the correct geographic location. Therefore, we drop these municipalities from our analyses.[34] Apart from this exclusion, we are able to

[29] For example, the higher level of aggregation in our data does not allow us to distinguish public and historical buildings from other types of non-built-up land cover.

[30] Since 2016, the land use categories follow a definition different from that in AdV (1991). The new definition no longer allows us to consistently identify the six types of land cover corresponding to the surface groups.

[31] We chose the territorial status of January 1, 2017, because it was the most current one when we started work on this chapter. The territorial reforms did not affect the county level, at which the administrative data provides GDP.

[32] The number of municipalities diminished from 12,485 on January 1, 2008 to 11,266 on January 1, 2017.

[33] In the rare cases of municipal separations, we divide the original land use information relative to the areas of the new municipalities.

[34] Apart from a reduction in sample size, we do not expect that dropping municipalities in the federal state of Lower Saxony affects our results, because the remaining parts of Germany still exhibit much variation in land cover.

calculate $builtup_{adm}$, $grass_{adm}$, $crops_{adm}$, $forest_{adm}$, $noveg_{adm}$, and $water_{adm}$ annually for all municipalities.

To assess the value of the surface groups we derive from Landsat data as a proxy for economic activity, we aggregate the pixel-level surface groups information to the different regional units of the validation data. We do so by counting the number of pixels in each surface group per regional unit and year, thus generating, at the respective regional level, six variables indicating the number of pixels per surface group: $builtup_{sat}$, $grass_{sat}$, $crops_{sat}$, $forest_{sat}$, $noveg_{sat}$, and $water_{sat}$.[35] Moreover, to improve the evaluation by accounting for potential measurement error in the number of pixels per surface group, we calculate a region's percentage of pixels with values missing because of, for example, cloud cover as an indicator of potential measurement error.[36]

In sum, this set of validation data allows us to perform a precise validation analysis of surface groups as a novel six-dimensional proxy for economic activity. We argue that if the quality of the surface groups-based prediction is high in the years that the validation data cover, this quality is high for earlier periods as well because we consistently measure the surface groups over time (i.e., for the entire period from 1984–2020). Put differently, we have no reason to believe that our results on the validity of surface groups as a proxy for economic activity would change if the validation data were already available from 1984. Therefore, we assume that the conclusions we draw from the validation analysis also hold for earlier periods for which validation data are not available (1984–1999 for GDP and 1984–2008 for household income) and, consequently, that the surface groups proxy economic activity equally well from 1984 through 2020.

[35] For better efficiency, we perform the aggregation tasks of surface groups (and that of any other regionally aggregated variables in our analyses such as night light intensity) to the different regional units using Esri's ArcPy package. However, these tasks can be achieved using freeware such as PyQGIS with similar results. The polygon shapefiles indicating the regional borders of the validation data in our analyses are available from the Federal Agency for Cartography and Geodesy (BKG) at https://daten.gdz.bkg.bund.de/produkte/vg/vg250_ebenen_0101/ (last retrieved on November 3, 2021; administrative regional borders in Germany, including historical shapefiles we use for determining municipalities' territorial status as of January 1, 2017, following Egger et al., 2022), from the ESDAC at https://esdac.jrc.ec.europa.eu/content/european-reference-grids (last retrieved on August 13, 2019; grid-cell borders for EU25 countries), and from the Database of Global Administrative Areas (GADM) at https://gadm.org/download_country.html (last retrieved on November 22, 2021; administrative borders of African and other countries).

[36] Other reasons for missing values could be implausible spectral values or inexistence of imagery (see appendix 3.5.1.2). However, cloud cover is the most likely reason.

3.5.2.3 Validation of Surface Groups as a Proxy for Economic Activity

Methodology. To assess the external validity of surface groups as a proxy for economic activity and to compare them to night light intensity—which has become a widely accepted proxy in economic research—we perform OLS regressions of the following form:

$$Y_{i,t} = \beta_0 + \beta_1 X_{i,t} + \beta_2 C_{i,t} + \nu_{i,t} \tag{3.4}$$

with i denoting the regional unit of observation (i.e., counties for the GDP analysis and grid cells for the household income analysis), t denoting the year of observation, and Y denoting the dependent variable $ln(GDP)$ or $ln(HHI)$. X denotes the independent variables, that is, the vector of surface groups (including $ln(builtup_{sat}+1)$, $ln(grass_{sat}+1)$, $ln(crops_{sat}+1)$, $ln(forest_{sat}+1)$, $ln(noveg_{sat}+1)$, and $ln(water_{sat}+1)$) or $ln(NL_{DMSPOLS}+1)$. C represents a vector of control variables and ν constitutes the error term.

To compare the surface groups-based prediction to the night lights-based prediction, we restrict the observation periods to those years for which all variables entering the equation are available. The years of observation are thus 2000 through 2013 for the GDP analysis and 2009 through 2013 for the household income analysis.[37]

To assess whether the combination of surface groups is a valid proxy for economic activity, we follow Henderson et al. (2012) by using the natural logarithms of the dependent variables and the independent variables. We add the value one to the variables in X before taking their natural logarithms, because they contain values of zero. As the variables in X do not represent percentage points, they do not add up to 100. Thus including all six surface groups in the regressions does not lead to multicollinearity. In an assessment of night light intensity as a country-level proxy for GDP, Henderson et al. (2012) argue that night light intensity might be more sensitive to a growth in GDP than to a decline in it, because technology and other factors constantly change over time. The same logic applies to surface groups. For example, while a growth in GDP and the construction of

[37] The household income data are also available for 2005, but we exclude this year to consistently examine patterns in the temporal distribution of the regression residuals by maintaining a data structure of consecutive years.

new buildings might occur simultaneously, a decline in GDP might involve a stagnation of construction activities or an abandonment of buildings rather than a remotely sensible reduction in built-up surfaces. Therefore, surface groups might also be more sensitive to a growth in GDP than to a decline in it.

The vector C comprises two control variables that cancel out any bias due to potential measurement error in the dependent or independent variables. First, year FE account for potential quality differences between years in Landsat or DMSP OLS data. Such differences might occur due to, for example, the technological performance of satellites or weather conditions. Second, federal state FE control for potential differences in administrative data collected by the statistical offices of the federal states.[38]

County-Level Analysis of GDP. The results of the county-level analysis with real GDP as the dependent variable in Table 3.10 show that surface groups explain more of the variation in GDP than DMSP OLS night light intensity. In the specifications without control variables, surface groups explain 43.9 percent of the variation in GDP (column 1), whereas night light intensity explains only 23.0 percent of this variation (column 3). Including the control variables does not affect this pattern, with surface groups explaining 62.3 percent (column 2) and night light intensity explaining 47.1 percent of the variation in GDP (column 4). As the specifications with control variables explain a larger percentage of the variation in GDP for both surface groups and night light intensity, controlling for potential measurement error improves the prediction but neither affects the predictive properties of surface groups nor those of DMSP OLS night light intensity. At the disaggregated regional level of counties, the combination of surface groups and control variables thus explains a significant percentage of the variation in GDP.

Figures 3.2a and 3.2b show that the statistical distribution of the residuals from the OLS regressions with control variables (columns 2 and 4 of table 3.10) looks smoother and narrower for surface groups than for DMSP OLS night light intensity. This finding is in

[38] As we compare the surface groups-based prediction to the night lights-based prediction, we do not include the percentage of cloud cover (see appendix 3.5.2.2) as a control variable for potential measurement error in the number of pixels per surface group. The results do not change when we include this control variable in the prediction using surface groups (see tables 3.11 and 3.13).

Table 3.10: OLS prediction of GDP using surface groups and using DMSP OLS night light intensity (county level, 2000–2013)

	Surface groups		DMSP OLS night light intensity	
Dep. var.: $ln(GDP)$	(1)	(2)	(3)	(4)
$ln(builtup_{sat}+1)$	1.625***	1.368***		
	(0.029)	(0.035)		
$ln(grass_{sat}+1)$	-0.050***	-0.132***		
	(0.015)	(0.013)		
$ln(crops_{sat}+1)$	-0.354***	-0.269***		
	(0.012)	(0.012)		
$ln(forest_{sat}+1)$	-0.095***	-0.162***		
	(0.011)	(0.011)		
$ln(noveg_{sat}+1)$	-0.408***	-0.246***		
	(0.016)	(0.015)		
$ln(water_{sat}+1)$	-0.153***	0.002		
	(0.017)	(0.015)		
$ln(NL_{DMSPOLS}+1)$			0.532***	0.432***
			(0.015)	(0.017)
Year FE	No	Yes***	No	Yes***
Federal state FE	No	Yes***	No	Yes***
N	5,402	5,402	5,402	5,402
Adj. R^2	0.439	0.623	0.230	0.471

Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, GFSO data (GENESIS), and BKG data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3.11: OLS prediction of GDP using surface groups (county level, 2000–2013)

Dep. var.: $ln(GDP)$	(1)	(2)
$ln(builtup_{sat}+1)$	1.642***	1.360***
	(0.029)	(0.035)
$ln(grass_{sat}+1)$	-0.030**	-0.116***
	(0.015)	(0.014)
$ln(crops_{sat}+1)$	-0.357***	-0.282***
	(0.012)	(0.012)
$ln(forest_{sat}+1)$	-0.104***	-0.172***
	(0.012)	(0.012)
$ln(noveg_{sat}+1)$	-0.407***	-0.241***
	(0.016)	(0.015)
$ln(water_{sat}+1)$	-0.151***	0.002
	(0.017)	(0.015)
Year FE	No	Yes***
Federal state FE	No	Yes***
Cloud cover (%)	-2.327**	-4.247***
	(0.960)	(0.923)
N	5,402	5,402
Adj. R^2	0.439	0.624

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data. Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

line with surface groups explaining more of the variation in GDP than night light intensity, as indicated by the adjusted R^2 of the regressions. Moreover, for both surface groups and night light intensity, the residuals are normally distributed, although the distribution has more pronounced local maxima in the night lights specification. Surface groups thus proxy GDP more precisely than DMSP OLS night light intensity.

Furthermore, using surface groups to compare GDP over time and between regions requires that the prediction error be neither temporally nor spatially biased. Temporal bias would occur if the prediction error is constant for a given region throughout all observation years, and spatial bias would occur if the prediction error is equal for clusters of regions. To assess the existence of such biases, Figure 3.6 illustrates the temporal and spatial distribution of the residuals from the regressions in column 2 of Table 3.10. For reference, Figure 3.5 provides a map indicating the names of the federal states and the locations of their capitals. In four-year intervals evenly spread over our observation period, Figure 3.6 shows the estimated residuals for all counties in the respective year, that is, the degree to which GDP is overestimated (blue counties) or underestimated (red counties). For comparison, Figure 3.7 proceeds similarly for DMSP OLS night light intensity, illustrating the residuals from the regression in column 4 of Table 3.10.

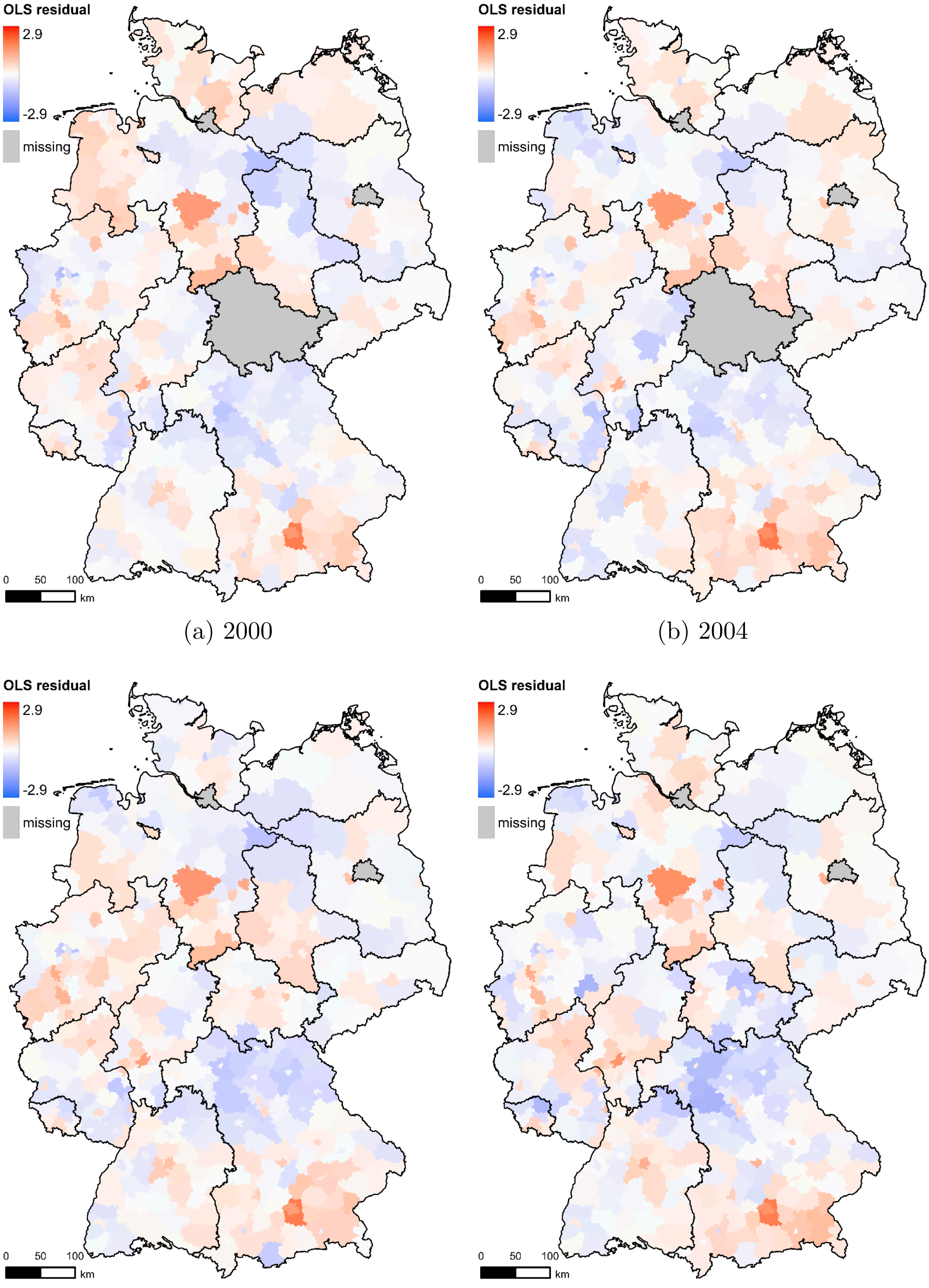
Figures 3.6 and 3.7 suggest that the surface groups-based prediction yields a considerably smaller temporal bias than the night lights-based prediction. If a temporal bias in prediction error existed, the color of a given region would stay the same over the entire observation period. For surface groups, such a pattern exists for 179 counties (44.9 percent), and, for the remaining regions, the color varies over time in Figure 3.6. For DMSP OLS night light intensity, this pattern appears for 339 counties (85.0 percent), leading to the four maps in Figure 3.7 hardly differing in color. Therefore, although we cannot definitely rule out the existence of a temporal bias for some regions when proxying GDP with surface groups, this temporal bias is far less severe than that of proxying GDP with night light intensity.

The distribution of the residuals across regions in Figures 3.6 and 3.7 suggests a somewhat larger spatial bias in prediction error for surface groups than for DMSP OLS night light intensity. If such a bias existed, clusters of similarly colored regions would

Figure 3.5: Reference map of German federal states and their capitals

Source: Authors' illustration with BKG data.

Figure 3.6: Spatial and temporal distribution of GDP residuals for surface groups



(a) 2000

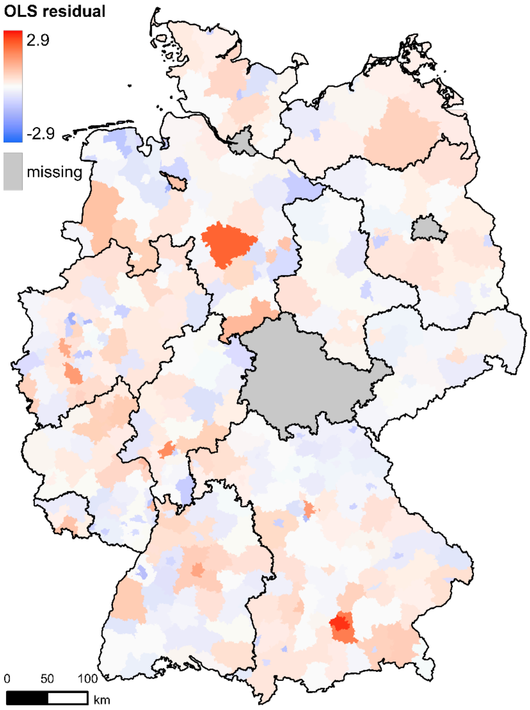
(b) 2004

(c) 2008

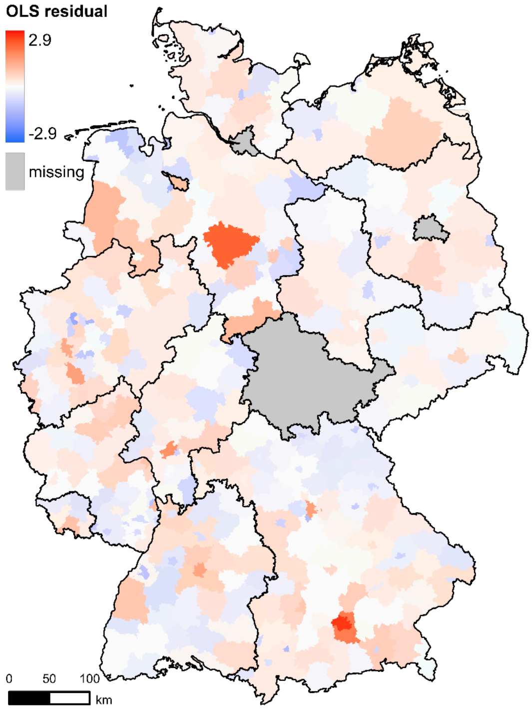
(d) 2012

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Note: Maps illustrate residuals from the regression in column 2 of Table 3.10.

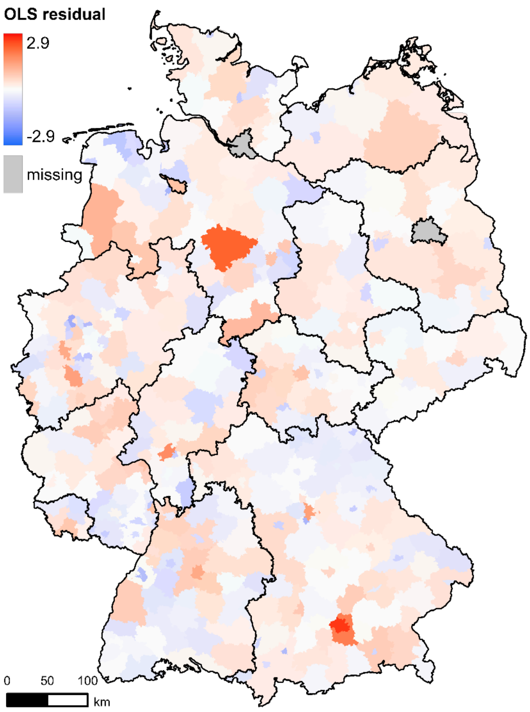
Figure 3.7: Spatial and temporal distribution of GDP residuals for DMSP OLS night light intensity



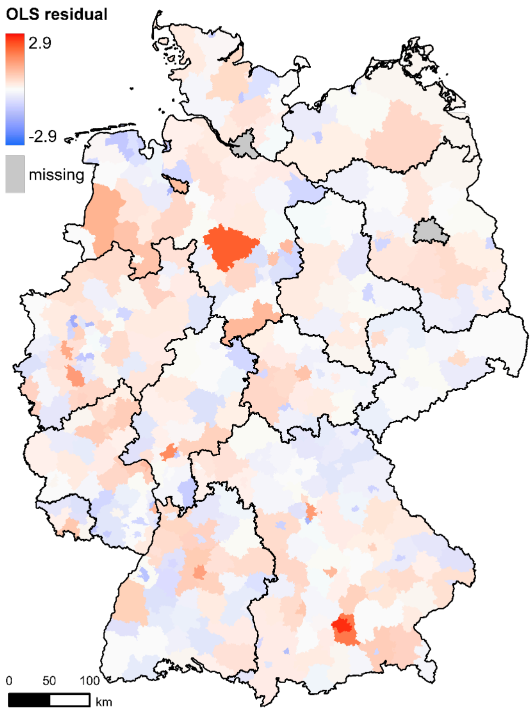
(a) 2000



(b) 2004



(c) 2008



(d) 2012

Source: Authors' calculations based on DMSP OLS data, GFSO data (GENESIS), and BKG data.
Note: Maps illustrate residuals from the regression in column 4 of Table 3.10.

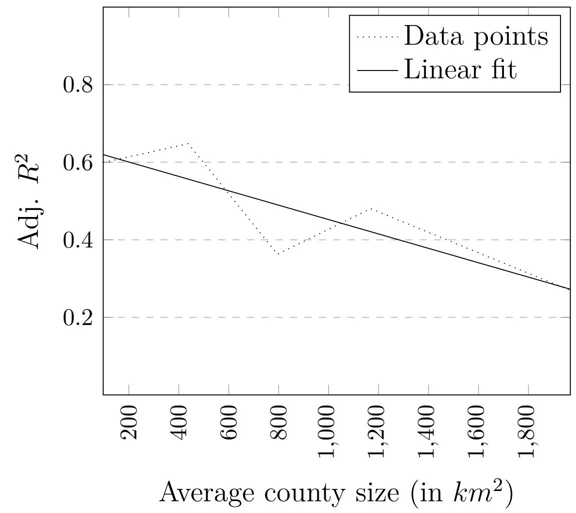
appear. For surface groups, 992 observations (18.4 percent) have the same color as all their geographically neighboring observations, whereas for night light intensity, this pattern shows for only 565 observations (10.5 percent). However, for both surface groups and night light intensity, the clusters appear randomly distributed across the country rather than concentrated in specific parts (e.g., clusters not only in rural areas, clusters not only in the north). Therefore, the spatial distribution of the prediction error appears random but yields a larger bias for surface groups.

Combining the indicators of temporal and spatial bias shows that the smaller temporal bias of the surface groups-based prediction outweighs the prediction's larger spatial bias as compared to the night lights-based prediction. For surface groups, only 11 counties (2.8 percent) have the same color as all their neighboring observations and, simultaneously, the same color throughout all observation years. For DMSP OLS night light intensity, this pattern appears for 26 counties (6.5 percent). This finding reflects in the small clusters of similarly colored counties not showing up in consecutive years in Figure 3.6.

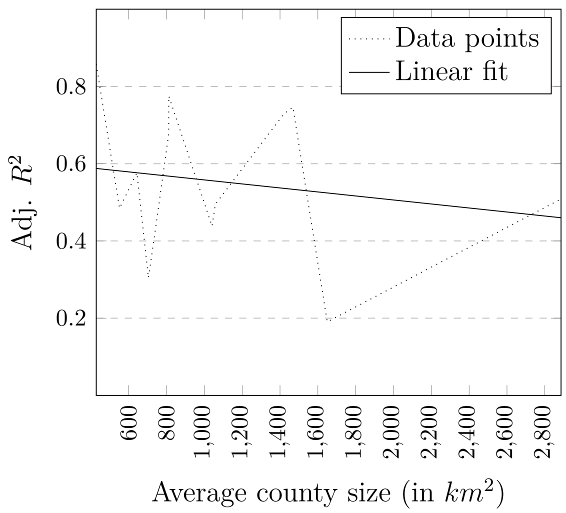
In addition, to show that the value of surface groups as a proxy for economic activity increases with the degree of regional disaggregation, we estimate our OLS model separately by county-size groups. Figure 3.8 plots average county size within a group against the adjusted R^2 obtained from the separate regressions. As county-size groups, we use quintiles of the county-size distribution (figure 3.8a) and federal states (figure 3.8b). In addition to the original data points obtained from the regressions, Figure 3.8 also plots the linear fitted values to visualize the trend in the data. For both county-size groups, the plots show a declining trend, that is, the percentage of the variation in GDP explained by surface groups declines with an increase in county size. Put differently, the smaller the county size the better the proxy. This finding emphasizes the potential of surface groups as a valuable measure for analyses at detailed regional levels.

In essence, the county-level analysis of the surface groups-based prediction of GDP yields the finding that surface groups are a highly suitable proxy for GDP. They explain a significant percentage of the variation in GDP. Moreover, in comparison to the DMSP OLS night lights-based prediction, the surface groups-based prediction shows a smaller

Figure 3.8: Adjusted R^2 by county-size groups



(a) By area-size quintile



(b) By federal state

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Note: Values from separate regressions of surface groups on GDP corresponding to the specification in column 2 of Table 3.10.

bias in the regression residuals. Therefore, surface groups provide a useful alternative for proxying GDP at disaggregated regional levels such as German counties.

Grid-Level Analysis of Household Income. In the grid-level analysis of surface groups as a proxy for household income, we find the same patterns as in the county-level analysis of surface groups as a proxy for GDP. Table 3.12 presents the estimation results for this grid-level analysis. At this very detailed regional level, the surface groups-based prediction explains a much larger percentage of the variation in household income than the DMSP OLS night lights-based predictions (63.6 percent vs. 27.2 percent in the specifications without control variables in columns 1 and 3, and 67.5 percent vs. 30.7 percent in the specifications with control variables in columns 2 and 4). In comparison to the GDP analysis, the control variables (year FE and federal state FE) improve the prediction only slightly in the household income analysis, probably because the number of observation years is smaller and because the dependent variable is not collected within administrative borders.

Figures 3.2c and 3.2d confirm the findings of the regressions. The statistical distribution of the prediction error for household income is much narrower (although slightly left-skewed) for surface groups than for night light intensity. The distribution of the prediction error for night light intensity is slightly right-skewed and, instead of a peak at the value zero, the distribution exhibits a plateau around this value. Therefore, surface groups proxy household income at the grid level much more precisely than DMSP OLS night light intensity.

Furthermore, the assessment of the temporal and spatial distribution of the prediction error in the household income analysis yields results similar to those in the GDP analysis. Figures 3.9 and 3.10 show the spatial and temporal distribution of the prediction error in household income for surface groups and DMSP OLS night light intensity, respectively. For a better illustration of the very small grid cells, the map shows an area at the borders of four federal states, with the metropolitan region of Ludwigshafen-am-Rhein/Mannheim in the south-west and the rural Odenwald region in the east. The gray cells are those with missing values (i.e., uninhabited or only sparsely inhabited areas).

Again, the smaller temporal bias in the surface groups-based prediction in comparison to the night lights-based prediction outweighs the larger spatial bias. For surface groups 90,054 grid cells (59.5 percent) have the same color throughout all observation years, whereas this number amounts to 131,704 grid cells (87.0 percent) for DMSP OLS night light intensity. Moreover, the spatial bias of the surface groups-based prediction is only slightly larger than the spatial bias of the night lights-based prediction, with 167,095 observations (22.7 percent) for surface groups and 126,703 observations (17.2 percent) for night light intensity having the same color as all their geographical neighbors. Combining the two types of biases shows that for surface groups, 8,166 grid cells (5.4 percent) have the same color as their neighbors and, simultaneously, the same color throughout all observation years. For DMSP OLS night light intensity, this pattern applies to 15,058 grid cells (9.9 percent). Therefore, the smaller temporal bias of surface groups again outweighs their slightly larger spatial bias.

Summary. To summarize our main analyses of the surface groups' external validity, we show that at the county level (GDP) and at the grid level (household income) surface groups can serve as a valid proxy for economic activity. At both levels, the surface groups predict a significant percentage of the variation in economic activity, and this prediction is more precise (i.e., less biased) for surface groups than for DMSP OLS night light intensity. Furthermore, the comparative advantage of surface groups as a proxy for economic activity becomes more pronounced in the grid-level analysis than in the county-level analysis. This finding, in combination with the GDP analysis by county-size group, suggests that surface groups are particularly useful for applications that investigate very small regional units. Although we derive these findings from external validation data with limited time series, we argue that, due to the high and temporally stable internal validity of the surface groups measure (see appendix 3.5.1.5), surface groups can also function as a valid proxy for economic activity for earlier years.

To ensure the surface groups' validity across all years in economic or other applications,

Table 3.12: OLS prediction of household income using surface groups and using DMSP OLS night light intensity (grid level, 2009–2013)

	Surface groups		DMSP OLS night light intensity	
Dep. var.: $ln(HHI)$	(1)	(2)	(3)	(4)
$ln(builtup_{sat} + 1)$	1.449***	1.412***		
	(0.002)	(0.002)		
$ln(grass_{sat} + 1)$	-0.090***	-0.126***		
	(0.002)	(0.002)		
$ln(crops_{sat} + 1)$	-0.422***	-0.371***		
	(0.002)	(0.002)		
$ln(forest_{sat} + 1)$	-0.053***	-0.066***		
	(0.001)	(0.001)		
$ln(noveg_{sat} + 1)$	-0.200***	-0.173***		
	(0.001)	(0.001)		
$ln(water_{sat} + 1)$	-0.268***	-0.211***		
	(0.001)	(0.001)		
$ln(NL_{DMSPOLS} + 1)$			0.936***	0.953***
			(0.002)	(0.002)
Year FE	No	Yes***	No	Yes***
Federal state FE	No	Yes***	No	Yes***
N	737,626	737,626	737,626	737,626
Adj. R^2	0.636	0.675	0.272	0.307

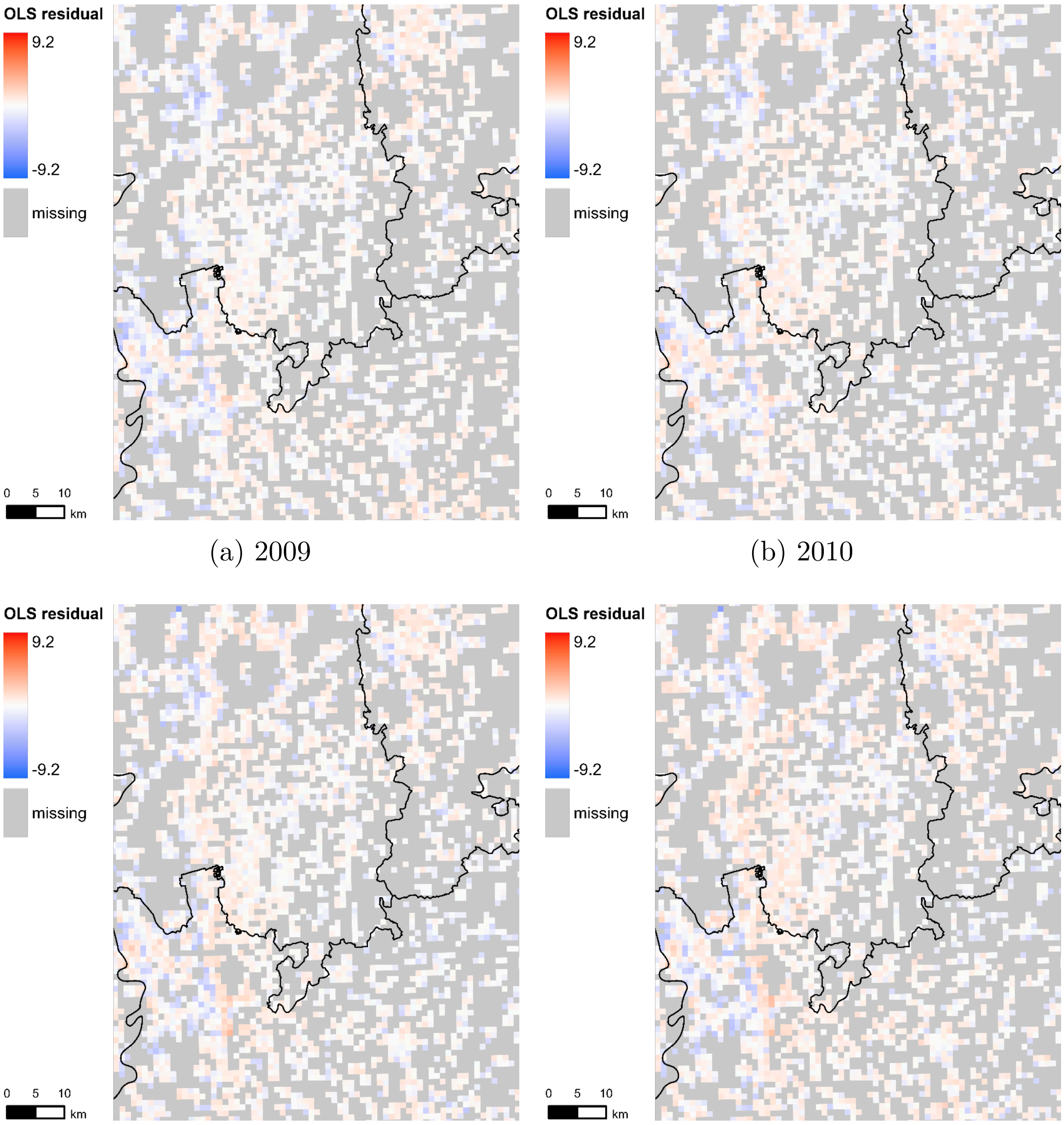
Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, RWI and microm (2019) data, and ESDAC data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3.13: OLS prediction of household income using surface groups (grid level, 2009–2013)

Dep. var.: $ln(HHI)$	(1)	(2)
$ln(builtup_{sat} + 1)$	1.462***	1.426***
	(0.002)	(0.002)
$ln(grass_{sat} + 1)$	-0.0832***	-0.118***
	(0.002)	(0.002)
$ln(crops_{sat} + 1)$	-0.413***	-0.360***
	(0.001)	(0.001)
$ln(forest_{sat} + 1)$	-0.044***	-0.057***
	(0.001)	(0.001)
$ln(noveg_{sat} + 1)$	-0.200***	-0.173***
	(0.001)	(0.001)
$ln(water_{sat} + 1)$	-0.270***	-0.214***
	(0.001)	(0.001)
Year FE	No	Yes***
Federal state FE	No	Yes***
Cloud cover (%)	0.804***	0.904***
	(0.052)	(0.053)
N	737,626	737,626
Adj. R^2	0.637	0.675

Source: Authors' calculations based on Landsat data, CLC data, RWI and microm (2019), and BKG data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 3.9: Spatial and temporal distribution of household income residuals for surface groups



(a) 2009

(b) 2010

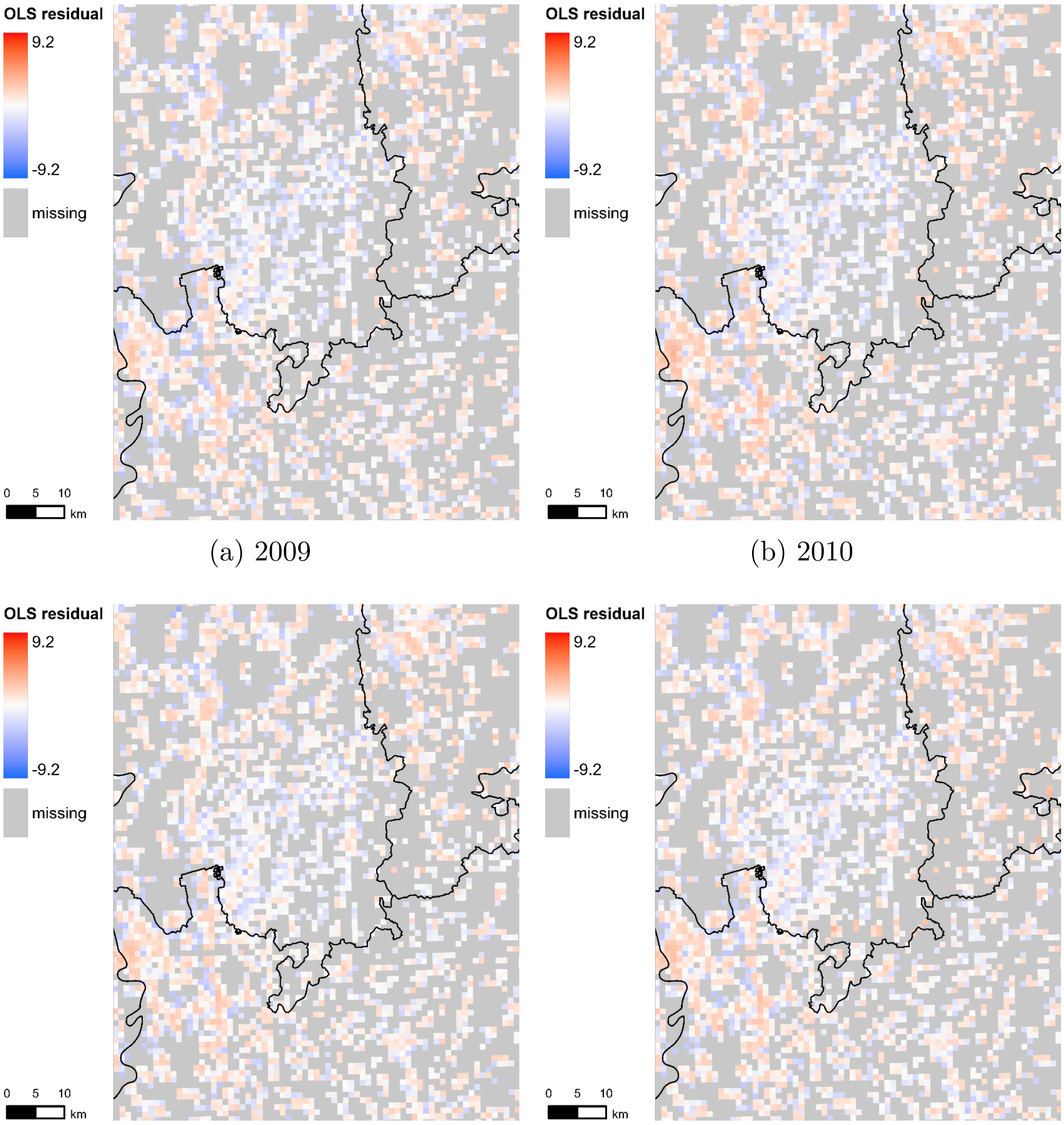
(c) 2011

(d) 2012

Source: Authors' calculations based on Landsat data, CLC data, RWI and microm (2019) data, and ESDAC data.

Note: Maps illustrate residuals from the regression in column 2 of Table 3.12. Maps show an area at the borders of the four federal states Rhineland-Palatinate, Hesse, Baden-Württemberg, and Bavaria.

Figure 3.10: Spatial and temporal distribution of household income residuals for DMSP OLS night light intensity



(a) 2009 (b) 2010 (c) 2011 (d) 2012

Source: Authors' calculations based on DMSP OLS data, RWI and microm (2019) data, and ESDAC data.

Note: Maps illustrate residuals from the regression in column 4 of Table 3.12. Maps show an area at the borders of the four federal states Rhineland-Palatinate, Hesse, Baden-Württemberg, and Bavaria.

we recommend (a) including the number of cloud-covered pixels as a control variable and (b) checking the data for outlier observations and remove those from empirical analyses for particular years and regions. Such outliers can occur in few regions in years with scarce Landsat imagery (particularly in the 1980s). For these years, our greenest pixel composite features higher percentages of cloud-covered pixels, pixels showing cloud shadow, or otherwise invalid pixels. As the filters we apply in constructing the greenest pixel composite cannot detect some of these pixels, our algorithm potentially produces an erroneous classification for these pixels.[39] To obtain more valid results, we apply outlier corrections in the comparison of GDP developments across German counties (figure 3.3). For details on the outlier removal procedure, see Appendix 3.5.2.5.

While surface groups offer substantial advantages in proxying economic activity at disaggregated levels, night light intensity might still be the more appropriate proxy for cross-country studies or other larger regions. The reason is that land use characteristics might have heterogeneous meanings for a country's economy, depending on the country's historical development (Henderson et al., 2018). However, for small regional units and early time series, surface groups constitute a valuable and more accurate proxy for economic activity.

3.5.2.4 Additional Validation Analyses on Surface Groups as a Proxy for Economic Activity

We present three additional analyses on the surface groups' external validity. First, we use VIIRS night light intensity data as a benchmark to show that surface groups offer higher precision in predicting economic activity than night light intensity data with higher spatial resolution compared to DMSP OLS data. Second, we analyze within-region heterogeneity in predicted GDP to demonstrate that surface groups enable the isolation of subregional changes in economic activity. Third, a comparison to Yeh et al.'s (2020) work in Africa suggests that surface groups can function as a proxy for economic conditions also in developing countries.

[39] Visual inspections of the classification show that most of these undetected invalid pixels are classified as *builtup*.

VIIRS Night Light Intensity as Benchmark. To analyze whether surface groups outperform night light intensity data with higher spatial resolution than DMSP OLS data in proxying economic activity, we reestimate the OLS model specified in Equation 3.4 both at the county level (with GDP as outcome) and at the grid level (with household income as outcome) with VIIRS night light intensity as a benchmark. The observation periods of this analysis start in 2014 (first consistent year in the VIIRS data). They end in 2018 for the county-level analysis (last year in the GDP data) and in 2016 for the grid-level analysis (last year in the household income data).

Table 3.14 presents the county-level analysis that compares surface groups and VIIRS night light intensity as proxies for GDP. Our surface groups proxy achieves 142.2 percent of the VIIRS precision in predicting GDP, thus offering a much higher precision. While VIIRS night light intensity explains only 46.9 percent of the variation in GDP in the specification with control variables (column 4), surface groups explain 66.7 percent of this variation (column 2). Therefore, at the county level our surface groups proxy outperforms even night light intensity data with a higher spatial resolution than DMSP OLS data.

The grid-level analysis of household income in Table 3.15 supports the county-level finding that surface groups outperform VIIRS night light intensity in predicting regional economic activity. With 51.8 percent, VIIRS night light intensity explains a lower percentage of the variation in household income than surface groups with 70.0 percent (columns 2 and 4). While VIIRS night light intensity thus appears to perform better in proxying household income than DMSP OLS night light intensity, our surface groups proxy outperforms both sources of night light intensity data.

Within-Region Predictive Power To analyze the surface groups' predictive power of within-region changes in economic activity, we (a) reestimate our model specified in Equation 3.4 with region unit (i.e., county for the GDP analysis and grid cells for the household income analysis) FE and (b) conduct additional analyses on the usefulness of surface groups for proxying within-county changes in economic activity. The results show that (a) region unit and year FE alone explain almost all of the variation in economic activity

Table 3.14: OLS prediction of GDP using surface groups and using VIIRS night light intensity (county level, 2014–2018)

	Surface groups		VIIRS night light intensity	
Dep. var.: $ln(GDP)$	(1)	(2)	(3)	(4)
$ln(builtup_{sat}+1)$	1.419***	1.249***		
	(0.040)	(0.049)		
$ln(grass_{sat}+1)$	-0.054**	-0.151***		
	(0.026)	(0.024)		
$ln(crops_{sat}+1)$	-0.312***	-0.233***		
	(0.018)	(0.019)		
$ln(forest_{sat}+1)$	-0.165***	-0.205***		
	(0.018)	(0.019)		
$ln(noveg_{sat}+1)$	0.028	0.043		
	(0.033)	(0.030)		
$ln(water_{sat}+1)$	-0.239***	-0.043*		
	(0.024)	(0.022)		
$ln(NL_{VIIRS}+1)$			0.482***	0.382***
			(0.025)	(0.026)
Year FE	No	Yes***	No	Yes
Federal state FE	No	Yes***	No	Yes***
N	1,995	1,995	1,995	1,995
Adj. R^2	0.499	0.667	0.213	0.469

Source: Authors' calculations based on Landsat data, CLC data, VIIRS data, GFSO data (GENESIS), and BKG data.
Notes: Robust standard errors in parentheses. All models include intercept. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.15: OLS prediction of household income using surface groups and using VIIRS night light intensity (grid level, 2014–2016)

	Surface groups		VIIRS night light intensity	
Dep. var.: $ln(HHI)$	(1)	(2)	(3)	(4)
$ln(builtup_{sat}+1)$	1.297***	1.275***		
	(0.002)	(0.002)		
$ln(grass_{sat}+1)$	-0.123***	-0.160***		
	(0.002)	(0.002)		
$ln(crops_{sat}+1)$	-0.356***	-0.324***		
	(0.002)	(0.002)		
$ln(forest_{sat}+1)$	-0.076***	-0.074***		
	(0.002)	(0.002)		
$ln(noveg_{sat}+1)$	-0.061***	0.058***		
	(0.002)	(0.002)		
$ln(water_{sat}+1)$	-0.222***	-0.184***		
	(0.002)	(0.001)		
$ln(NL_{VIIRS}+1)$			1.394***	1.377***
			(0.003)	(0.003)
Year FE	No	Yes***	No	Yes***
Federal state FE	No	Yes***	No	Yes***
N	446,524	446,524	446,524	446,524
Adj. R^2	0.671	0.700	0.497	0.518

Source: Authors' calculations based on Landsat data, CLC data, VIIRS data, RWI and microm (2019) data, and ESDAC data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

with neither surface groups nor night light intensity adding any significant informative value, but (b) surface groups are more useful than night light intensity in disentangling which subregional units contribute to changes in economic activity.

Reestimation of our model specified in Equation 3.4 with region unit FE corresponds to, for example, Henderson et al.'s (2012) cross-country analysis of DMSP OLS night light intensity as a predictor for economic activity. The reason that Henderson et al. (2012) include region-level FE (in this case countries) is to control for differences in night light intensity resulting from cultural or economic differences. Such differences can affect the country-wide use of night lights because of, for example, the relative importance of daytime activities in comparison to nighttime activities or the level of technological advancement for producing electricity. However, for within-country applications analyzing small subnational regions—the type of application that we develop our proxy for—such differences are less likely to create heterogeneity over time.

The FE estimations show that county and year FE explain almost all of the variation in economic activity. That is, neither surface groups nor DMSP OLS night light intensity have enough within-county variation over time to significantly contribute to explaining within-region changes. Table 3.16 shows the results of three different FE models that illustrate this finding for the county-level GDP analysis: The first model includes only county and year FE without any of the two proxies; the second includes the surface groups in addition to county and year FE; and the third includes DMSP OLS night light intensity in addition to county and year FE. The models thus correspond to the OLS regressions in Table 3.10, with the difference of containing county instead of federal state FE. We estimate all three models using two different estimation methods, one including the county FE as covariates to obtain an estimate of the overall variance explained by the models and one considering the county FE by subtracting the county-level mean of the dependent variable to obtain an estimate of the within-county variation explained by the model. Both estimation methods show that the inclusion of any proxy leads only to a negligibly small increase in (adjusted) R^2, with the county and year FE explaining 99.6 percent of the overall variation in GDP. Similarly, the grid-cell and year FE explain 99.7 percent of the overall variation in household income (table 3.17). Due to the shorter time series in the grid-level analysis (2009-2013) compared to the county-level analysis (2000-2013), neither including surface groups nor including night light intensity increases adjusted R^2 at all.

Further analyses illustrate that surface groups contribute to a better understanding of within-county changes in regional economic activity. We conduct these analyses at the level of municipalities, the smallest administrative regional unit in Germany.[40] Although GDP data do not exist at the municipality level, we can use the surface groups to derive a prediction of GDP at this level. We then compare the municipality-level change over time in this GDP prediction to the county-level change in the administrative GDP measure. If the change in GDP is similar at both geographic levels, the municipality-level prediction of GDP does not add any informative value to the county-level measure. However, if the change in municipality-level predicted GDP differs from the change in county-level GDP, the new municipality-level prediction can be informative about within-county heterogeneity in economic development, thus allowing assessments of which municipalities drive county-level economic activity (i.e., how economic activity develops heterogeneously within a county). To investigate which proxy offers more insight into within-county heterogeneity, we also compare the surface groups-based and the DMSP OLS night light intensity-based municipality-level GDP predictions.

We proceed in two steps to analyze municipality-level GDP. First, we predict it. Because both the continuous independent variables and the dependent variable are natural logarithms of their original values in the county-level prediction in Table 3.10, the estimation coefficients are not directly transferable to the municipality level. Therefore, we standardize these county-level variables to have a mean of 0 and a standard deviation of 1, then estimate the OLS model specified in Equation 3.4 using the standardized variables. As the standardization does not affect the variables' distributional properties except for the mean and the standard deviation, the OLS result in Table 3.18 has the same properties (adjusted R^2, F-value, coefficients' t-values) as the original unstandardized result. Assuming that the distributional properties of the variables in the model are identical at the county level

[40] As of January 1, 2017, a county consists on average of 28.1 municipalities, with one municipality belonging to only one county.

Table 3.16: FE prediction of GDP using surface groups and using DMSP OLS night light intensity (county level, 2000–2013)

	County FE as covariates			County FE through within-estimator		
	No proxy	Surface groups	DMSP OLS night light intensity	No proxy	Surface groups	DMSP OLS night light intensity
Dep. var.: $ln(GDP)$	(1)	(2)	(3)	(4)	(5)	(6)
$ln(builtup_{sat}+1)$		0.023***			0.023***	
		(0.006)			(0.007)	
$ln(grass_{sat}+1)$		-0.002			-0.002	
		(0.006)			(0.006)	
$ln(crops_{sat}+1)$		-0.021***			-0.021***	
		(0.005)			(0.005)	
$ln(forest_{sat}+1)$		0.007			0.007	
		(0.005)			(0.007)	
$ln(noveg_{sat}+1)$		-0.012***			-0.012***	
		(0.003)			(0.003)	
$ln(water_{sat}+1)$		0.001			0.001	
		(0.003)			(0.004)	
$ln(NL_{DMSPOLS}+1)$			0.076***			0.076***
			(0.010)			(0.010)
Year FE	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
N	5,402	5,402	5,402	5,402	5,402	5,402
Adj. R^2	0.996	0.996	0.996			
Adj. within-R^2				0.295	0.301	0.307

Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, GFSO data (GENESIS), and BKG data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3.17: FE prediction of household income using surface groups and using DMSP OLS night light intensity (grid level, 2009–2013)

	County FE as covariates			County FE through within-estimator		
	No proxy	Surface groups	DMSP OLS night light intensity	No proxy	Surface groups	DMSP OLS night light intensity
Dep. var.: $ln(HHI)$	(1)	(2)	(3)	(4)	(5)	(6)
$ln(builtup_{sat} + 1)$		0.003***			0.003***	
		(0.001)			(0.001)	
$ln(grass_{sat} + 1)$		0.000			0.000	
		(0.000)			(0.000)	
$ln(crops_{sat} + 1)$		0.004***			0.004***	
		(0.000)			(0.004)	
$ln(forest_{sat} + 1)$		0.002***			0.002***	
		(0.000)			(0.000)	
$ln(noveg_{sat} + 1)$		-0.001***			-0.001***	
		(0.000)			(0.000)	
$ln(water_{sat} + 1)$		-0.001***			-0.001***	
		(0.000)			(0.000)	
$ln(NL_{DMSPOLS} + 1)$			-0.000			-0.000
			(0.001)			(0.001)
Year FE	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
N	737,626	737,626	737,626	737,626	737,626	737,626
Adj. R^2	0.997	0.997	0.997			
Adj. within-R^2				0.044	0.044	0.044

Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, RWI and microm (2019) data, and ESDAC data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

and at the municipality level, we can use the coefficients from the county-level estimation with standardized variables to predict standardized GDP at the municipality level. We produce one prediction of standardized municipality-level GDP using surface groups as predictor and one using DMSP OLS night light intensity.

Second, we construct an indicator for the difference between the municipality-level change in predicted GDP and the county-level change in administrative GDP. To obtain the municipality-level change in predicted GDP, for each municipality and for both surface groups and DMSP OLS night light intensity we calculate the difference between the prediction of standardized GDP in 2013 (the last year in the DMSP OLS night light intensity data) and that in 2000 (the first year in the administrative GDP data). To obtain the county-level change in standardized administrative GDP, we proceed similarly at the county level by calculating the difference in administrative GDP between 2013 and 2000. As final indicators, we then calculate for both surface groups and DMSP OLS night light intensity the difference between the municipality-level change in the prediction of standardized GDP and the county-level change in standardized administrative GDP. These indicators measure at the municipality-level whether and to what extent the municipality-level change in GDP over time deviates from the county-level change in GDP over time.

Figure 3.11 plots the distribution of the two indicators. The figure shows that DMSP OLS night light intensity yields a lower degree of additional information at the municipality level in comparison to the county level, that is, surface groups have higher within-region predictive power for geographies below the county level than DMSP OLS night light intensity. Figure 3.11 reveals this relationship through the stronger concentration towards its mean in the indicator for DMSP OLS night light intensity compared to the larger variation in the indicator for surface groups. Therefore, surface groups offer more additional information at the municipality level. The change in the municipality-level prediction of standardized GDP using surface groups thus yields substantially more information on within-county heterogeneity in GDP change in comparison to DMSP OLS night light intensity.

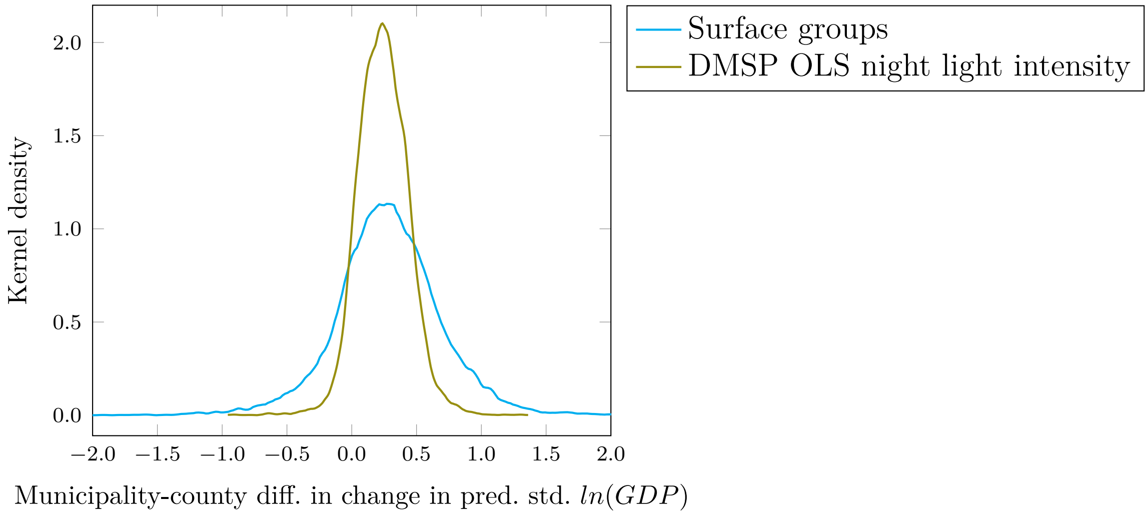
The higher degree of additional municipality-level information obtainable from surface

Table 3.18: OLS prediction of GDP using surface groups and using DMSP OLS night light intensity with standardized variables (county level, 2000–2013)

	Surface groups		DMSP OLS night light intensity	
Dep. var.: standardized $ln(GDP)$	(1)	(2)	(3)	(4)
standardized $ln(builtup_{sat}+1)$	1.975***	1.642***		
	(0.035)	(0.041)		
standardized $ln(grass_{sat}+1)$	-0.109***	-0.285***		
	(0.032)	(0.028)		
standardized $ln(crops_{sat}+1)$	-0.771***	-0.585***		
	(0.026)	(0.025)		
standardized $ln(forest_{sat}+1)$	-0.224***	-0.381***		
	(0.025)	(0.027)		
standardized $ln(noveg_{sat}+1)$	-0.782***	-0.471***		
	(0.032)	(0.029)		
standardized $ln(water_{sat}+1)$	-0.296***	0.003		
	(0.033)	(0.030)		
standardized $ln(NL_{DMSPOLS}+1)$			0.486***	0.395***
			(0.014)	(0.015)
Year FE	No	Yes***	No	Yes***
Federal state FE	No	Yes***	No	Yes***
N	5,402	5,402	5,402	5,402
Adj. R^2	0.439	0.623	0.230	0.471

Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, GFSP data (GENESIS), and BKG data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 3.11: Distribution of municipality-county difference in the change in predicted standardized $ln(GDP)$ between 2000 and 2013 for surface groups-based and DMSP OLS night light intensity-based prediction



Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, GFSO data (GENESIS), and BKG data.
Note: Figure shows univariate kernel density estimates at 300 points using the Epanechnikov kernel function with a kernel half-width of 0.025.

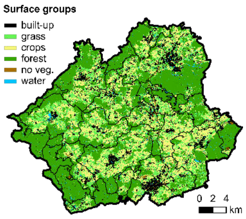
groups also becomes obvious in Figure 3.12, which illustrates for one county (Wunsiedel) as an example the two indicators plotted in Figure 3.11. In essence, surface groups detect much more variation in economic activity in this county's municipalities, represented by the higher intensity of colors in Figures 3.12k and 3.12l.

Figures 3.12a and 3.12b show the surface groups classification underlying the GDP prediction for 2000 and 2013, and Figures 3.12c and 3.12d the corresponding raw DMSP OLS night light intensity. Figures 3.12e and 3.12f illustrate the surface groups-based prediction of standardized municipality-level GDP for these two years, and Figures 3.12g and 3.12h the DMSP OLS night light intensity-based prediction. Figures 3.12i and 3.12j indicate the difference between Figures 3.12e and 3.12f and that between Figures 3.12g and 3.12h, respectively, that is, the changes in the GDP predictions between 2000 and 2013. Figure 3.12k then shows the municipality-county difference in the change in predicted standardized GDP using surface groups as predictor and Figures 3.12l shows this difference using DMSP OLS night light intensity (i.e., the same indicators for which figure 3.11 plots the distribution).

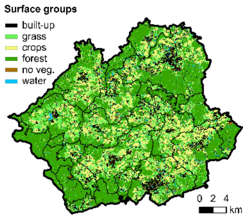
Two properties become noticeable. First, the colors in Figures 3.12e, 3.12f, 3.12i, and 3.12k (surface groups) are much more intense than in Figures 3.12g, 3.12h, 3.12j, and 3.12l (night light intensity). This higher intensity is in line with Figure 3.11, confirming that surface groups offer substantially more information on within-county heterogeneity by detecting variation in economic activity at the municipality level. Second, the municipalities at the south-western border of the county exhibit a substantially lower growth in GDP than the county when using surface groups for prediction (blue-colored municipalities in figure 3.12k), a pattern that is not visible when using DMSP OLS night light intensity (figure 3.12l). These municipalities differ from the other municipalities by being unincorporated areas, that is, typically uninhabited areas (e.g., forests) belonging to the county but without their own municipal governments. Therefore, that these uninhabited municipalities exhibit a substantially lower growth in GDP is a logical consequence of their characteristics. The surface groups detect these characteristics, whereas DMSP OLS night light intensity does not.

Figure 3.12: Surface groups, DMSP OLS night light intensity, predictions of standardized $ln(GDP)$, changes in predicted standardized $ln(GDP)$, and municipality-county differences in the changes in predicted standardized $ln(GDP)$.

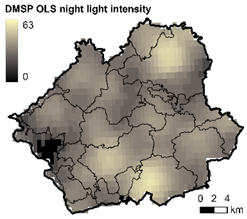
Surface groups **DMSP OLS night light intensity**



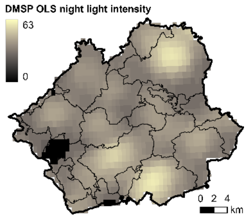
(a) Surface groups (2000)



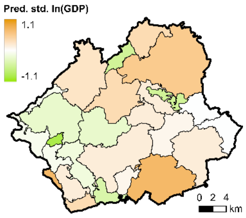
(b) Surface groups (2013)



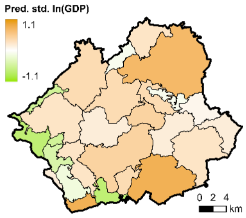
(c) DMSP OLS night light intensity (2000)



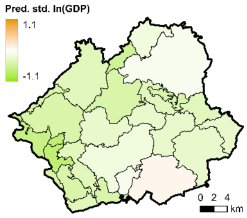
(d) DMSP OLS night light intensity (2013)



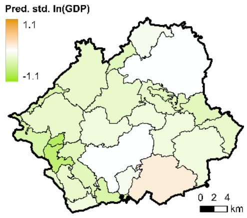
(e) GDP prediction using surface groups (2000)



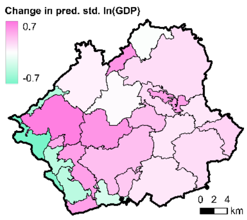
(f) GDP prediction using surface groups (2013)



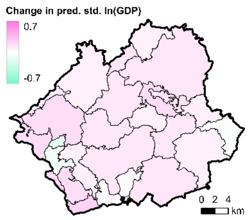
(g) GDP prediction using DMSP OLS night light intensity (2000)



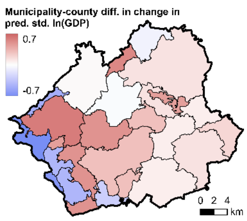
(h) GDP prediction using DMSP OLS night light intensity (2013)



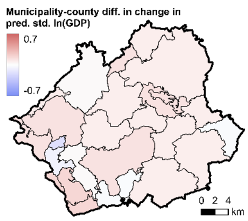
(i) Change in GDP prediction using surface groups between 2000 and 2013



(j) Change in GDP prediction using DMSP OLS night light intensity between 2000 and 2013



(k) Difference between municipality-level and county-level change in GDP prediction using surface groups



(l) Difference between municipality-level and county-level change in GDP prediction using DMSP OLS night light intensity

Source: Authors' calculations based on Landsat data, CLC data DMSP OLS data, GFSO data (GENESIS), and BKG data.
Note: Maps show the county of Wunsiedel (situated in south-east Germany at the border to the Czech Republic).

Validation of Surface Groups for Developing Countries. To investigate whether surface groups can serve as a proxy for economic activity in developing countries, we compare our approach to a Yeh et al.'s (2020) approach for African countries. While both approaches provide indicators for economic conditions, the approaches differ in the type of economic conditions they proxy. Yeh et al.'s (2020) approach uses African DHS data to construct an index for village-level asset wealth, including measures for quality of living (e.g., if households have running water), and then trains a neural network to directly predict this index from a combination of Landsat and DMSP OLS night light intensity data. In contrast, our approach intends to proxy regional economic activity as indicated in administrative statistics, and thus represents primarily industrial economic activity rather than asset wealth of villages. Moreover, by classifying Landsat pixels into the six surface groups before using them to predict economic activity, our approach offers a direct measure for land cover with a potential for applications in regional science studies. Yeh et al. (2020) thus demonstrate that satellite data can be used to predict a particular developmental characteristic (village asset wealth), while our approach demonstrates that satellite data can be trained to predict both disaggregated and potentially missing or erroneous economic activity data (e.g., GDP at disaggregated levels within a county).

To make the comparison, we produce our surface groups proxy for four African countries—Guinea, Togo, Uganda, and Zimbabwe—using the procedure we outline in Appendix 3.5.1.6.[41] Choosing these four countries ensures the fairest possible comparison, because for them Yeh et al.'s (2020) approach yields an above-average prediction quality (according to R^2 reported in figure 2 in Yeh et al., 2020). Yeh et al. (2020) provide both their village-level asset wealth index and their prediction of this index for the years available in the underlying DHS data—2012 for Guinea; 2013 for Togo; 2009, 2011, and 2014 for Uganda; and 2010 and 2015 for Zimbabwe. The locations of villages are indicated by the coordinates of their geographic centers. Similar to Yeh et al. (2020), we consider the area within a radius of 6.72 kilometers of a village's center for predicting the village's asset wealth with surface groups. For each of the four countries separately, we run an

[41] We produced the surface groups dataset we use in this chapter for the four African countries in December 2021.

OLS regression of the surface groups, the percentage of cloud cover, and year FE (if applicable) on Yeh et al.'s (2020) DHS-based asset wealth index (see table 3.19 for the regression results). The predictions derived from these regressions allow us to calculate the percentage of the variation in the asset wealth index our approach explains and to compare it to the corresponding percentage Yeh et al. (2020) explain.

The results of this comparison show that our approach also contributes to explaining the variation in Yeh et al.'s (2020) DHS-based asset wealth index. Pooling over all villages in the four countries, our approach explains 59.7 percent of the variation in the asset wealth index, compared to 73.6 percent with Yeh et al.'s (2020) approach (corresponds to red R^2 in figure 2a in Yeh et al., 2020).[42] Our surface groups proxy thus explains a significant percentage of the index, although lower than Yeh et al.'s (2020) approach. Despite our metric not being designed to identify asset wealth like Yeh et al.'s (2020) metric, our approach performs 81.1 percent as well as Yeh et al.'s (2020) metric in predicting asset wealth.

While Yeh et al.'s (2020) approach in Africa is designed to optimally predict the asset wealth index they construct from DHS data, our approach focuses on predicting a much broader proxy for regional economic activity. Both approaches explain substantial variation in the outcome variables they respectively predict. Each approach has comparative advantages and disadvantages depending on the research question (e.g., advantage for focused, village-level analyses in developing countries with the approach of Yeh et al., 2020, advantage for broader regional-level analyses in developed countries with our new approach). Satellite data can provide insight, predictability, and accuracy to various developmental indicators when trained specifically toward predicting the outcome in

[42] Calculating this indicator separately for each of the four countries and then averaging it, our approach explains 56.9 percent of the variation in the asset wealth index, compared to 78.8 percent with Yeh et al.'s (2020) approach (corresponds to black R^2 in figure 2a in Yeh et al., 2020). Conducting the analyses at the administrative district level (see table 3.20 for the OLS regression results) yields indicators of 69.4 percent vs. 81.8 percent when pooling over all districts in the four countries and weighting by the number of villages (corresponds to red weighted R^2 in figure 2b in Yeh et al., 2020), 71.4 percent vs. 90.9 percent when separating by country and weighting (corresponds to black weighted R^2 in figure 2b in Yeh et al., 2020), 50.9 percent vs. 63.2 percent when pooling and not weighting (corresponds to red unweighted R^2 in figure 2b in Yeh et al., 2020), and 48.7 percent vs. 78.8 percent when separating and not weighting (corresponds to black unweighted R^2 in figure 2b in Yeh et al., 2020).

Table 3.19: OLS prediction of asset wealth index in African countries using surface groups (village level)

	Guinea	Togo	Uganda	Zimbabwe
Dep. var.: $ln(AWI)$	(1)	(2)	(3)	(4)
$ln(builtup_{sat}+1)$	0.430***	0.559***	0.774***	0.668***
	(0.066)	(0.051)	(0.041)	(0.034)
$ln(grass_{sat}+1)$	0.321***	-0.038	-0.050	-0.248***
	(0.093)	(0.083)	(0.043)	(0.093)
$ln(crops_{sat}+1)$	-0.452***	-0.458***	-0.559***	-0.403***
	(0.115)	(0.089)	(0.052)	(0.073)
$ln(forest_{sat}+1)$	-0.298***	0.030	0.007	-0.182***
	(0.076)	(0.054)	(0.028)	(0.060)
$ln(noveg_{sat}+1)$	-0.039	-0.042	-0.265***	0.014
	(0.054)	(0.040)	(0.021)	(0.063)
$ln(water_{sat}+1)$	0.236***	-0.067***	0.021	0.184***
	(0.059)	(0.024)	(0.017)	(0.032)
Year FE	n/a	n/a	Yes***	Yes***
Cloud cover (%)	0.466	-0.081	-0.365	0.198
	(0.627)	(0.379)	(0.268)	(0.644)
N	300	300	778	793
Adj. R^2	0.624	0.663	0.533	0.457

Source: Authors' calculations based on Landsat data, CLC data, data from Yeh et al. (2020), and GADM data.
Notes: Robust standard errors in parentheses. All models include intercept. AWI denotes the DHS-based asset wealth index from Yeh et al. (2020). Available years are 2012 for Guinea; 2013 for Togo; 2009, 2011, and 2014 for Uganda; and 2010 and 2015 for Zimbabwe. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3.20: OLS prediction of asset wealth index in African countries using surface groups (district level)

	Guinea	Togo	Uganda	Zimbabwe
Dep. var.: $ln(AWI)$	(1)	(2)	(3)	(4)
$ln(builtup_{sat}+1)$	0.226	0.738**	0.424***	0.242*
	(0.211)	(0.249)	(0.076)	(0.123)
$ln(grass_{sat}+1)$	0.751**	-0.300	0.106**	-0.485**
	(0.320)	(0.307)	(0.046)	(0.224)
$ln(crops_{sat}+1)$	-0.766**	-0.611*	-0.462***	-0.125
	(0.371)	(0.290)	(0.074)	(0.205)
$ln(forest_{sat}+1)$	-1.034***	0.239	0.018	-0.085
	(0.265)	(0.197)	(0.029)	(0.128)
$ln(noveg_{sat}+1)$	-0.230	-0.016	-0.201***	0.118
	(0.198)	(0.140)	(0.030)	(0.116)
$ln(water_{sat}+1)$	0.934***	-0.043	-0.034***	0.151*
	(0.259)	(0.150)	(0.010)	(0.081)
Year FE	n/a	n/a	Yes***	Yes***
Cloud cover (%)	-19.351***	-1.979	-8.864**	-16.622
	(5.772)	(8.518)	(4.218)	(16.197)
N	34	21	397	120
Adj. R^2	0.657	0.675	0.389	0.227

Source: Authors' calculations based on Landsat data, CLC data, data from Yeh et al. (2020), and GADM data.
Notes: Robust standard errors in parentheses. All models include intercept. AWI denotes the DHS-based asset wealth index from Yeh et al. (2020). Available years are 2012 for Guinea; 2013 for Togo; 2009, 2011, and 2014 for Uganda; and 2010 and 2015 for Zimbabwe. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

context.

3.5.2.5 Surface Groups Economic Proxy

The six surface groups can be combined into a single-variable proxy by computing a predicted indicator of economic activity using our OLS model specified in Equation 3.4. To establish the external validity of such a single-variable proxy, for both GDP and household income we estimate Equation 3.4 using only one randomly selected half of the sample (the training sample). With the OLS coefficients obtained from the training-sample estimation, we predict GDP and household income for the second half of the sample (the left-out sample), that is, we predict $\widehat{ln(GDP)}$ and $\widehat{ln(HHI)}$. To assess whether this predicted single-variable proxy is as valid as the original proxy, we then reestimate the OLS model using only the left-out sample and using the single-variable proxy as independent variable instead of the original proxy. Again, we proceed similarly for DMSP OLS night light intensity to have a benchmark comparison.

Tables 3.21 and 3.22 present the estimation results for GDP and household income, respectively. In the specifications using the single-variable proxy as independent variable (columns 2 and 4), the surface groups-based proxy explains a higher percentage of the variation in economic activity than the night lights-based proxy (63.4 percent vs. 48.8 percent for GDP and 67.4 percent vs. 30.8 percent for household income). This finding corroborates the findings of the county-level analysis of GDP and of the grid-level analysis of household income. Therefore, the surface groups can provide a valid single-variable proxy of economic activity, which might be desirable when economic activity is the dependent variable in an analysis.

Finally, Table 3.23 shows the results of an OLS estimation that uses all available GDP data (2000–2018) to train the single-variable surface groups-based economic proxy. Moreover, to improve the quality of the prediction, this estimation also includes the regional percentage of pixels with cloud cover as a further indicator of potential measurement error (see appendix 3.5.2.2). This estimation underlies the time series plots of predicted GDP in Figure 3.3. In producing Figure 3.3, we follow our recommendation in Appendix 3.5.2.3

Table 3.21: OLS prediction of single-variable proxy for GDP (county level, 2000–2013)

	Surface groups		DMSP OLS night light intensity	
	Training sample	Left-out sample	Training sample	Left-out sample
Dep. var.: $ln(GDP)$	(1)	(2)	(3)	(4)
$ln(builtup_{sat}+1)$	1.371***			
	(0.051)			
$ln(grass_{sat}+1)$	-0.136***			
	(0.019)			
$ln(crops_{sat}+1)$	-0.260***			
	(0.017)			
$ln(forest_{sat}+1)$	-0.159***			
	(0.016)			
$ln(noveg_{sat}+1)$	-0.261***			
	(0.022)			
$ln(water_{sat}+1)$	-0.001			
	(0.022)			
$ln(NL_{DMSPOLS}+1)$			0.434***	
			(0.024)	
$\widehat{ln(GDP)}$ from (1)		1.014***		
		(0.033)		
$\widehat{ln(GDP)}$ from (3)				0.995***
				(0.054)
Year FE	Yes***	Yes*	Yes**	Yes***
Federal state FE	Yes***	Yes***	Yes***	Yes***
N	2,687	2,715	2,687	2,715
Adj. R^2	0.610	0.634	0.457	0.488

Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, GFSO data (GENESIS), and BKG data.
Notes: Robust standard errors in parentheses. All models include intercept. * $p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$.

Table 3.22: OLS prediction of single-variable proxy for household income (grid level, 2009–2013)

	Surface groups		DMSP OLS night light intensity	
	Training sample	Left-out sample	Training sample	Left-out sample
Dep. var.: $ln(HHI)$	(1)	(2)	(3)	(4)
$ln(builtup_{sat} + 1)$	1.414***			
	(0.003)			
$ln(grass_{sat} + 1)$	-0.126***			
	(0.002)			
$ln(crops_{sat} + 1)$	-0.371***			
	(0.002)			
$ln(forest_{sat} + 1)$	-0.066***			
	(0.002)			
$ln(noveg_{sat} + 1)$	-0.172***			
	(0.002)			
$ln(water_{sat} + 1)$	-0.212***			
	(0.002)			
$ln(NL_{DMSPOLS} + 1)$			0.955***	
			(0.003)	
$\widehat{ln(HHI)}$ from (1)		0.996***		
		(0.001)		
$\widehat{ln(HHI)}$ from (3)				0.996***
				(0.003)
Year FE	Yes***	Yes**	Yes***	Yes***
Federal state FE	Yes***	Yes**	Yes***	Yes***
N	367,668	369,958	367,668	369,958
Adj. R^2	0.676	0.674	0.307	0.308

Source: Authors' calculations based on Landsat data, CLC data, DMSP OLS data, RWI and microm (2019) data, and ESDAC data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3.23: OLS prediction of GDP using surface groups (county level, 2000–2018)

Dep. var.: $ln(GDP)$	(1)
$ln(builtup_{sat} + 1)$	1.307***
	(0.029)
$ln(grass_{sat} + 1)$	-0.114***
	(0.012)
$ln(crops_{sat} + 1)$	-0.259***
	(0.010)
$ln(forest_{sat} + 1)$	-0.187***
	(0.010)
$ln(noveg_{sat} + 1)$	-0.185***
	(0.013)
$ln(water_{sat} + 1)$	-0.006
	(0.013)
Year FE	Yes***
Federal state FE	Yes***
Cloud cover (%)	-5.029***
	(0.792)
N	7,397
Adj. R^2	0.630

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Notes: Robust standard errors in parentheses. Model includes intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

and remove outlier observations. More specifically, we consider a county-year observation an outlier if the number of *builtup* pixels in that year is more than twice as large as the median number of *builtup* pixels among all observations from the same county or if more than ten percent of the observation's pixels are covered by clouds.

3.5.2.6 Validation of Surface Groups as Proxies for Land Cover

After analyzing the quality of surface groups as a proxy for economic activity, we investigate their usefulness for economic research analyzing actual land cover (e.g., urbanization). In so doing, we use external data from administrative statistics to examine the quality of surface groups as proxies for their corresponding types of land cover. This analysis allows us to draw conclusions on the external validity of surface groups as proxies for administrative land cover, and thus to determine the usefulness of surface groups for economic applications investigating different types of land cover. Moreover, the analysis complements the assessment of internal validity in Section 3.2.2 and in Appendix 3.5.1.5.

To assess the validity of surface groups as proxies for their corresponding types of land cover, we perform separate OLS regressions for each of the six surface groups. The dependent variables in these regressions are the natural logarithms of land cover as indicated in administrative data: $ln(builtup_{adm}+1)$, $ln(grass_{adm}+1)$, $ln(crops_{adm}+1)$, $ln(forest_{adm}+1)$, $ln(noveg_{adm}+1)$, or $ln(water_{adm}+1)$. The main independent variables are the natural logarithms of the corresponding surface groups we retrieve from satellite data: $ln(builtup_{sat}+1)$, $ln(grass_{sat}+1)$, $ln(crops_{sat}+1)$, $ln(forest_{sat}+1)$, $ln(noveg_{sat}+1)$, or $ln(water_{sat}+1)$. To control for potential measurement error in the dependent and independent variables, we again include year FE and federal state FE (see appendix 3.5.2.3).

As a comparison to night light intensity makes no sense for assessing land cover, we also include the percentage of missing pixels (most likely due to cloud cover) to capture potential measurement error in the number of pixels per surface group (see appendix 3.5.2.2). The unit of observation in these regressions is the municipality, because the administrative land cover data is available at this level. The observation period is 2008 through 2015.

Table 3.24 shows the results of the regressions for each surface group. For all six surface groups, an increase in the satellite-based measure is significantly associated with an increase in the administrative measure, whether we include the control variables or not. As adjusted R^2 indicates, with 80.0 percent *builtup* is the surface group that explains most of the variation in the corresponding administrative measure (column 2), closely followed by *crops*, with 78.9 percent (column 6), and *forest*, with 76.6 percent (column 8). The surface groups *grass* (44.0 percent, column 4), *noveg* (53.5percent, column 10), and *water* (59.4 percent, column 12) explain less of the variation in the corresponding administrative measure. Although one reason for their lower predictive value may be that our classification algorithm detects these surface groups with a lower accuracy, the results of the five-fold cross-validation in Section 3.2.2 and in Appendix 3.5.1.5 suggest otherwise. Therefore, more likely is that the administrative measures do not indicate the six types of land cover with equal accuracy, due to the level of aggregation in the data (see appendix 3.5.2.2). In other words, the lower predictive values likely result from the differing aggregations of land cover types in the administrative data.

Figure 3.13 depicts the statistical distribution of the residuals by surface group, stemming from the regressions in columns 2, 4, 6, 8, 10, and 12 of Table 3.24. In line with the assessment of adjusted R^2, the distribution is much narrower and smoother for *builtup*, *crops*, and *forest* than for the other three groups. Thus differences exist in the predictive values of the surface groups.

Figures 3.14 through 3.19 show the temporal and spatial distribution of the residuals by surface group. Assessing this distribution allows us to draw further conclusions on how valid a proxy each surface groups is for its corresponding type of land cover. Regarding the combined temporal and spatial bias, *builtup* performs best. We find the smallest temporal bias for *builtup* and the smallest spatial bias for *grass*. The assessment of temporal and spatial bias again confirms differences in the predictive value of the six surface groups.[43]

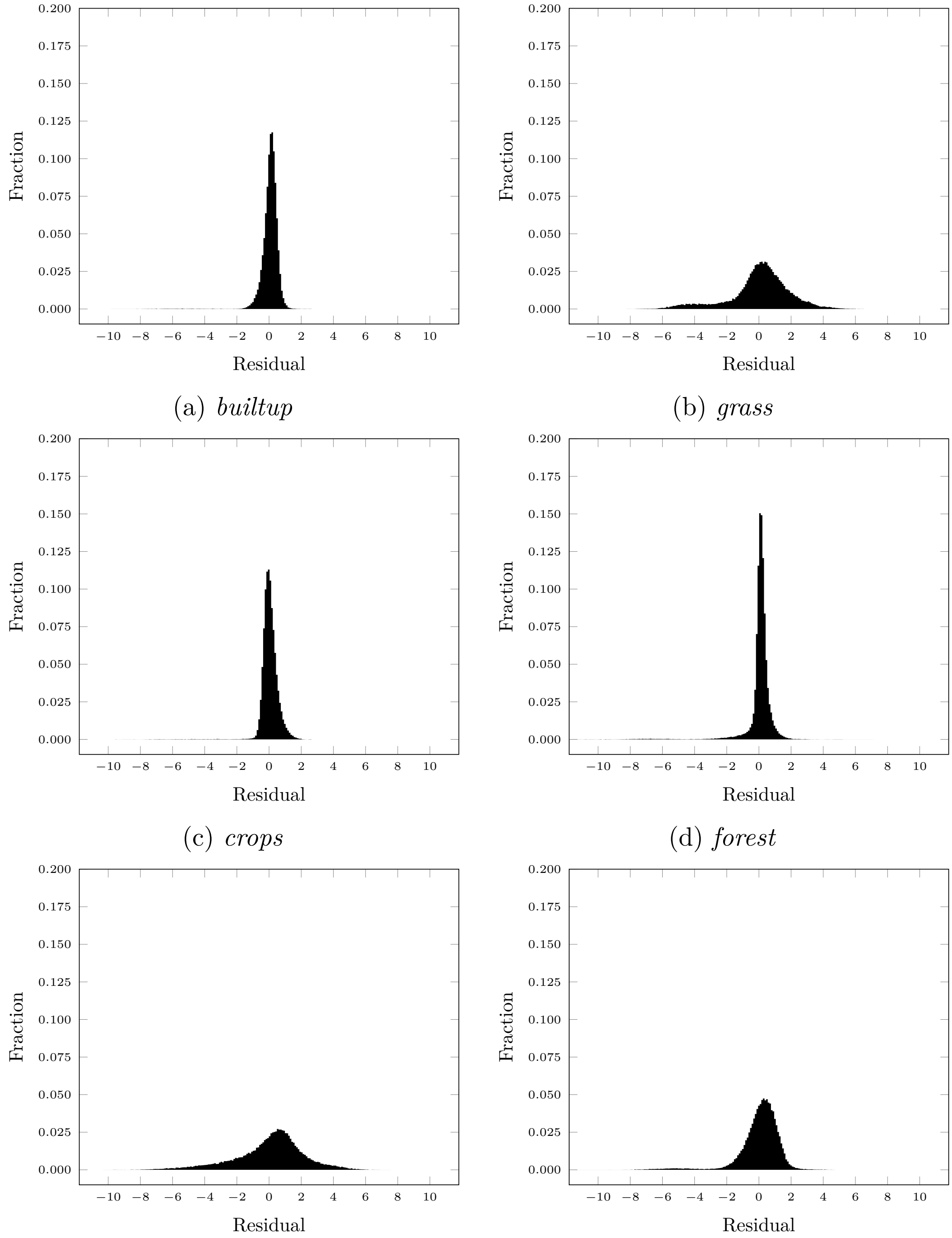
[43] More specifically, we find the following indicators of temporal and spatial bias: Regarding temporal bias, the residual has the same direction throughout all observation years in 3,799 municipalities (37.2 percent) for *builtup*, 7,516 municipalities (73.5 percent) for *grass*, 5,841 municipalities (57.1 percent) for *crops*, 4,188 municipalities (41.0 percent) for *forest*, 5,699 municipalities (55.8 percent) for *noveg*,

Table 3.24: OLS prediction of administrative land cover using surface groups (municipality level, 2008–2015)

	$ln(builtup_{adm}+1)$		$ln(grass_{adm}+1)$		$ln(crops_{adm}+1)$		$ln(forest_{adm}+1)$		$ln(noveg_{adm}+1)$		$ln(water_{adm}+1)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$ln(builtup_{sat}+1)$	0.823***	0.870***										
	(0.005)	(0.004)										
$ln(grass_{sat}+1)$			0.721***	0.795***								
			(0.009)	(0.008)								
$ln(crops_{sat}+1)$					0.769***	0.791***						
					(0.005)	(0.004)						
$ln(forest_{sat}+1)$							1.081***	1.125***				
							(0.005)	(0.004)				
$ln(noveg_{sat}+1)$									1.136***	0.881***		
									(0.006)	(0.006)		
$ln(water_{sat}+1)$											0.898***	0.926***
											(0.004)	(0.004)
Year FE	No	Yes***	No	Yes***	No	Yes***	No	Yes***	No	Yes***	No	Yes***
Federal state FE	No	Yes***	No	Yes***	No	Yes***	No	Yes***	No	Yes***	No	Yes***
Cloud cover (%)		5.200***		7.065***		5.902***		5.996***		7.134***		5.196***
		(0.161)		(0.188)		(0.113)		(0.227)		(0.202)		(0.101)
N	80,535	80,535	80,535	80,535	80,535	80,535	80,535	80,535	80,535	80,535	80,535	80,535
Adj. R^2	0.723	0.800	0.116	0.440	0.720	0.789	0.725	0.766	0.370	0.535	0.487	0.594

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Notes: Robust standard errors in parentheses. All models include intercept. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 3.13: Statistical distribution of land cover residuals by surface group



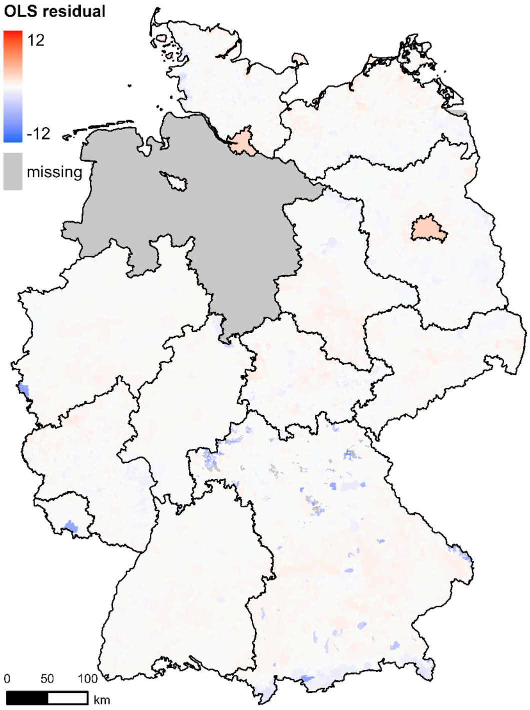
(a) *builtup* (b) *grass*

(c) *crops* (d) *forest*

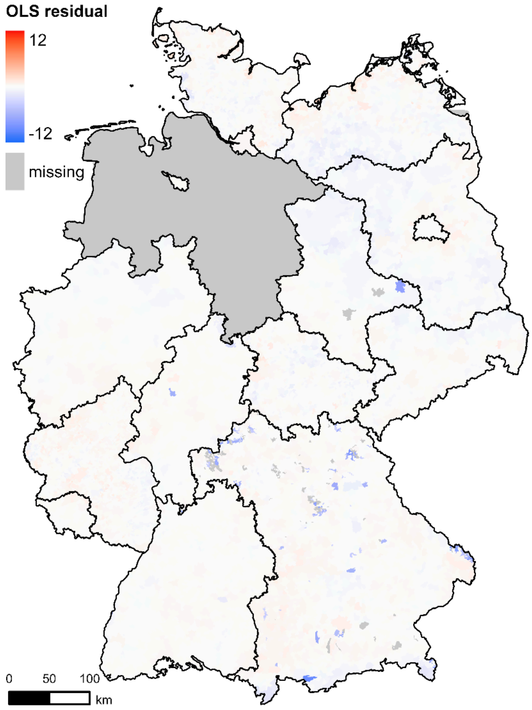
(e) *noveg* (f) *water*

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data. Note: Residuals stem from the regressions in Table 3.24 (specifications including cloud cover, year FE, and federal state FE). Figure shows histograms with a bin width of 0.1.

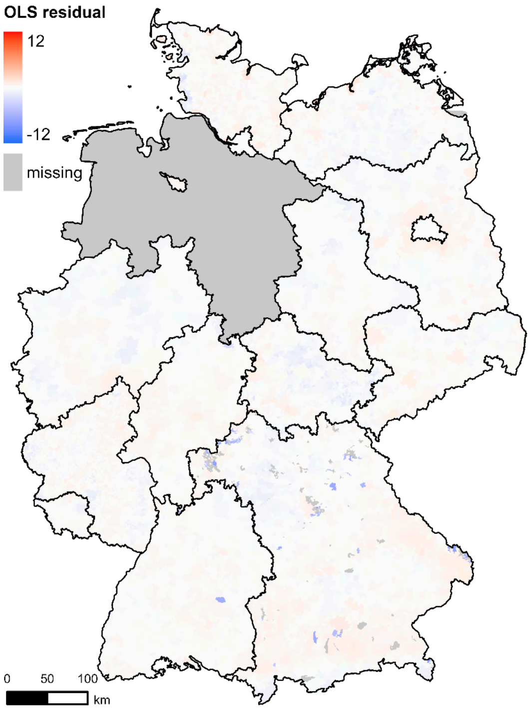
Figure 3.14: Spatial and temporal distribution of land cover residuals for built-up surfaces (surface group *builtup*)



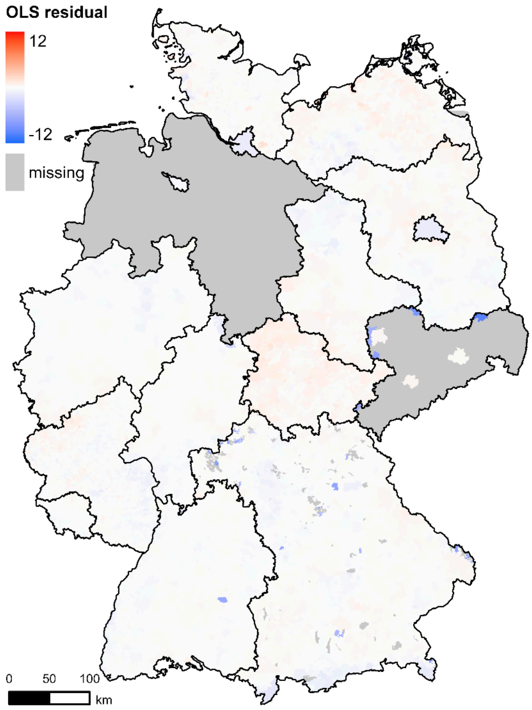
(a) 2008



(b) 2010



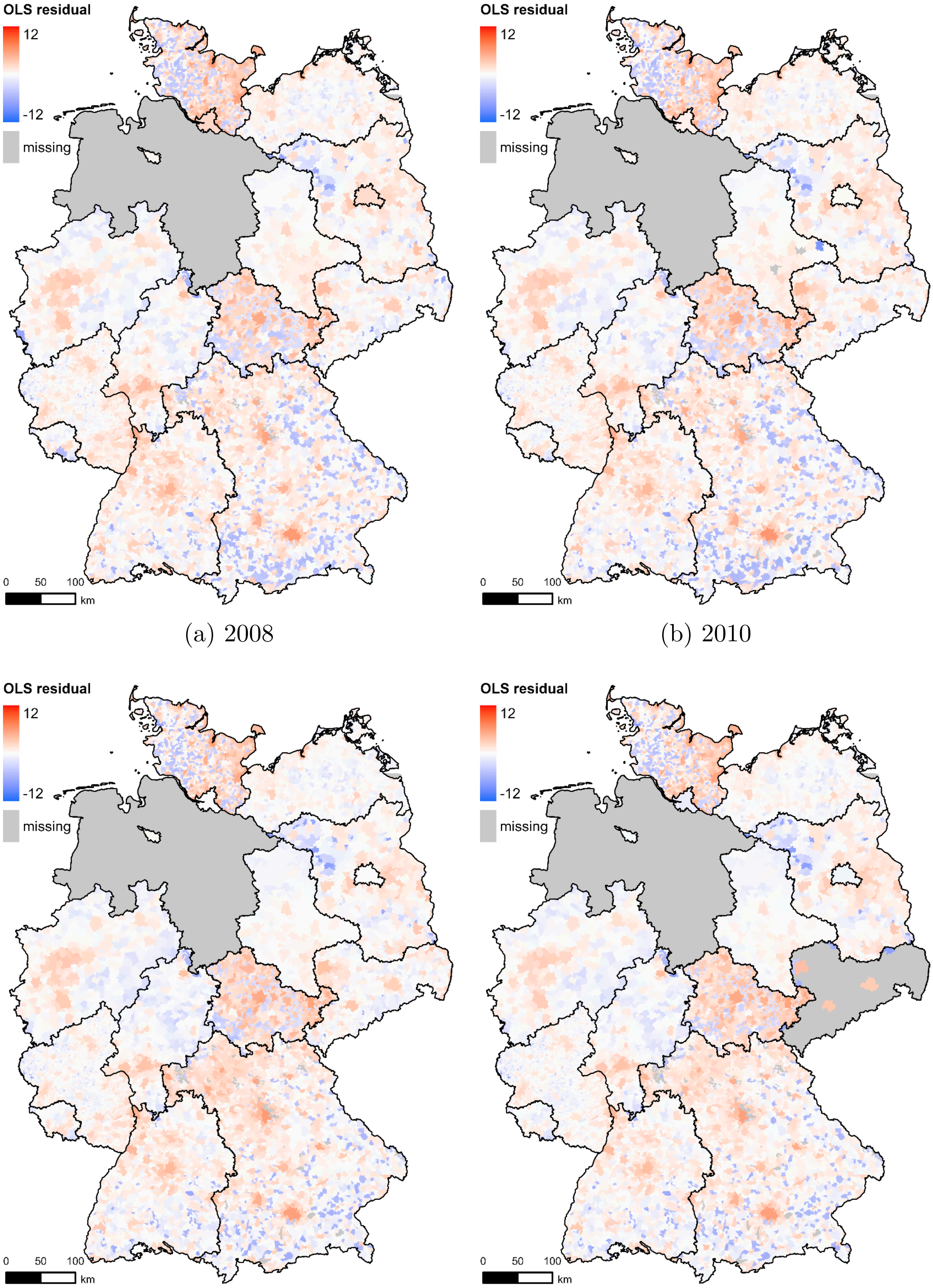
(c) 2012



(d) 2014

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Note: Maps illustrate residual that stems from the regression in column 2 of Table 3.24.

Figure 3.15: Spatial and temporal distribution of land cover residuals for grassy surfaces (surface group *grass*)



(a) 2008

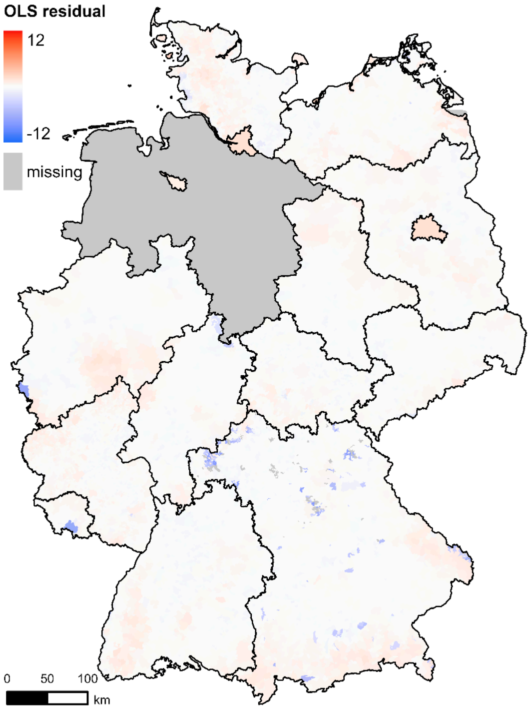
(b) 2010

(c) 2012

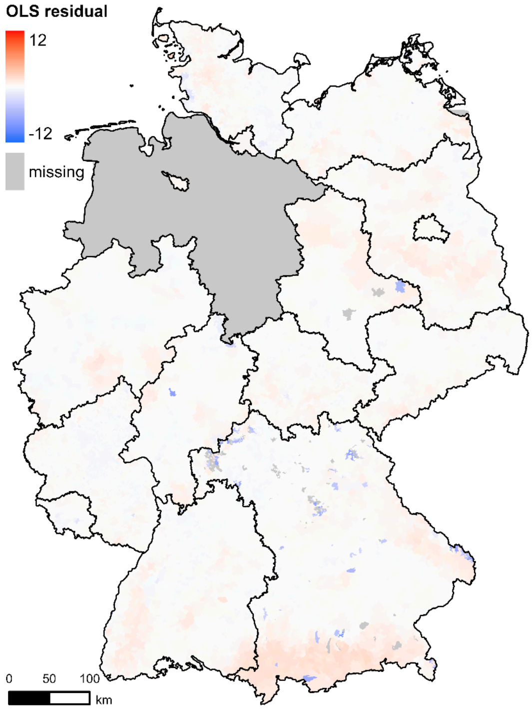
(d) 2014

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Note: Maps illustrate residual that stems from the regression in column 4 of Table 3.24.

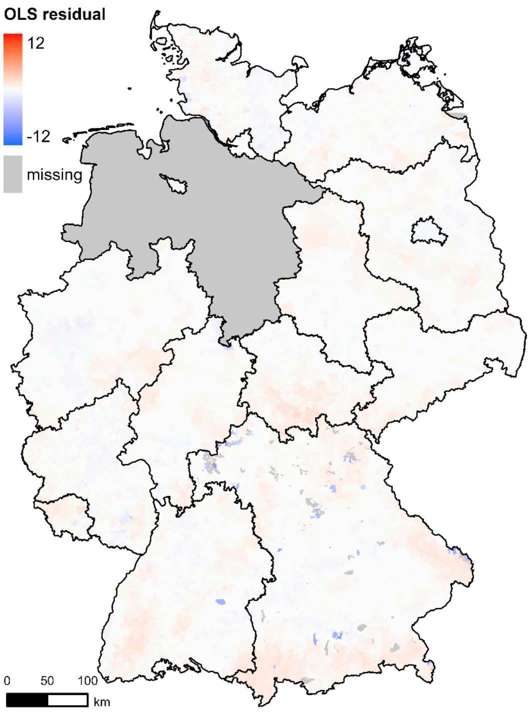
Figure 3.16: Spatial and temporal distribution of land cover residuals for surfaces with crop fields (surface group *crops*)



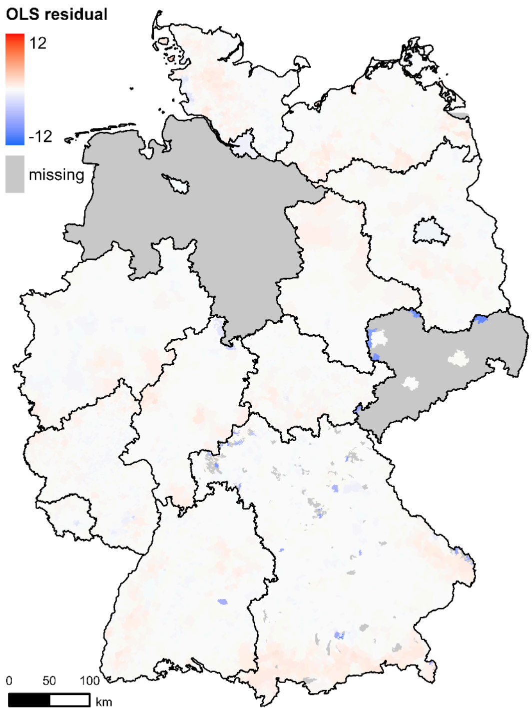
(a) 2008



(b) 2010



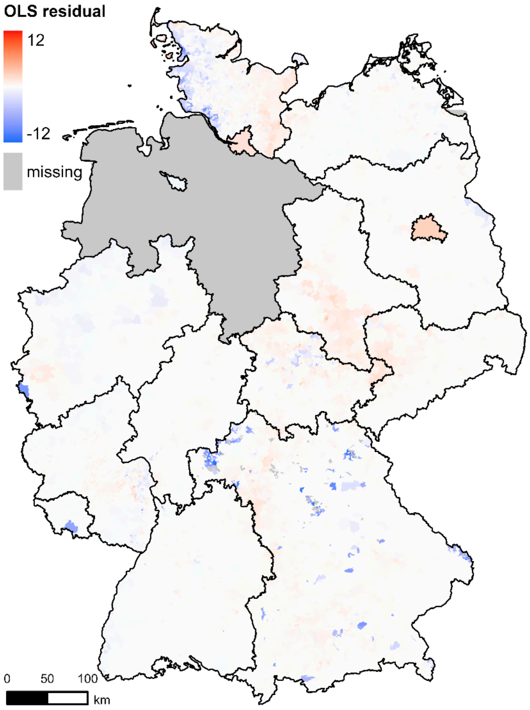
(c) 2012



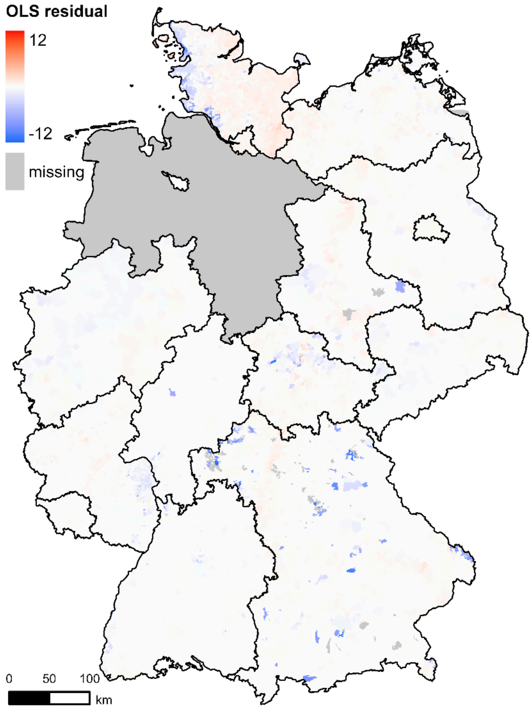
(d) 2014

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Note: Maps illustrate residual that stems from the regression in column 6 of Table 3.24.

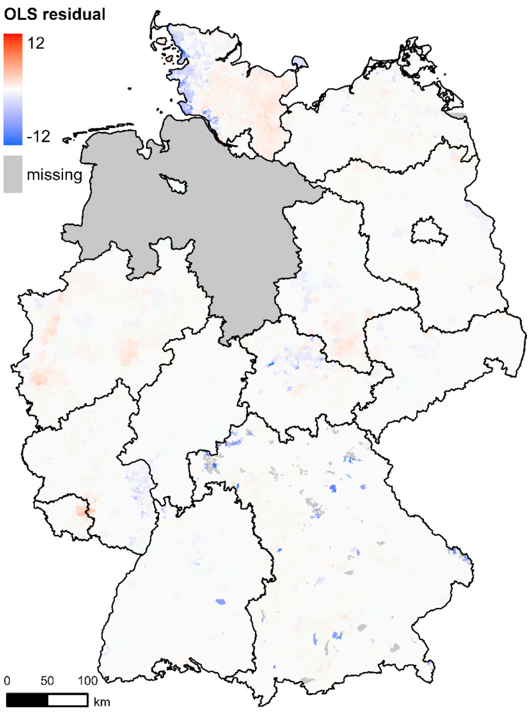
Figure 3.17: Spatial and temporal distribution of land cover residuals for forest-covered surfaces (surface group *forest*)



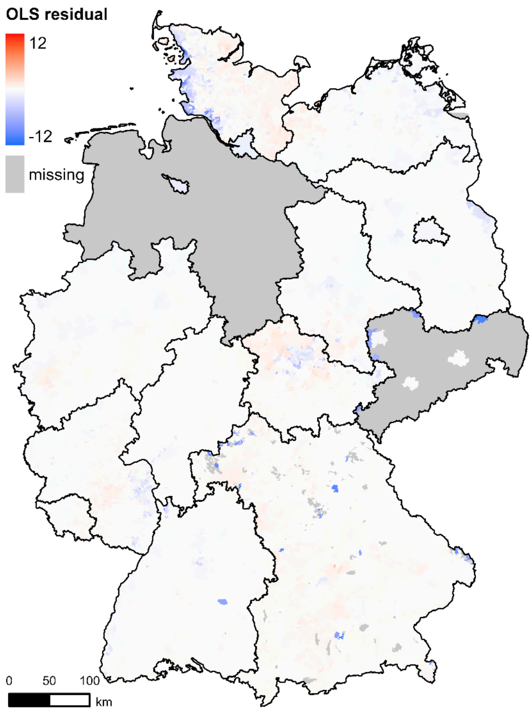
(a) 2008



(b) 2010



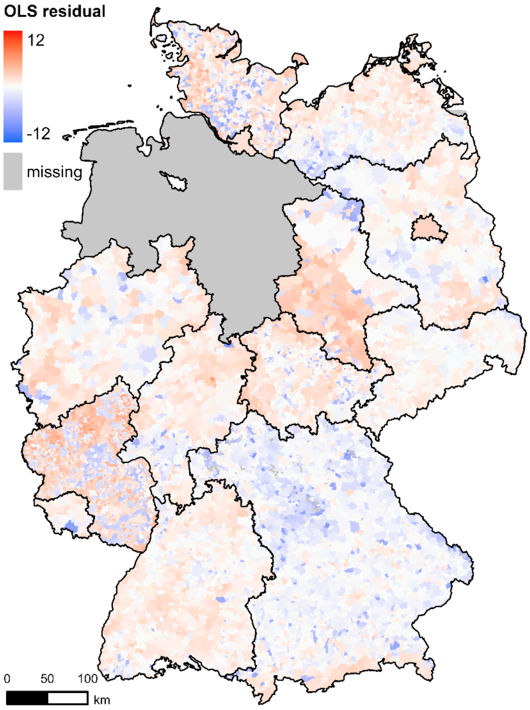
(c) 2012



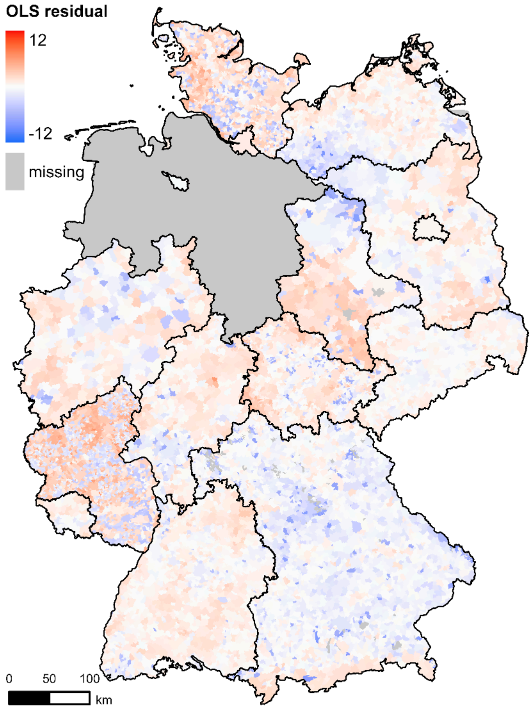
(d) 2014

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Note: Maps illustrate residual that stems from the regression in column 8 of Table 3.24.

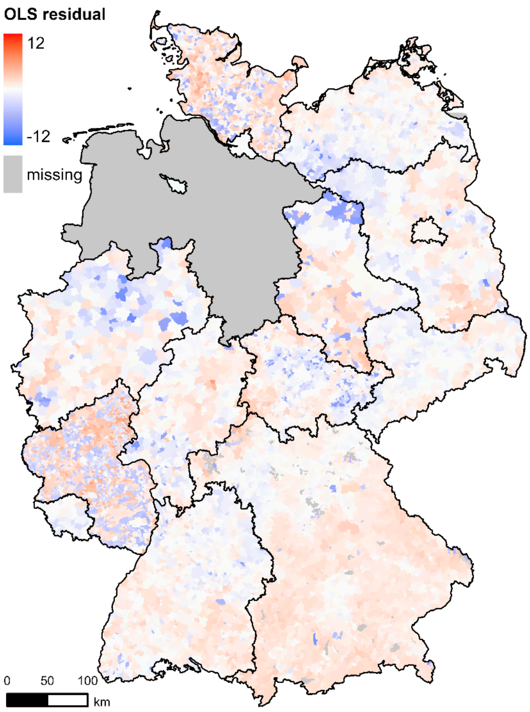
Figure 3.18: Spatial and temporal distribution of land cover residuals for surfaces without vegetation (surface group *noveg*)



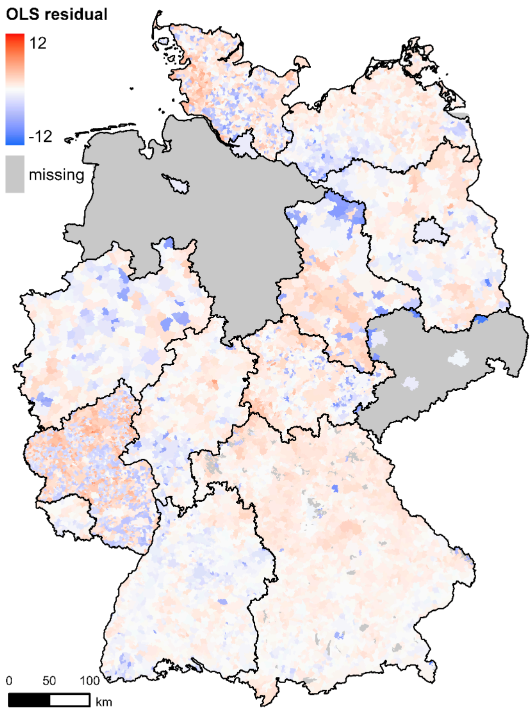
(a) 2008



(b) 2010



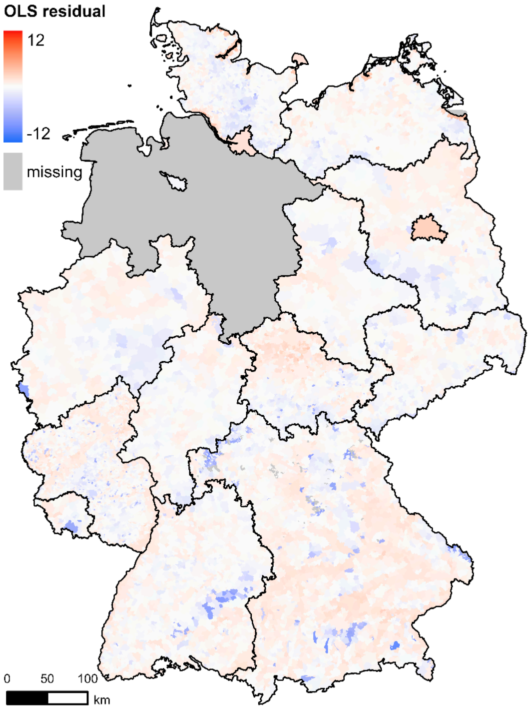
(c) 2012



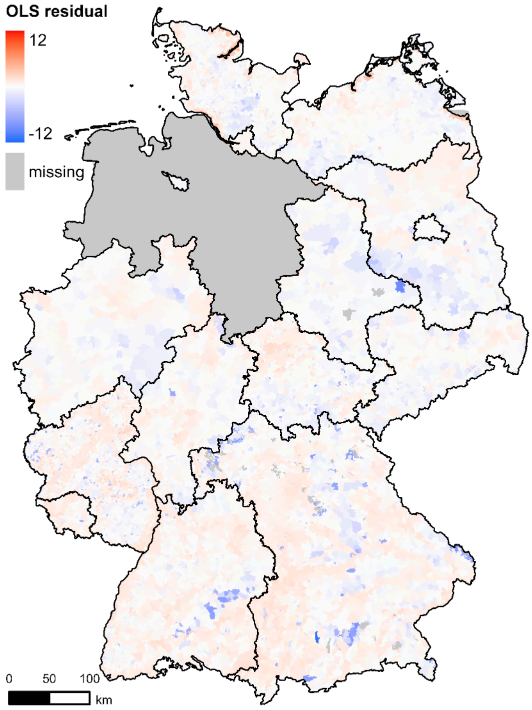
(d) 2014

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Note: Maps illustrate residual that stems from the regression in column 10 of Table 3.24.

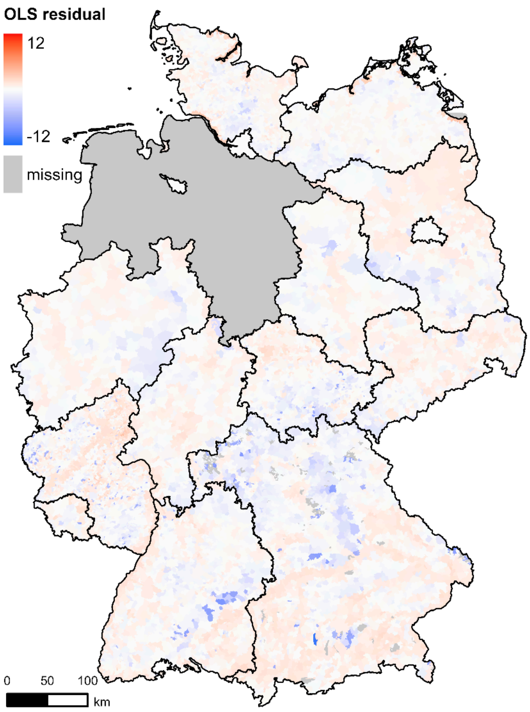
Figure 3.19: Spatial and temporal distribution of land cover residuals for water surfaces (surface group *water*)



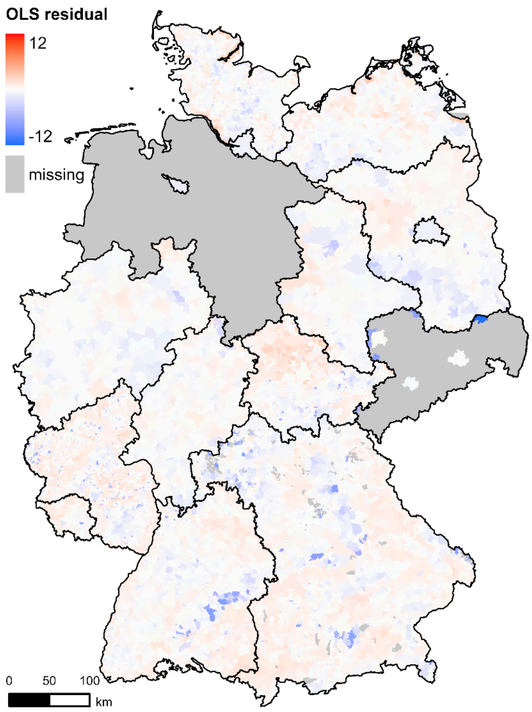
(a) 2008



(b) 2010



(c) 2012



(d) 2014

Source: Authors' calculations based on Landsat data, CLC data, GFSO data (GENESIS), and BKG data.
Note: Maps illustrate residual that stems from the regression in column 12 of Table 3.24.

In sum, we argue that the six surface groups—*builtup*, *grass*, *crops*, *forest*, *noveg*, and *water*—are useful measures for economic research that investigates actual land cover. These surface groups provide information on land cover for time periods and spatial units for which other data do not exist. The six surface groups are valid proxies for their corresponding types of land cover in administrative statistics. Based on the explained percentage of the variation in the administrative measure and the combined bias indicator, the surface group *builtup* performs best in predicting its administrative counterpart. The prediction error we observe in our analysis for some of the remaining surface groups likely originate in the differing aggregations of land cover types caused by the data structure of the administrative statistics.

and 5,172 municipalities (50.6 percent) for *water*. Regarding spatial bias, we find the same direction of the residual in all neighboring regions for 24,293 observations (30.2 percent) for *builtup*, 17,329 observations (21.5 percent) for *grass*, 27,043 observations (33.6 percent) for *crops*, 28,104 observations (34.9 percent) for *forest*, 19,393 observations (24.1 percent) for *noveg*, and 18,452 observations (22.9 percent) for *water*. The combined bias persists in 376 municipalities (3.7 percent) for *builtup*, 973 municipalities (9.5 percent) for *grass*, 998 municipalities (9.8 percent) for *crops*, 546 municipalities (5.3 percent) for *forest*, 570 municipalities (5.6 percent) for *noveg*, and 447 municipalities (4.4 percent) for *water*.

Chapter 4

The Role of Knowledge Complementarities Between Universities of Applied Sciences and Research Institutions for Regional Innovation

Part of this chapter is an extended version of early parts of the working paper "Knowledge Complementarities and Patenting: Do New Universities of Applied Sciences Foster Regional Innovation?" by Lehnert, Pfister, Harhoff, and Backes-Gellner (2020).

4.1 Introduction

While a great number of studies have identified the positive innovation effects of basic research performed at traditional UNIs (e.g., Anselin et al., 1997; Demircioglu and Audretsch, 2019; Jaffe, 1989; Toivanen and Väänänen, 2016), only a few studies have investigated the innovation effects of other types of research institutions, particularly those with a more application-oriented research focus. For example, for the U.S. Adams et al. (2003) find positive innovation effects for federal laboratories cooperating with industrial laboratories, and Popp (2017) identifies application-oriented PROs as more important for private firms in the energy sector than traditional UNIs. For Switzerland, Pfister et al. (2021) examine whether UASs lead to an increase in regional patenting activities, finding strong positive effects. In this chapter, we study the German context and explore if UASs also have an impact on regional innovation activities. The novel contribution of this chapter is that we investigate whether—in addition to the effects of stand-alone UASs—complementarity effects between UASs and other research institutions arise in a diverse landscape of coexisting research institutions.

Thus, on top of innovation effects that stem from a single institution itself, we address the question of whether additional effects arise due to the existence of an ecosystem of UASs and other research institutions in close proximity. We expect that different types of knowledge creation in regions where research institutions coexist lead to complementarity effects, thus enhancing the capacity for innovation in regions with bundles of different research orientations. In particular, we expect UASs to foster the transfer of the scientific knowledge that basic research institutions produce into applied research, leading to a higher degree of innovation in comparison to regions with a stand-alone UAS.

This chapter thus analyzes the single and combined regional innovation effects of UASs and neighboring research institutions. We are able to (a) identify the regional innovation effects of the introduction of UASs in Germany and (b) to separate which part of the effect goes back to UASs themselves and which part goes back to complementarities resulting from the embeddedness of UASs in a diverse research landscape of coexisting research institutions.

UASs in Germany have started to conduct applied research in the 1980s and offer three-year study programs awarding bachelor-level degrees. Since adapting to the Bologna Process in the 2000s, in many cases they also offer master-level degrees,[1] but do not bestow doctoral degrees. To analyze the UAS effect on regional innovation, we examine the evolution of patents in Germany in treated and untreated regions since the 1980s. We investigate at the regional level whether, in the regions where UASs are located, the UAS

[1] In 2000, the first UASs began to offer master-level study programs and by the end of the 2000s, the large majority of German HEIs had adapted their study programs to the Bologna Process (Key and Seeßelberg, 2012; Petzina, 2005). The findings of this chapter are robust to restricting the observation period to the pre-2005 period, that is, the period before the first UAS students graduated from master-level programs (see appendix 4.8.8).

effect on innovation depends on the existence of other research institutions—UNIs and PROs. A larger UAS effect on innovation in regions with a basic research institution (i.e., a UNI or a PRO that focuses on basic research) would clearly support the existence of complementarities between basic and applied research.

For our analysis, we draw on two main data sources to compose a novel dataset covering all research institutions (more than 700) in Germany. First, we collect information on the exact locations and opening years of all research institutions. For this purpose, we augment official directories of HEIs (including UNIs and UASs) and PROs with extensively researched information on the particular campus and institute locations of each HEI and PRO, and the opening years of all identified locations. Second, to measure innovation outcomes we use patent data from the EPO Worldwide Patent Statistical Database (October 2019 version). Combining the two data sources provides us with a rich dataset that is highly suitable for analyzing the impact of knowledge complementarities between different types of research institutions on innovation outcomes. As we use patent data to measure innovation outcomes, we focus on institutions specializing in STEM, because patents would not adequately represent the potential innovation effects of other fields such as the social sciences or the arts.

As an identification strategy, we follow a growing literature that uses the openings of HEIs to estimate causal effects (e.g., Eyles and Machin, 2019; Kamhöfer et al., 2019; Pfister et al., 2021; Toivanen and Väänänen, 2016). In so doing, we exploit variation in the location and the timing of UAS openings to investigate if patenting activities increase after these openings.

As we do not want to claim that the spatial and temporal distribution of UASs is entirely random in Germany, we need to account for potential endogeneity in the location of UASs to identify their effect on innovation activities. That is, we need to control for time-invariant and time-variant regional economic factors that may determine both the existence of a UAS in a region and innovation activities. To control for time-invariant regional economic factors, we apply FE estimation. To control for time-variant regional economic factors, we employ disaggregated historical data. However, such disaggregated economic data is unavailable for a time period reaching as far back as the 1980s (when UASs began to conduct applied research) and at the very detailed regional level required for our analysis. Therefore, we use a novel self-developed measure based on daytime satellite data to proxy economic activity at highly disaggregated regional units, a method that studies analyzing night lights data show to be valuable when other data is unavailable (e.g., Castelló-Climent et al., 2018; Henderson et al., 2012).

We derive this novel proxy in Chapter 3 with machine-learning techniques and show that it validly proxies economic activity in an analysis using data on economic activity available for limited time series and regional units. As the proxy dates back until the 1980s, it covers a sufficiently long time series for our analysis of German UASs. Moreover, the proxy contains detailed information on different types of land cover—measured in six different surface groups (e.g., built-up land, cropland, forests)—thereby allowing a precise approximation of regional economic activity. Therefore, using the surface groups as control variables in our estimations provides a novel solution to solving endogeneity problems resulting from the non-random spatial and temporal distribution of UAS openings under the assumption that the proxy captures all factors determining regional economic activity and, in turn, UAS locations.

Our results show that stand-alone UASs have a statistically significant positive effect on patenting activities. Furthermore, the UAS effect increases substantially in regions where other research institutions exist at the time of the UAS opening. A diverse research landscape featuring different types of research knowledge thus leads to strong knowledge complementarities.

This chapter makes two contributions to the literature on the innovation effects of different types of research institutions (e.g., Intarakumnerd and Goto, 2018; Popp, 2017; Toivanen and Väänänen, 2016). First, we study UASs, a type of HEI that focuses on applied research and that only very few studies have considered so far (e.g., Pfister et al., 2021). Second, previous research has largely neglected interactions between different types of research institutions and thus the role of the surrounding research landscape. The distinct specializations among research institutions in Germany allow us to explicitly analyze such interactions.

The chapter proceeds as follows. Section 4.2 reviews the relevant literature and develops our hypotheses. Section 4.3 provides detailed information on UASs in Germany and basic information on other research institutions in the German education and innovation system (UNIs and PROs). Section 4.4 describes the data we use for our analysis and Section 4.5 our methodological approach. Section 4.6 presents our estimation results and provides further robustness checks. Section 4.7 concludes.

4.2 Literature Review

Many studies have emphasized the positive influence of research institutions performing basic research on regional innovation activities. In an important study, Jaffe (1989) finds evidence of spillovers from UNI research to private-sector R&D, in STEM-related industries in particular, thereby contributing to local patenting activities. Other studies attesting to the positive relationship between UNI research and the evolution of patents include Autant-Bernard (2001), Cowan and Zinovyeva (2013), Leten et al. (2014), and Toivanen and Väänänen (2016). This literature thus extensively documents the positive effect of basic research institutions on regional innovation.

In addition, other types of PROs, in particular those with a more applied research focus, positively affect innovation. In their review of the traditional linear model of innovation, Leyden and Menter (2018) argue that basic public research alone—without accompanying applied research—is insufficient for generating knowledge spillovers from the public to the private sector. A variety of PROs can serve this applied function within an innovation system, with empirical studies providing evidence for the positive effect of these PROs (e.g., Comin et al., 2019; Intarakumnerd and Goto, 2018; Popp, 2017). Case studies of regional innovation systems in Germany also show that PROs can play a key role in fostering innovation (e.g., Broekel and Graf, 2012; Graf, 2011).

A particular type of applied research institution belonging to the higher education sector is the UAS. Pfister et al. (2021) show that UASs have a large positive effect on patenting in Switzerland. For Germany, previous research has not analyzed the innovation effect of UASs separately, but only in combination with other institutions in the research and innovation system. Fritsch and Slavtchev (2007) provide evidence for a positive innovation effect of HEIs (UNIs and UASs combined) using six years of patent data from the German Patent Office. Furthermore, using two waves of Community Innovation Survey data, Robin and Schubert (2013) find that collaboration between firms and public research institutions (UNIs, UASs, and PROs combined) increases innovation at the firm level. Other studies use cross-sectional surveys to investigate the role of public research institutions (differentiating between UNIs, UASs, and PROs) for innovation in firms, providing mixed results (Beise and Stahl, 1999; Fritsch and Schwirten, 1999).

Few studies explicitly address cooperation between research institutions. Fritsch and Schwirten (1999) find that in the regions surveyed in their analysis, cooperation with other research institutions is more common for UNIs and PROs than for UASs, which in turn cooperate more often with firms. Investigating how the number of cooperation partners affects innovation in firms (but without explicitly differentiating between different types of cooperation partners), Becker and Dietz (2004) argue that a mix of heterogeneous cooperation partners creates synergies that further increase R&D activities. However, the innovation effect resulting from knowledge complementarities between different types of research institutions remains unexplored.

In sum, based on the empirical evidence we expect that both basic and applied public research institutions increase innovation and that cooperation between the two potentially enhances this effect. In this chapter, we contribute to the literature by studying the introduction of UASs, allowing us to disentangle the UAS effect on patenting from the effects of other research institutions and the joint effects.

Assuming that knowledge spillovers are geographically concentrated, as many studies argue (e.g., Audretsch and Feldman, 1996; Cabrer-Borrás and Serrano-Domingo, 2007; Holl et al., 2022; Ponds et al., 2007), we expect that opening a UAS in a region where it can cooperate with a basic research institution yields a larger innovation effect than opening a UAS elsewhere. Moreover, if the two types of institutions produce complementary applied research knowledge, opening a UAS in a region where another type of applied research institution already exists might also lead to positive knowledge spillovers. In regions where no other institution exists, we expect UASs to increase innovation, but to a lesser extent.

4.3 The German Landscape of Research Institutions

4.3.1 Higher Education Institutions (HEIs)

In Germany, two types of public[2] HEIs perform research in the STEM fields: UNIs and UASs.[3] Like all education institutions in Germany, UNIs and UASs fall under the jurisdiction of the Länder[4] governments (BMBF, 2018a).[5] Although both institutions award equivalent bachelor's and master's degrees (ISCED 2011 levels 6 and 7), UNIs hold the exclusive right to award doctoral degrees throughout the years that we analyze in this chapter (Enders, 2010; BMBF, 2004).[6] Moreover, UNIs focus on basic research and impart basic research skills to their students, whereas UASs emphasize vocational practice and applied research (BMBF, 2004).

While some UNIs have a centuries-long tradition, UASs are rather new institutions.[7] In comparison to UNIs, teaching at UASs targets students who have completed an apprenticeship (a dual VET program) and aims at knowledge relevant for vocational practice and problem-solving (BMBF, 2004). This focus also manifests at many UASs in, for example, bachelor-degree programs that include a practical semester in the form of an internship at a firm (BMBF, 2004; Lackner, 2019).

[2] We restrict our analysis to public HEIs. We exclude private ones primarily because relatively fewer of them are active in STEM-related research (see Buschle and Haider, 2016).

[3] In addition to UNIs and UASs, the German higher education sector also comprises universities of education (*Pädagogische Hochschulen*), of theology (*Theologische Hochschulen*), of art and music (*Kunsthochschulen*), and of public administration (*Verwaltungsfachhochschulen*) (BMBF, 2018a). However, as these institutions specialize in subjects unrelated to STEM and thus produce innovations that do not usually result in patents, we do not consider them in our analysis.

[4] Germany is divided into 16 federal states called Länder.

[5] Very few HEIs that train civil servants fall partly under federal jurisdiction, such as the German University of Administrative Sciences in Speyer, the German Police University in Münster, or the Universities of the German Federal Armed Forces in Hamburg and Munich.

[6] Recently, the federal state of Hesse granted research-intensive UASs a limited right to award doctoral degrees (Meurer, 2018).

[7] The first UASs in Germany were introduced only in 1968 as teaching institutions (BMBF, 2004).

Only in 1985 did an amendment of the German Higher Education Framework Act add applied research to the purpose of UASs (Enders, 2010; Kulicke and Stahlecker, 2004; Wissenschaftsrat, 2002). By engaging in applied research projects jointly conducted with firms, faculty members at UASs maintain their vocational practice, which they can then pass on to their students (Hinz et al., 2016; Kulicke and Stahlecker, 2004). A large part of the research projects that UASs undertake is in STEM fields such as IT, materials science, or mechanical engineering (Kulicke and Stahlecker, 2004).

Due to their focus on applied research and vocational practice, UASs are an important research partner for local firms. Small and medium-sized firms in particular profit from the knowledge that UASs generate, because these firms value the application-oriented knowledge and the vocational practice, and because these firms do not possess the capacities for carrying out research projects by themselves (Hachmeister et al., 2015; Kulicke and Stahlecker, 2004). Participation in joint research projects with larger firms or in publicly funded research projects is also important for UASs (Hachmeister et al., 2015).

The UASs' focus on applied—in comparison to basic—research and on combining vocational practice with applied research in their teaching systematically distinguishes UASs from UNIs. Thus, by providing applied research knowledge to local players in the R&D ecosystem, UASs can positively contribute to regional innovation activities. When cooperating with research institutions that have a stronger focus on basic research, UASs can contribute to the transfer of basic scientific knowledge, thereby increasing regional innovation. Through this cooperation, a local combination of basic and applied research institutions can be an important driver of (regional) innovation outputs.

4.3.2 Public Research Organizations (PROs)

In addition to UNIs and UASs, which combine teaching and research, four PROs exclusively conduct research. These four PROs—the Max Planck Society, the Leibniz Association, the Helmholtz Association, and the Fraunhofer Society—thus constitute another unique pillar of the German research and innovation system. They are publicly funded and act independently, serving functions that are complementary to those of HEIs (see, e.g., EFI, 2010). Each PRO maintains several research institutes throughout Germany, often employing hundreds and sometimes even thousands of researchers (BMBF, 2018b).

The four PROs serve different functions and have different ranges of activity:

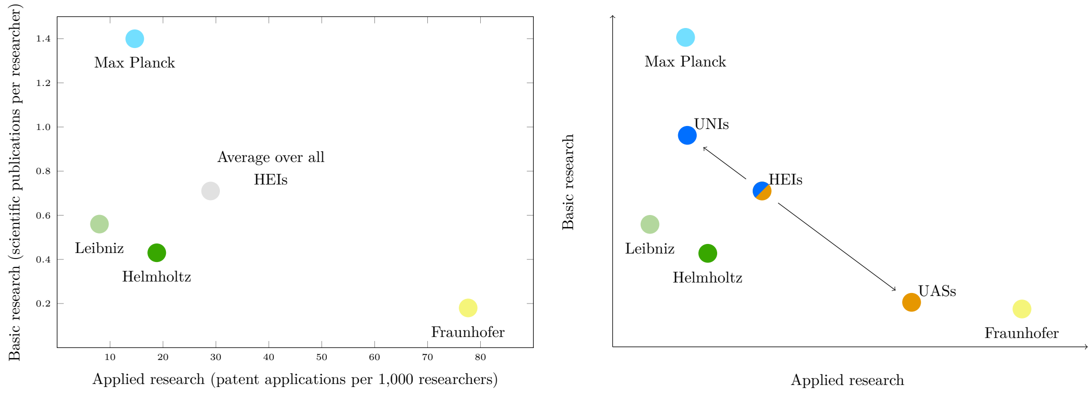
- The Max Planck Society conducts basic research in the natural sciences, the life sciences, the social sciences, and the humanities (BMBF, 2018a), that is, in both STEM and non-STEM fields. The Max Planck Society autonomously chooses its research subjects and aspires to scientific excellence, a goal also reflected in the organization's high degree of internationalization (BMBF, 2018a; Hohn, 2010). Examples include the Max Planck Institute for Plasmaphysics or the Max Planck Institute of Biochemistry (see BMBF, 2018b).
- The Fraunhofer Society engages in applied research projects on a range of topics (e.g., health, environment, mobility, energy), with the majority being STEM-related (BMBF, 2018a). As the Fraunhofer Society receives government funding that calls for research collaboration, both private and public demand drive the organization's choice of research projects (Hohn, 2010). Examples include the Fraunhofer Institute for Solar Energy Systems or the Fraunhofer Institute for Laser Technology (see BMBF, 2018b).
- The Helmholtz Association comprises research centers that are active in technology-intensive (i.e., STEM-related) fields with a long-term perspective, such as aeronautics and materials science (BMBF, 2018a). Therefore, the organization is involved in the transfer of basic research knowledge to technological products (Hohn, 2010). In comparison to the Max Planck Society and the Fraunhofer Society, the public funding that the Helmholtz Association receives is tied to specific subjects or projects (Hohn, 2010). Examples include the German Aerospace Center or the Helmholtz Centre for Infection Research (see BMBF, 2018b).
- The Leibniz Association originally put those institutes that did not fit into the structures of the other three organizations under one umbrella (Hohn, 2010). Consequently, the Leibniz Association has the broadest research scope of the four PROs, ranging from basic to applied research in both STEM and non-STEM fields, and

having a decentralized organizational structure (BMBF, 2018a; Hohn, 2010). In addition to research institutes, the Leibniz Association also comprises non-research institutes such as museums and further education institutes (e.g., training centers) (BMBF, 2018a; Hohn, 2010). Examples include the Leibniz Institute of Polymer Research or the Leibniz Institute of Plant Genetics and Crop Plant Research (see BMBF, 2018b).

Figure 4.1a graphically depicts the profiles of research institutions by comparing their publication and patenting activities (EFI, 2010). Corresponding to their mandates, the Max Planck Society focuses on basic research (mainly publications) and the Fraunhofer Society focuses on applied research (mainly patents). The Helmholtz and Leibniz Associations range somewhere in between, with the Leibniz Association tending towards basic research. With respect to HEIs, the study unfortunately did not differentiate between UNIs and UASs, but the average over all HEIs locates in the middle of the spectrum. However, if one were to depict UNIs and UASs separately, their positioning would correspond to that in Figure 4.1b in line with the two HEIs' mandates, that is, UNIs tending towards basic research and UASs towards applied research.

In accordance with their differing profiles, PROs strategically engage in different types of research cooperation with HEIs (BMBF, 2018a). Again, such cooperation can play a key role in fostering the transfer of basic scientific knowledge to actual applications. For PROs with a basic research focus, such as the Max Planck Society, cooperation with UASs can provide an important source of applied research knowledge, while UASs, in turn, might value such cooperation as a source of basic research knowledge. Furthermore, PROs with an applied research focus, such as the Fraunhofer Society, might also profit from knowledge complementarities with UASs.

Figure 4.1: Profiles of research institutions in Germany



(a) EFI (2010): Positioning of institutions according to empirical analyses

(b) Authors' extension: Approximate positioning of UNIs and UASs according to their legal mandates

Source: Figure 4.1a shows an illustration based on EFI (2010, p. 40). Figure 4.1b shows authors' extensions based on the legal mandates of HEIs.
Notes: The original analysis in EFI (2010) is based on an analysis of publications in the Science Citation Index and patent applications per researcher (in full-time equivalents) for three periods, 1994–1996, 1999–2001, and 2004–2006 (dots show averages over these three periods). The original analysis does not differentiate between UASs and UNIs. The dots for UASs and UNIs in Figure 4.1b represent authors' assessment of the activity of UASs and UNIs according to their legal mandates.

4.4 Data

To analyze the innovation effect of UASs, we use three different datasets. First, we use patent data to measure innovation. Second, we use self-collected data on the locations and opening years of all HEI campuses and PRO institutes in Germany to identify regions treated by one or more of these institutions. Third, we use the proxy measure for regional economic activity from Chapter 3 to control for endogeneity in the spatial and temporal distribution of UAS openings in our empirical analysis. Combining these three datasets provides ample information for investigating the innovation effect of UASs and the role of knowledge complementarities between UASs and other research institutions.

The first dataset is the EPO Worldwide Patent Statistical Database (October 2019 version), from which we extract two measures of regional innovation as our outcome variables. This data offers complete information on patents from 1980 and thus goes back farther in time than 1985, when UASs began to conduct applied research. The patent information includes, among other items, the exact geographic locations of inventors, the application date, and the number of patent citations three years after publication.[8] To assign every inventor to a German municipality, we geocode[9] the inventor locations in the EPO data and then link the geocoded addresses to administrative geodata provided by the BKG.[10] Using these assignments, we follow Pfister et al. (2021) and compute the fractionated[11] number of patents per municipality and year (patent quantity, $PQUAN$), as well as a patent's average number of citations per municipality and year (patent quality, $PQUAL_{cit}$) as innovation outcomes. To ensure a complete citation window of three years for all patents in the EPO data, we end the observation period in 2015.

From the second dataset, we construct our treatment variables by determining which

[8] The EPO data also includes citation information for five and ten years after publication, respectively. We analyze the citation lag of three years to get a conservative estimate and to be able to use a longer time series.

[9] We geocode addresses using the HERE application programming interface. See https://developer.here.com/ (last retrieved on November 21, 2019).

[10] Available from http://www.geodatenzentrum.de/geodaten/gdz_rahmen.gdz_div?gdz_spr=deu&gdz_akt_zeile=5&gdz_anz_zeile=1&gdz_unt_zeile=15&gdz_user_id=0 (last retrieved on May 5, 2017).

[11] For example, if a patent lists one inventor from municipality A, one inventor from municipality B, and one inventor from abroad, the patent counts as $\frac{1}{3}$ of a patent for municipality A and $\frac{1}{3}$ for municipality B. The remaining $\frac{1}{3}$ does not enter our estimations.

types of HEIs and PROs exist in each municipality at a given point in time. To create this dataset, we had to exert extensive efforts in data collection to get the location and timing of the openings of all HEI campuses and all PRO institutes.

As a starting point for the HEI data collection, we used data from the German Rectors' Conference.[12] This data contains, among other items, the name, type, main address, and opening year of every HEI (i.e., every UNI and every UAS) that existed in Germany on January 1, 2017,[13] but it does not contain their individual campus locations or study fields. Furthermore, to identify HEIs that closed before January 1, 2017 (and are thus not part of the German Rectors' Conference data),[14] we use data on student numbers ranging as far back as 1998 from the GFSO's GENESIS database.[15] This data contains the names and (indirectly through disappearance in the data) the closing years of these HEIs. For the construction of our treatment variables, we had to augment these HEI data by performing extensive online searches in which we studied for every single HEI the available information on the web to nail down the HEI's individual history, campus locations, and study fields. To obtain this detailed information, we researched online the campus addresses, campus opening years,[16] and campus profiles (i.e., the departments or study programs that each offers),[17] gathering this information primarily from the HEI's websites. Drawing on the campus profiles, we categorize HEI campuses according to whether they are active in STEM or not. If a campus offers at least one study program in STEM or has a STEM department, we categorize it as a STEM campus.

[12] Available from http://www.hs-kompass2.de/kompass/xml/download/hs_liste.txt (last retrieved on January 20, 2017).

[13] We exclude HEIs that offer only distance learning, because we do not expect their innovation effects, if any, to be locally concentrated.

[14] Nonetheless, we might miss some HEIs that closed before 1998. However, drawing on historical descriptions of HEIs (see also section 4.3.1), we find this number to be very small, therefore not biasing our estimations.

[15] Available from https://www-genesis.destatis.de/genesis/online/data;jsessionid=7CF9926E22E6DDE18B248F7571AD3949.tomcat_GO_2_3?operation=abruftabelleBearbeiten&levelindex=1&levelid=1500539721416&auswahloperation=abruftabelleAuspraegungAuswaehlen&auswahlverzeichnis=ordnungsstruktur&auswahlziel=werteabruf&selectionname=21311-0002&auswahltext=&werteabruf=Werteabruf (last retrieved on July 20, 2017).

[16] If the websites or other sources do not indicate an opening year for a campus, we assume it to be the opening year of the respective institution as indicated in the German Rectors' Conference data.

[17] For example, the main campus of the Weihenstephan-Triesdorf UAS with the department of bioengineering sciences (among others) is in Freising. However, the Weihenstephan-Triesdorf UAS has a second campus in Weidenbach (more than 100 kilometers from Freising), which contains the departments of agriculture, food, and nutrition, and of environmental engineering.

To obtain initial information for the PRO data collection (i.e., for the collection of data on all institutes belonging to one of the four PROs Max Planck Society, Leibniz Association,[18] Helmholtz Association,[19] and Fraunhofer Society), we draw on BMBF (2018b). This source contains the name and city of all such institutes that existed in 2018, but not the exact addresses, opening years, or information on closed institutes. Again, we perform extensive online searches to augment this data by collecting the addresses, opening years, and profiles of all institutes belonging to one of the four PROs (including information on closed institutes) and categorize the institutes according to whether they are active in STEM or not.

The second dataset thus contains the exact addresses of all UAS campuses, UNI campuses, and PRO institutes. This information allows us to determine the spatial distribution of these campuses and institutes, providing us with a rich longitudinal dataset. From 1980 through 2015 (the observation period of this chapter), 134 public UASs have campuses in 347 locations in Germany, 212 of which are STEM campus locations.[20] These STEM campus locations are geographically distributed across 142 (of a total of 11,266) municipalities.[21] Furthermore, 99 UNIs have 310 STEM campus locations distributed across 74 municipalities, the Max Planck Society has 145 STEM institute locations distributed across 50 municipalities, the Leibniz Association has 101 STEM institute locations distributed across 49 municipalities, the Helmholtz Association has 114 STEM institute locations distributed across 52 municipalities, and the Fraunhofer Society has 161 STEM institute locations distributed across 84 municipalities. Appendix 4.8.2 shows the distribution of STEM campus and STEM institute locations across municipalities for the year 2015. For better readability, if not explicitly stated otherwise, we drop the specification "STEM" when we refer to UAS STEM campuses, UNI STEM campuses, or

[18] The Leibniz Assocation, formerly named Blue List Partnership and Blue List Science Assocation, emerged in the 1990s from the Blue List institutes (see, e.g., Brill, 2017). Former Blue List institutes also enter our analysis as Leibniz institutes.

[19] In 1995, the Association of National Research Centres was transfered into the Helmholtz Association (see, e.g., Hoffmann and Trischler, 2015). Institutes belonging to the former Association of National Research Centres also enter our analysis as Helmholtz institutes.

[20] Here, a location comprises all departments or branches with the same postal address. Thus one location can consist of multiple departments or branches.

[21] Throughout all observation years, we use the territorial status of January 1, 2017, because it was the most current one when we started work on this chapter.

PRO STEM institutes.

Third, to account for endogeneity in the spatial and temporal distribution of UAS campus locations, we use a novel proxy for regional economic activity. Using this proxy is necessary, because no other direct or indirect measure of regional economic activity or other factors related to UAS openings and patenting exists for a time series dating as far back as 1985 (when UASs began to conduct applied research), and at a sufficiently disaggregated regional level. This chapter uses the novel proxy for regional economic activity based on daytime satellite imagery, a proxy that we develop in Chapter 3.[22]

We demonstrate in Chapter 3 that our proxy—like other satellite-based proxies such as night light intensity (e.g., Chen and Nordhaus, 2011)—constitutes a valuable proxy for regional economic activity. Closely following the geographic literature, the proxy is constructed by applying machine-learning techniques to daytime satellite data, thereby classifying the surfaces of extremely small regional units (30 times 30 square meter pixels) into six different surface groups: built-up land (i.e., artificial materials such as buildings), grassland, cropland, forest, land without vegetation (e.g., bare rock, pure soil), and water. This information on surface groups can then be aggregated to any regional unit, including municipalities,[23] to provide detailed information on regional surface structures. The proxy is available from 1984, and thus earlier than any other satellite-based proxy for regional economic activity. Using fine-graind regional data on economic activity that is only available for a limited number of years, we show in Chapter 3 that the regional combination of surface groups is a highly valid proxy for regional economic activity. Further information on the construction of the proxy and its validation can be found in Chapter 3.

22 In accordance with Chapter 3, this chapter uses the most recent surface groups dataset produced in January 2022.

23 On average, a German municipality consists of about 56,000 pixels.

4.5 Methodology

Combining all data sources described in Section 4.4, we exploit the temporal and spatial variation in the openings of UAS campuses and estimate the following FE model:

$$
\begin{aligned}
Y_{i,t} = &\beta_0 + \beta_1 UAS_{i,t-3} + \beta_2 UAS_{i,t-3} \times UNI_i + \\
&\beta_3 UAS_{i,t-3} \times MaxPlanck_i + \beta_4 UAS_{i,t-3} \times Leibniz_i + \\
&\beta_5 UAS_{i,t-3} \times Helmholtz_i + \beta_6 UAS_{i,t-3} \times Fraunhofer_i + \\
&i + t + X_{i,t-3} + SG_{i,t} + \mu_{i,t}
\end{aligned}
\tag{4.1}
$$

with i indicating the municipality, t the year ranging from 1984 (the first year in the surface groups data)[24] through 2015 (the last year in the patent data containing complete citation information); Y the dependent variable (patent quantity or patent quality); UAS the treatment status by a UAS campus; UNI, $MaxPlanck$, $Leibniz$, $Helmholtz$, and $Fraunhofer$ the existence of a campus or institute belonging to the respective institution or organization at the time of the UAS campus opening; and μ the error term. To investigate change rates instead of changes in absolute numbers, we follow previous studies that use patenting activities as an indicator for regional innovation (Feldman and Florida, 1994; Pfister et al., 2021; Schlegel et al., 2022) and take the natural logarithms of the dependent variables Y after adding the value 1, that is, we observe $ln(PQUAN + 1)$ and $ln(PQUAL_{cit} + 1)$, respectively (appendix 4.8.1 shows the descriptive statistics for $PQUAN$ and $PQUAL_{cit}$). The vector X includes controls for the openings of other campuses or institutes during the observation period (i.e., the vector includes $UNI_{i,t-3}$, $MaxPlanck_{i,t-3}$, $Leibniz_{i,t-3}$, $Helmholtz_{i,t-3}$, and $Fraunhofer_{i,t-3}$). The vector SG contains the surface groups.

To estimate Equation 4.1 and thus to identify the innovation effect of a UAS opening, we assign each German municipality either to a treatment group or to the control group. To the treatment groups, we assign all municipalities that are treated by a UAS campus.

[24] Appendix 4.8.6 provides the estimation results without using surface groups as control variables for 1980 through 2015. In these estimations, the pattern of the results is identical to that in our main results, but the coefficients are larger in magnitude. Therefore, surface groups do indeed account for potential endogeneity in the spatial and temporal distribution of UAS locations.

We consider a municipality treated by a UAS campus if the municipality is located within a 25-kilometer[25] travel-distance[26] radius of the campus. For our analysis, we consider 1985 (i.e., when UASs began to conduct applied research according to their legal mandate) as the earliest possible treatment year, even if a UAS already existed as a teaching institution before 1985. We allow for a treatment lag of three years in Equation 4.1, because the regional research structures of UASs need some time to establish after the announcement of applied research becoming a goal of UASs.[27] To the control group, we assign all municipalities not treated by a UAS campus.

To investigate knowledge complementarities between UASs and other research institutions, we form six treatment groups by differentiating treated municipalities according to whether they are simultaneously treated by other types of research institutions. In so doing, we apply the same 25-kilometer radius to retrieve indicators for whether a municipality is treated by a UNI campus or a PRO institute at the time of the UAS campus opening. We then interact the treatment variable UAS with these indicators to assign treated municipalities to six, non-mutually exclusive, treatment groups: municipalities treated by a UAS campus but not by any other research institution (term $UAS_{i,t-3}$ in equation 4.1), municipalities treated by a UAS campus and by a UNI campus ($UAS_{i,t-3} \times UNI_i$), municipalities treated by a UAS campus and by a Max Planck institute ($UAS_{i,t-3} \times MaxPlanck_i$), municipalities treated by a UAS campus and by a Leibniz institute ($UAS_{i,t-3} \times Leibniz_i$), municipalities treated by a UAS campus and by a Helmholtz institute ($UAS_{i,t-3} \times Helmholtz_i$), and municipalities treated by a UAS campus and by a Fraunhofer institute ($UAS_{i,t-3} \times Fraunhofer_i$). These interaction terms allow us to

[25] We choose the threshold of 25 kilometers because in Germany, the majority of the working population (79.2 percent in 2016) commute 25 kilometers or less. See https://www.destatis.de/DE/ZahlenFakten/GesamtwirtschaftUmwelt/Arbeitsmarkt/Erwerbstaetigkeit/TabellenPendler/Pendler1.html (last retrieved on February 19, 2019). Empirical evidence on the innovation effects of research institutions also suggests a concentration of these effects within a 25-kilometer radius of an institution (Helmers and Overman, 2017; Pfister et al., 2021). We provide estimation results with varying treatment radii as a robustness check in Appendix 4.8.3 to demonstrate that the choice of treatment radius does not alter our main results.

[26] We compute the travel distance between the geographical center of a municipality and a campus or institute location using the HERE application programming interface. See https://developer.here.com/ (last retrieved on July 10, 2019).

[27] For example, graduates are one important channel of knowledge transfer (see, e.g., chapter 2 and Andrews, 2020) and the minimum number of years a UAS student needs to graduate is three years.

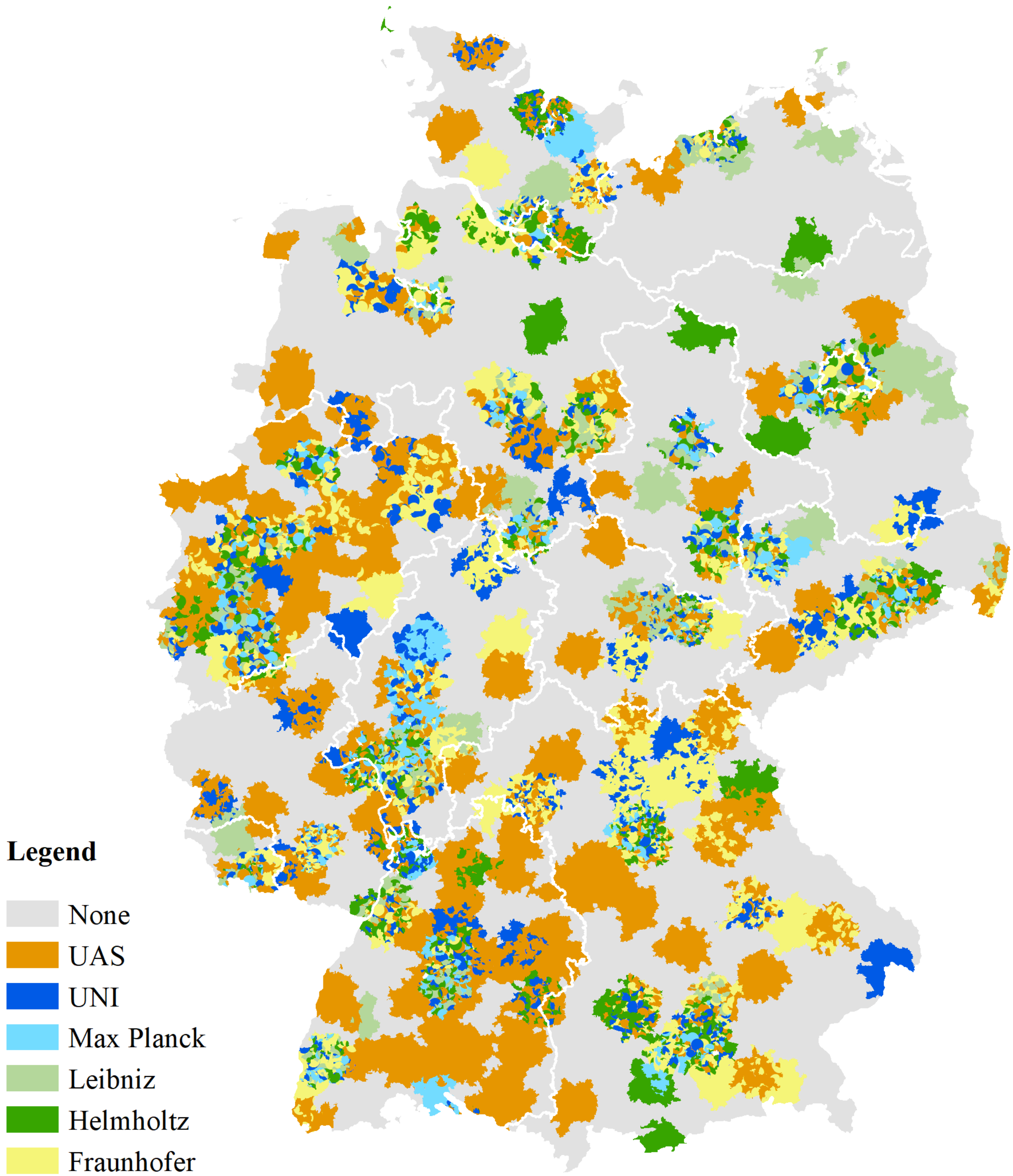
disentangle the effect of opening a UAS in regions with no preexisting research knowledge to draw upon from the effect resulting from knowledge complementarities between UASs and other research institutions. Moreover, by including the vector X in Equation 4.1 we control for the openings of UNI campuses or PRO institutes after the opening of a UAS campus.

Figure 4.2 shows the distribution of treatment regions in Germany for 2015, the last year of observation in our analysis. Of the 11,266 municipalities, 3,720 (33.0 percent) lie within the 25-kilometer treatment radii of UAS campuses, 2,263 (20.1 percent) within the radii of UNI campuses, 1,842 (16.4 percent) within the radii of Max Planck institutes, 1,206 (10.7 percent) within the radii of Leibniz institutes, 1,404 (12.5 percent) within the radii of Helmholtz institutes, and 2,278 (20.2 percent) within the radii of Fraunhofer institutes. Of the 3,720 municipalities treated by a UAS campus, at the time of the UAS campus opening 1,364 (36.7 percent) were also treated by a UNI campus, 828 (22.3 percent) by a Max Planck institute, 337 (9.1 percent) by a Leibniz institute, 362 (9.7 percent) by a Helmholtz institute, and 705 (19.0 percent) by a Fraunhofer institute.

To identify the causal UAS effect on innovation, we need to control for potential endogeneity in the locations of UAS campuses. As we cannot assume that the establishment of UASs in Germany followed a quasi-random pattern, we account for the potential endogeneity in UAS campus openings in three ways. First, to account for time-invariant municipality characteristics that determine both UAS locations and patenting activities, we use FE estimation (municipality FE i in equation 4.1). Second, to capture time trends that are common to all municipalities, we add year FE (t). Third, as the municipality FE and year FE still do not solve the problem of time-variant regional factors potentially determining both UAS locations and patenting, we account for (at least some of) these time-variant factors in our estimations by including the surface groups developed in Chapter 3 in Equation 4.1 as proxies for time-variant regional economic activity (vector SG).[28] Neglecting such factors would bias our results if these factors were correlated with

[28] To account for potential measurement error in the surface groups variables (see chapter 3), we make two adjustments. First, in addition to the six surface groups, we add a variable indicating the percentage of cloud cover. Second, we follow the procedure suggested in Chapter 3's Appendix 3.5.2.5 and identify municipality-year observations with twice as many built-up pixels than the median number of built-up

Figure 4.2: Municipalities treated by STEM campuses and STEM institutes in 2015 (25-kilometer travel-distance radius)



Source: Authors' illustration with geodata from the BKG, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.

both the treatment (i.e., the timing and locations of UAS campus openings) and patenting.

From 1984 through 2015, a German municipality (which comprises on average about 7,800 acres), consists on average of 12.4 percent built-up land, 22.8 percent grassland, 32.0 percent cropland, 24.9 percent forest, 1.9 percent land without vegetation, and 3.1 percent water.[29] Comparing the treatment and control groups shows that treated municipalities feature on average more built-up land (15.5 percent vs. 10.8 percent), cropland (33.0 percent vs. 31.5 percent), and land without vegetation (2.1 percent vs. 1.8 percent), but less grassland (20.3 percent vs. 24.0 percent) and forest (23.3 percent vs. 25.7 percent).[30] The treated and control regions thus differ in their regional surface structures. Therefore, our novel proxy captures differences in regional economic structures before and after UAS openings, thereby controlling for time-variant regional heterogeneity potentially determining UAS campus locations.

4.6 Results

4.6.1 Main Results

Table 4.1 shows the FE estimation results for the dependent variable patent quantity. These estimations show that UASs have a positive and significant effect on patent quantity. This effect is substantially larger in regions where other research institutions exist at the time of the UAS campus opening. According to our preferred specification in column 7, which includes the full set of interactions with UNI campuses and PRO institutes and thus considers the entire research knowledge that new UASs can draw upon, the opening of a UAS campus significantly increases patent quantity both in regions without another research institution and in regions with another research institution. In comparison

pixels within a given municipality over all years or with more than ten percent cloud cover as potential outliers in the surface groups data. If possible, we replace these outlier values for all surface groups by the average of a municipality's previous and following years' values. For all remaining potential outlier observation that are not replaceable through this procedure, we assume complete cloud cover.

[29] The remaining 3.0 percent are unclassified or classified as outliers due to, e.g., cloud cover in the satellite imagery (see chapter 3).

[30] The percentage of water surfaces is similar in both groups (3.1 percent).

to regions without another research institution, the UAS effect is larger in regions with institutions focusing on basic research (Max Planck institutes), applied research (Fraunhofer institutes) or a mixture of both (Helmholtz institutes). The full set of interaction terms thus unveils that complementarities between UASs and other research institutions boost the UAS effect on patent quantity.

The FE estimation results on patent quality in Table 4.2 reveal the same overall pattern as the results on patent quantity. Our preferred specification with the full set of interaction terms in column 7 shows that UASs significantly increase patent quality by 2.3 percent in regions where a Max Planck institute coexists, indicating that the research knowledge of Max Planck institutes constitutes a valuable resource for the quality of innovations that UASs produce. This finding might result from the Max Planck Society's mission of aspiring to scientific excellence, suggesting that UASs need strong accompanying basic research knowledge to produce high-quality innovations.

In sum, the FE estimations suggest that UASs can tap their full potential as drivers of regional innovation in regions where other research institutions coexist. For patent quantity, we also consistently find positive interaction effects if UASs open in regions where they can draw upon basic research knowledge that Max Planck institutes provide, applied research knowledge that Fraunhofer institutes provide, or a combination of both that Helmholtz institutes provide. For patent quality, we find a positive interaction effect of UASs only in regions where they can draw upon very strong basic research knowledge that Max Planck institutes generate. The complementarities between UASs and other research institutions can result from a variety of mechanisms, such as movements of faculty, joint research projects, or cooperation with local firms. Although we cannot determine these potential mechanisms, our results present strong evidence that knowledge complementarities between UASs and other research institutions foster regional innovation.

Table 4.1: FE estimation results on patent quantity (25-km travel-distance radius, 3-y. lag)

Dep. var.: $ln(PQUAN+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.061***	0.041***	0.041***	0.057***	0.052***	0.041***	0.027***
	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
$UAS_{i,t-3} \times UNI_i$		0.064***					0.014
		(0.012)					(0.013)
$UAS_{i,t-3} \times MaxPlanck_i$			0.114***				0.079***
			(0.015)				(0.016)
$UAS_{i,t-3} \times Leibniz_i$				0.062***			-0.016
				(0.022)			(0.023)
$UAS_{i,t-3} \times Helmholtz_i$					0.124***		0.058**
					(0.022)		(0.023)
$UAS_{i,t-3} \times Fraunhofer_i$						0.116***	0.078***
						(0.015)	(0.017)
Municipalities	11,266	11,266	11,266	11,266	11,266	11,266	11,266
N	341,885	341,885	341,885	341,885	341,885	341,885	341,885
R^2 (within)	0.095	0.096	0.096	0.096	0.096	0.096	0.097
R^2 (between)	0.173	0.187	0.213	0.179	0.183	0.201	0.226
R^2 (overall)	0.128	0.135	0.151	0.132	0.136	0.144	0.159

Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, surface groups, and controls for $UNI_{i,t-3}$, $MaxPlanck_{i,t-3}$, $Leibniz_{i,t-3}$, $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.2: FE estimation results on patent quality (citations) (25-km travel-distance radius, 3-y. lag)

Dep. var.: $ln(PQUAL_{cit}+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.010***	0.007**	0.006**	0.010***	0.010***	0.008***	0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$UAS_{i,t-3} \times UNI_i$		0.011**					0.003
		(0.005)					(0.006)
$UAS_{i,t-3} \times MaxPlanck_i$			0.024***				0.023***
			(0.006)				(0.007)
$UAS_{i,t-3} \times Leibniz_i$				0.007			-0.006
				(0.009)			(0.010)
$UAS_{i,t-3} \times Helmholtz_i$					0.005		-0.008
					(0.009)		(0.010)
$UAS_{i,t-3} \times Fraunhofer_i$						0.015**	0.009
						(0.007)	(0.008)
Municipalities	11,266	11,266	11,266	11,266	11,266	11,266	11,266
N	341,885	341,885	341,885	341,885	341,885	341,885	341,885
R^2 (within)	0.023	0.024	0.024	0.023	0.023	0.024	0.024
R^2 (between)	0.076	0.085	0.102	0.077	0.078	0.086	0.106
R^2 (overall)	0.038	0.040	0.043	0.038	0.039	0.040	0.044

Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, surface groups, and controls for $UNI_{i,t-3}$ $MaxPlanck_{i,t-3}$ $Leibniz_{i,t-3}$ $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$ $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.6.2 Robustness Checks

To check whether our main results are robust to alternative model specifications, we perform four robustness checks. First, we investigate whether these results are sensitive to decreasing or increasing the treatment radius of 25 kilometers we apply in the main specification. Second, we assess whether the main results are sensitive to choosing shorter or longer treatment lags than the three-year lag. Third, we exclude municipalities in East Germany (the former German Democratic Republic), for which the patent data have fewer observation years. Fourth, we consider two additional patent quality measures (number of claims and family size) to examine whether our main findings hold for other dimensions of patent quality.

First, to assess the role of the radius of 25 kilometers—and thus the spatial concentration of the UAS effect on innovation—that we choose for the assignment of municipalities to the treatment and control groups, we estimate Equation 4.1 with (a) decreased radii of 15 and 20 kilometers, and (b) increased radii of 30 and 35 kilometers. These estimations allow us to examine the spatial persistence of the UAS effect on innovation by analyzing (a) whether the effects of UASs are more locally concentrated (by decreasing the treatment radius) or (b) whether UASs exert effects on regions beyond the 25-kilometer radius (by increasing the treatment radius). Appendix 4.8.3 plots the coefficients of the corresponding estimations. We find the same overall pattern of effects as with the 25-kilometer radius for all specifications. For both patent quantity and patent quality, the UAS effect in regions without any coexisting research institutions becomes larger and significant when decreasing the treatment radius, but smaller and insignificant when increasing this radius. This finding suggests that the impact of a UAS campus opening concentrates in regions closer to the UAS campus.

Second, to investigate whether and, if so, how the timing of the treatment lag—and thus the persistence of the treatment effect over time—affects our results, we repeat our estimations with treatment lags varying between zero and ten years in Appendix 4.8.4. Again, the overall pattern of effects remains unchanged. Interestingly, the positive UAS effect on patent quality in regions without coexisting research institutions turns significant after a lag of five years, indicating a more long-term effect. One potential explanation for this finding is that establishing processes and cooperative projects leading to higher-quality innovations takes longer time. Moreover, for patent quality, the interaction of UASs with Max Planck institutes turns insignificant after seven years, whereas the interaction with Fraunhofer institutes turns significant.

Third, we perform a robustness check that excludes municipalities in East Germany. For these municipalities, the patent data begin only after the German reunification in 1991. The smaller number of observations for East German municipalities and the resulting smaller number of pre-treatment observations might affect our main estimation results. Therefore, we repeat our main analysis for West German municipalities only. Appendix 4.8.5 shows the results of these regressions. Again, the pattern of the results is similar to our main analysis that additionally includes East German municipalities. The only notable exception is, somewhat surprisingly, an insignificant effect on patent quantity and a significant negative effect (significant at the five-percent level) on patent quality for UASs in regions with Helmholtz institutes (column 7 in table 4.6). This finding suggests that in the production of high-quality innovation, UASs and Helmholtz institutes act as substitutes in West Germany, potentially because in these regions Helmholtz institutes already fulfill the research tasks of newly opened UASs.

Fourth, as patent citations might not capture all dimensions of a patent's quality, we consider two additional patent quality measures that complement the analysis of patent citations.[31] Following Pfister et al. (2021), we use the number of claims ($PQUAL_{claims}$, an indicator for a patent's technological scope) and family size ($PQUAL_{size}$, the number of countries where the invention has legal protection) as additional outcome variables. Appendix 4.8.7 shows the estimation results for these two variables.

The analysis of the additional patent quality measures yields two important insights. First, for both measures, number of claims and family size, UASs in regions without coexisting research institutions have a positive and highly significant effect (columns 7 in tables 4.9 and 4.10). Second, none of the interaction terms remains significantly positive.

[31] For an overview of patent quality indicators, see, e.g., Squicciarini et al. (2013).

Moreover, the interaction with Leibniz institutes becomes negative and significant (at the one percent level) for both additional patent quality measures (columns 7 in tables 4.9 and 4.10) and the interaction with UNIs becomes negative and significant (at the five percent level) for number of claims (column 7 in tables 4.9). Again, this finding suggest that UASs and other institutions can be substitutes in the production of high-quality innovations.

In essence, the robustness checks confirm our main finding of UASs and their complementarities with other research institutions positively influencing innovation, but also suggest heterogeneity in the underlying mechanisms. The robustness checks provide evidence that UASs yield more locally concentrated effects, indicating that spatial proximity to potential cooperation partners is important for UASs to increase innovation. Moreover, we find more long-term effects—particularly for patent quality as measured by patent citations—than specified in our main estimation parameters, implying that UASs take some time to produce high-quality innovations. Dynamic patterns in the complementarities also appear to arise over time, as the long-term positive interaction with Fraunhofer institutes for patent quality (citations) indicates. The few negative interactions that we detect for West Germany and for the additional patent quality measures suggest that some underlying mechanisms leading to these specific negative interactions exist. The patterns we find might result from movements of faculty or changing cooperation partners. Therefore, future research needs to explore the exact mechanisms behind the complementarities between UASs and other research institutions.

4.7 Conclusion

This chapter analyzes the innovation effect of UASs in Germany and focuses in particular on the role of knowledge complementarities between UASs and other institutions coexisting in a diverse research landscape. We exploit variation in the location and timing of the UAS openings in Germany to compare the development of patenting activities in municipalities with a UAS and those without one. We analyze whether the effect of UASs varies within regions in which other types of research institutions coexist to assess complementarities between these institutions and UASs. To econometrically deal with endogeneity in the location and timing decisions of the UAS openings, we (1) estimate FE models to account for time-invariant regional characteristics, (2) include year FE to capture time trends common to all regions, and (3) include the novel proxy for regional economic activity derived from daytime satellite data in Chapter 3 to control for time-variant regional economic factors potentially determining UAS campus locations.

Our results show that UASs have a statistically significant positive effect on patent quantity (i.e., the number of patent applications in a municipality) and on patent quality (i.e, the average number of patent citations three years after publication). This finding is in line with previous results for UASs in Switzerland (Pfister et al., 2021). In addition, the German context allows us to show that the UAS effect is substantially larger when the scientific knowledge of other research institutions is available in a UAS region. These results confirm the view that stand-alone UASs can contribute to regional innovation, but, most importantly, our results point out that strong complementarity effects arise on top of the stand-alone UAS effect: UASs can better develop their full potential when they have the opportunity to draw upon different types of research knowledge available in the surrounding research landscape. This finding suggests that UASs have a particularly pronounced role in technology transfer an in the adaptation of basic R&D to practical needs. In light of the growing complexity of technological innovation, such regional knowledge complementarities can be a key competitive advantage for producing innovation (Balland and Rigby, 2017).

More, specifically, we find strong knowledge complementarities between UASs and three other types of research institutions—Max Planck institutes (which perform basic research), Fraunhofer institutes (which perform applied research), and Helmholtz institutes (which perform a mixture of both). Complementarities between UASs and these three types of institutions lead to an additional increase in patent quantity. For patent quality, we find the same overall pattern of the effects. However, stand-alone UASs do not significantly increase patent quality and complementarities arise only between UASs and Max Planck institutes. To increase patent quality, Max Planck institutes thus provide a particularly valuable source of complementary basic research knowledge. This finding suggests that UASs can contribute to more radical innovations in regions where they can draw upon strong basic research knowledge, such as that of Max Planck institutes, and to more incremental innovations in other regions.

This chapter offers a novel solution to estimating the causal innovation effect of UASs, providing evidence for a positive UAS effect and for the existence of regional complementarities between UASs and other types of research institutions. Although our methodological approach using region FE, year FE, and satellite data as a proxy for regional economic activity to account for potential endogeneity in UAS campus locations might have certain limitations, we argue that by using daytime satellite imagery that provides information on six different types of surfaces related to economic activity, the proxy yields at least a more detailed insight into the overall structure of a region than, for example, simple measures of GDP. Combining the proxy for regional economic activity with region FE and year FE provides at least a better solution to the problem of endogenously determined UAS campus openings than was possible with previously available data.

Future research needs to shed light on the exact mechanisms behind the regional complementarities between UASs and other types of research institutions. More specifically, the question of whether knowledge complementarities arise through direct linkages between the different types of research institutions or through indirect linkages remains open. Direct linkages, such as movements of faculty, co-patenting and co-publication, exchanges between researchers at workshops or conferences, or cooperative research projects, can lead to the knowledge complementarities we find in this chapter. In addition, indirect linkages, such as local firms' drawing on the knowledge of regional research institutions by hiring their graduates, can contribute to such complementarities (e.g., chapter 2 and Schultheiss et al., 2021). An analysis of these direct an indirect linkages can also provide insights into the circumstances under which UASs and other research institutions might be substitutes, a relationship that our robustness checks reveal for some innovation tasks. Future research needs to further explore these potential mechanisms to achieve a better understanding of how knowledge complementarities between UASs and other research institutions arise in a diverse research landscape.

4.8 Appendix

4.8.1 Descriptive Statistics for Dependent Variables

Table 4.3: Descriptive statistics for patent quantity ($PQUAN$)

	Treatment group					Control group				
Year	N	Mean	SD	Min.	Max.	N	Mean	SD	Min.	Max.
1980	3,111	1.43	9.24	0.00	375.03	5,494	0.27	1.97	0.00	109.82
1981	3,111	1.51	8.89	0.00	317.33	5,494	0.30	1.90	0.00	97.62
1982	3,111	1.57	9.20	0.00	378.96	5,494	0.30	1.86	0.00	102.65
1983	3,111	1.80	9.87	0.00	393.11	5,494	0.37	2.12	0.00	114.90
1984	3,111	1.97	10.09	0.00	387.92	5,494	0.40	2.06	0.00	109.42
1985	3,111	2.14	11.00	0.00	425.26	5,494	0.42	2.25	0.00	119.70
1986	3,111	2.24	11.01	0.00	416.45	5,494	0.45	2.23	0.00	111.72
1987	3,111	2.50	11.78	0.00	427.02	5,494	0.53	2.61	0.00	129.78
1988	3,111	2.72	12.92	0.00	462.44	5,494	0.57	2.55	0.00	111.46
1989	3,111	2.76	12.88	0.00	475.64	5,494	0.57	2.74	0.00	138.49
1990	3,111	2.59	11.79	0.00	438.52	5,494	0.53	2.31	0.00	101.48
1991	3,720	2.20	11.14	0.00	390.32	7,546	0.40	2.07	0.00	114.99
1992	3,720	2.23	11.43	0.00	435.26	7,546	0.41	2.10	0.00	124.33
1993	3,720	2.26	11.05	0.00	379.63	7,546	0.43	1.98	0.00	104.56
1994	3,720	2.42	11.98	0.00	399.79	7,546	0.45	2.19	0.00	118.22
1995	3,720	2.52	12.67	0.00	465.47	7,546	0.48	2.13	0.00	91.48
1996	3,720	3.00	15.29	0.00	591.50	7,546	0.58	2.54	0.00	119.68
1997	3,720	3.40	17.66	0.00	711.32	7,546	0.64	2.60	0.00	114.72
1998	3,720	3.79	20.65	0.00	894.83	7,546	0.72	2.78	0.00	114.95
1999	3,720	4.12	22.06	0.00	899.20	7,546	0.74	2.92	0.00	131.64
2000	3,720	4.36	24.75	0.00	1,011.43	7,546	0.78	2.98	0.00	106.95
2001	3,720	4.34	24.04	0.00	940.36	7,546	0.76	2.91	0.00	112.59
2002	3,720	4.31	23.46	0.00	842.45	7,546	0.76	2.90	0.00	112.89
2003	3,720	4.35	23.31	0.00	799.75	7,546	0.80	3.08	0.00	122.04
2004	3,720	4.52	23.58	0.00	798.77	7,546	0.84	3.15	0.00	112.81
2005	3,720	4.67	24.59	0.00	805.28	7,546	0.89	3.38	0.00	134.82
2006	3,720	4.65	25.17	0.00	904.50	7,546	0.91	3.30	0.00	113.30
2007	3,720	4.72	25.24	0.00	878.40	7,546	0.91	3.23	0.00	108.88
2008	3,720	4.46	24.85	0.00	863.84	7,546	0.89	3.15	0.00	93.12
2009	3,720	4.53	24.07	0.00	798.90	7,546	0.90	3.27	0.00	98.15
2010	3,720	4.59	24.89	0.00	873.45	7,546	0.88	3.18	0.00	107.56
2011	3,720	4.49	24.65	0.00	839.19	7,546	0.87	3.20	0.00	91.58
2012	3,720	4.26	24.07	0.00	861.12	7,546	0.85	3.26	0.00	113.42
2013	3,720	4.20	24.50	0.00	867.33	7,546	0.84	3.10	0.00	101.44
2014	3,720	4.24	25.50	0.00	939.91	7,546	0.82	3.17	0.00	100.50
2015	3,720	4.20	26.71	0.00	1,041.00	7,546	0.81	3.08	0.00	96.93
Total	127,221	3.40	19.46	0.00	1,041.00	249,084	0.66	2.76	0.00	138.49

Source: Authors' calculations based on patent data from the EPO, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.

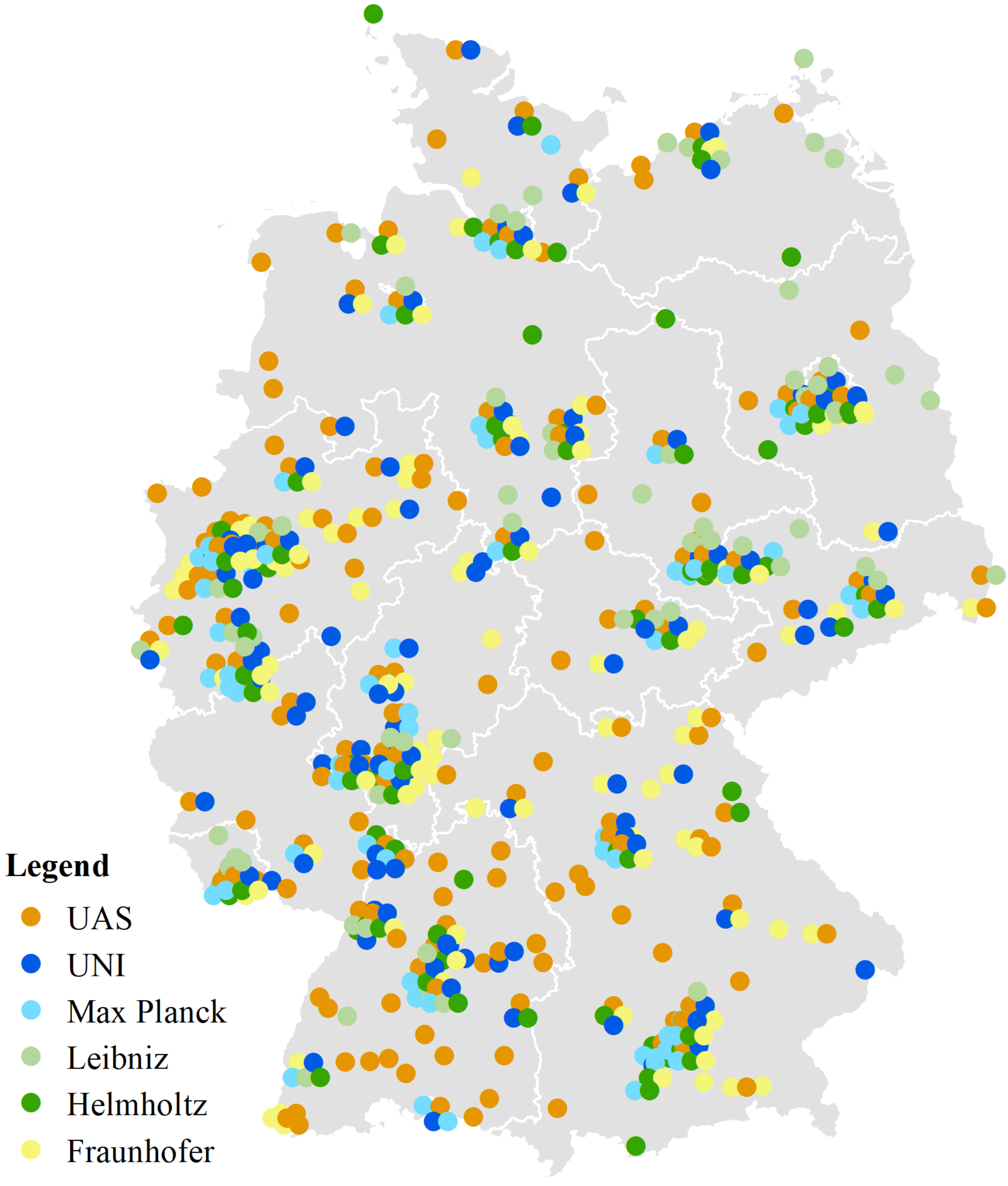
Table 4.4: Descriptive statistics for patent quality (citations) ($PQUAL_{cit}$)

	Treatment group					Control group				
Year	N	Mean	SD	Min.	Max.	N	Mean	SD	Min.	Max.
1980	3,111	0.11	0.33	0.00	5.00	5,494	0.04	0.23	0.00	4.00
1981	3,111	0.13	0.40	0.00	7.00	5,494	0.05	0.27	0.00	9.33
1982	3,111	0.13	0.36	0.00	5.00	5,494	0.05	0.26	0.00	4.67
1983	3,111	0.15	0.39	0.00	5.00	5,494	0.06	0.26	0.00	4.00
1984	3,111	0.18	0.47	0.00	10.00	5,494	0.07	0.31	0.00	6.00
1985	3,111	0.18	0.47	0.00	11.00	5,494	0.08	0.37	0.00	10.00
1986	3,111	0.20	0.44	0.00	4.00	5,494	0.10	0.39	0.00	7.00
1987	3,111	0.22	0.48	0.00	5.07	5,494	0.10	0.38	0.00	7.00
1988	3,111	0.22	0.54	0.00	16.00	5,494	0.10	0.35	0.00	7.00
1989	3,111	0.22	0.46	0.00	6.48	5,494	0.11	0.38	0.00	5.00
1990	3,111	0.23	0.56	0.00	10.15	5,494	0.12	0.43	0.00	12.00
1991	3,720	0.22	0.49	0.00	5.38	7,546	0.09	0.39	0.00	9.00
1992	3,720	0.23	0.54	0.00	10.62	7,546	0.10	0.42	0.00	14.00
1993	3,720	0.24	0.54	0.00	7.02	7,546	0.11	0.51	0.00	22.00
1994	3,720	0.27	0.58	0.00	9.17	7,546	0.11	0.45	0.00	10.00
1995	3,720	0.27	0.55	0.00	8.00	7,546	0.12	0.45	0.00	7.00
1996	3,720	0.31	0.63	0.00	9.00	7,546	0.15	0.50	0.00	9.67
1997	3,720	0.32	0.60	0.00	9.00	7,546	0.15	0.50	0.00	12.00
1998	3,720	0.34	0.65	0.00	12.00	7,546	0.15	0.49	0.00	10.00
1999	3,720	0.34	0.64	0.00	13.00	7,546	0.17	0.58	0.00	20.29
2000	3,720	0.33	0.83	0.00	37.00	7,546	0.17	0.53	0.00	11.27
2001	3,720	0.36	0.85	0.00	28.00	7,546	0.17	0.50	0.00	10.00
2002	3,720	0.36	0.71	0.00	12.00	7,546	0.18	0.54	0.00	7.00
2003	3,720	0.38	0.69	0.00	13.20	7,546	0.19	0.56	0.00	11.00
2004	3,720	0.35	0.62	0.00	9.00	7,546	0.20	0.62	0.00	17.00
2005	3,720	0.36	0.64	0.00	9.12	7,546	0.19	0.52	0.00	8.00
2006	3,720	0.34	0.59	0.00	7.00	7,546	0.19	0.51	0.00	8.67
2007	3,720	0.35	0.64	0.00	14.00	7,546	0.20	0.61	0.00	20.00
2008	3,720	0.38	0.69	0.00	11.00	7,546	0.22	0.65	0.00	18.00
2009	3,720	0.38	0.72	0.00	11.00	7,546	0.23	0.75	0.00	26.00
2010	3,720	0.36	0.72	0.00	19.00	7,546	0.19	0.72	0.00	33.00
2011	3,720	0.32	0.74	0.00	29.00	7,546	0.18	0.55	0.00	15.00
2012	3,720	0.29	0.55	0.00	11.00	7,546	0.17	0.54	0.00	14.00
2013	3,720	0.29	0.54	0.00	8.00	7,546	0.16	0.59	0.00	31.00
2014	3,720	0.27	0.56	0.00	11.00	7,546	0.15	0.47	0.00	13.00
2015	3,720	0.19	0.41	0.00	7.14	7,546	0.11	0.37	0.00	7.00
Total	127,221	0.28	0.60	0.00	37.00	249,084	0.14	0.50	0.00	33.00

Source: Authors' calculations based on patent data from the EPO, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.

4.8.2 Distribution of STEM Campus and STEM Institute Locations

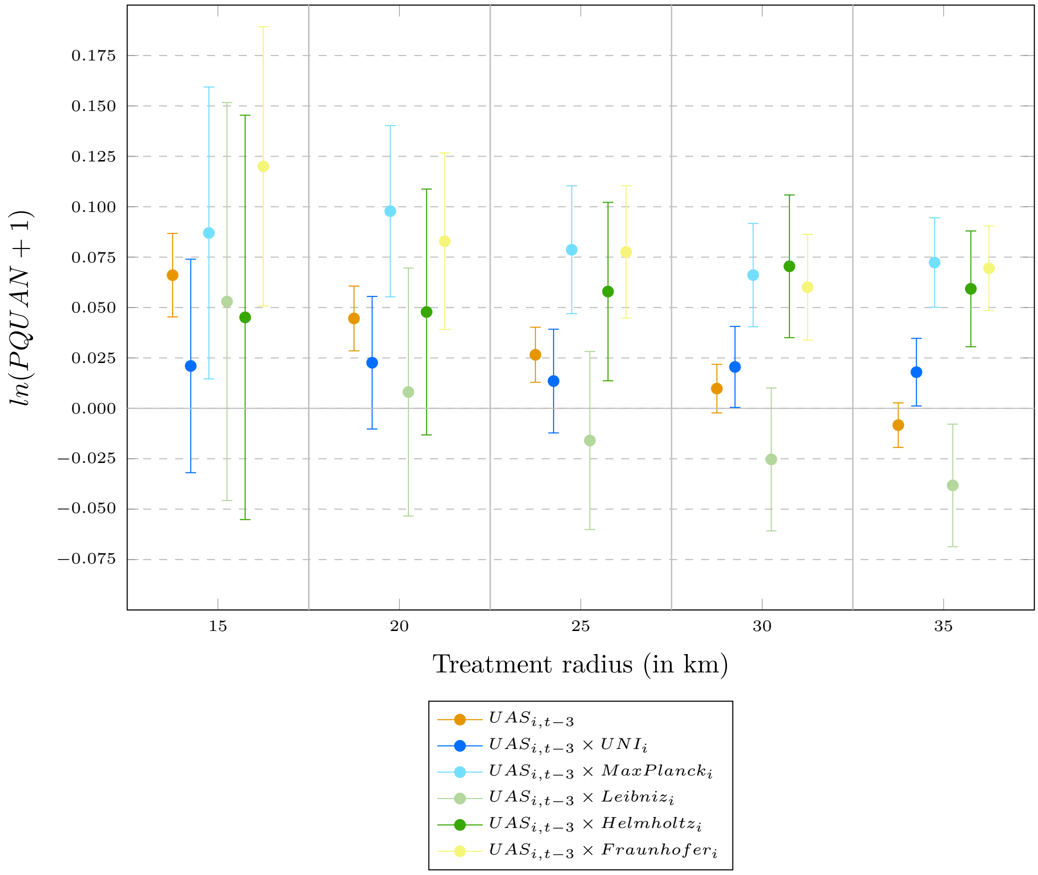
Figure 4.3: Distribution of STEM campus and STEM institute locations in 2015



Source: Authors' illustration with geodata from the BKG, data on HEIs from the German Rectors' Conference, data on higher education institutions from BMBF (2018b), and self-collected data on HEIs and PROs.

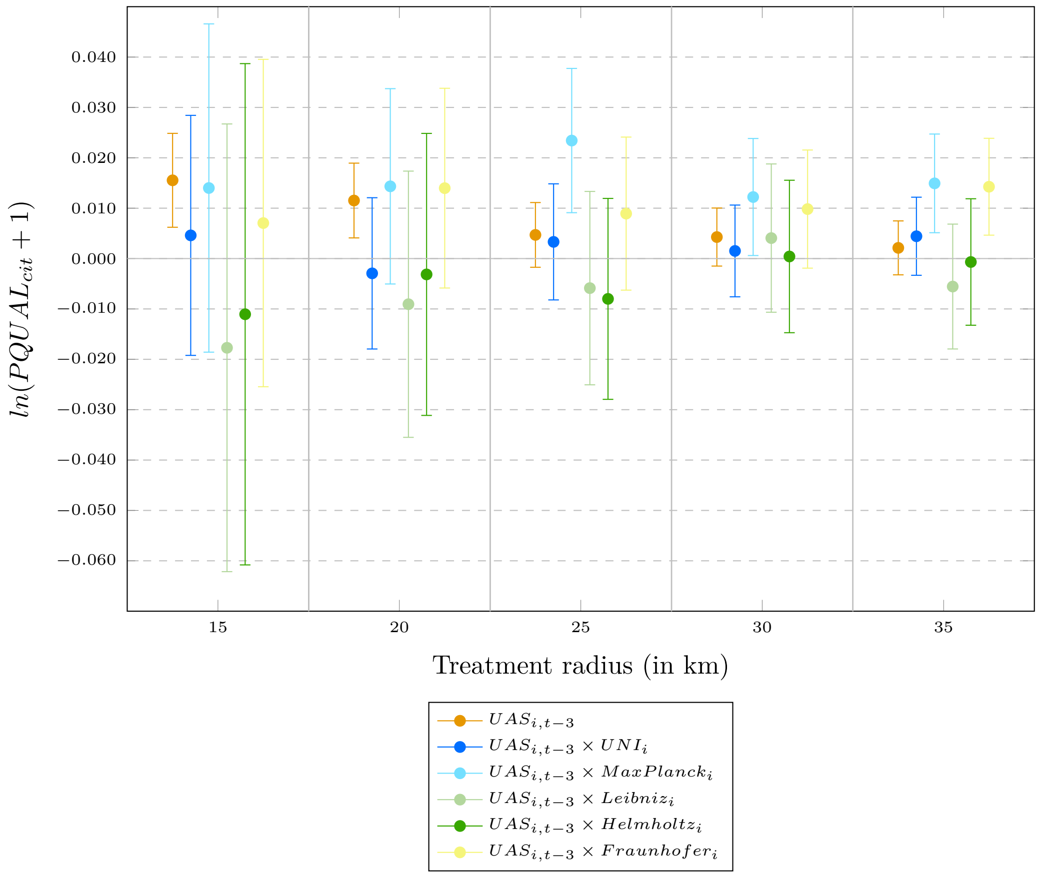
4.8.3 FE Estimation Results on Patent Quantity and Patent Quality (Citations) with Varying Treatment Radii

Figure 4.4: FE estimation results on patent quantity with varying treatment radii



Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Figure plots coefficients and their 95 percent confidence intervals from separate FE estimations for each treatment radius. With the exception of the varying treatment radii, the estimations are similar to those in column 7 of Table 4.1.

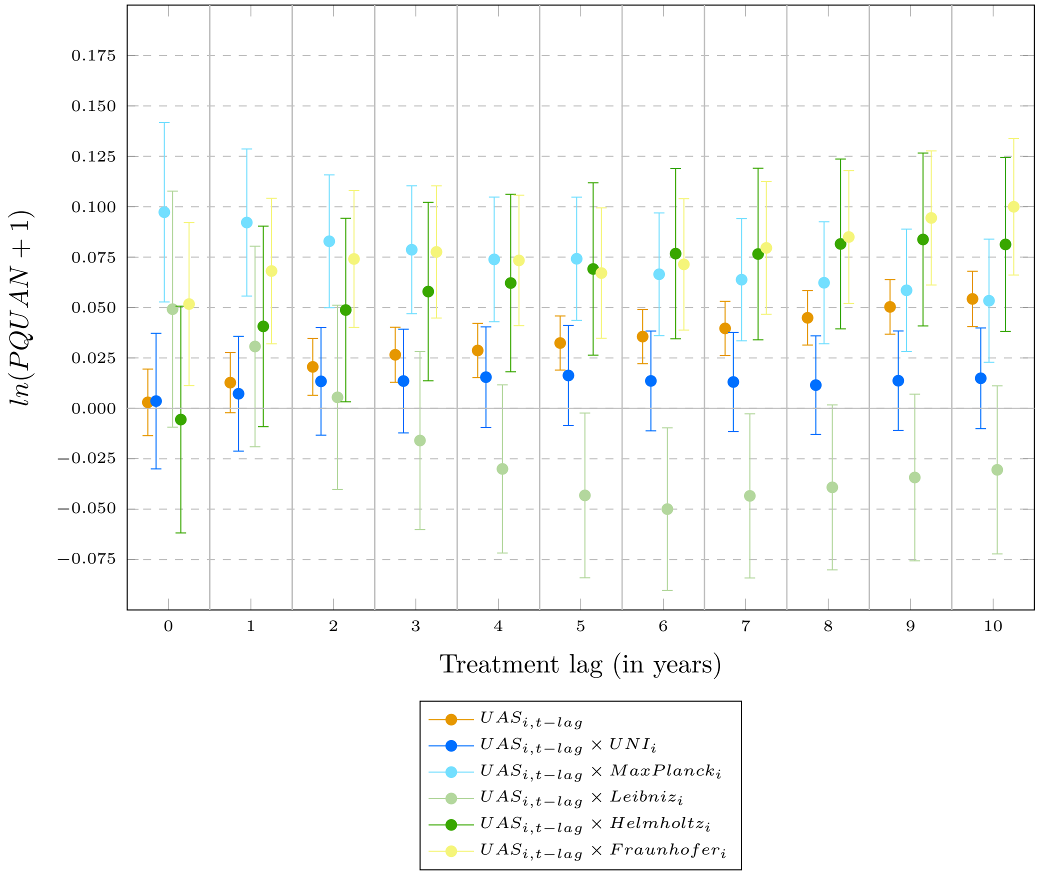
Figure 4.5: FE estimation results on patent quality (citations) with varying treatment radii



Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Figure plots coefficients and their 95 percent confidence intervals from separate FE estimations for each treatment radius. With the exception of the varying treatment radii, the estimations are similar to those in column 7 of Table 4.2.

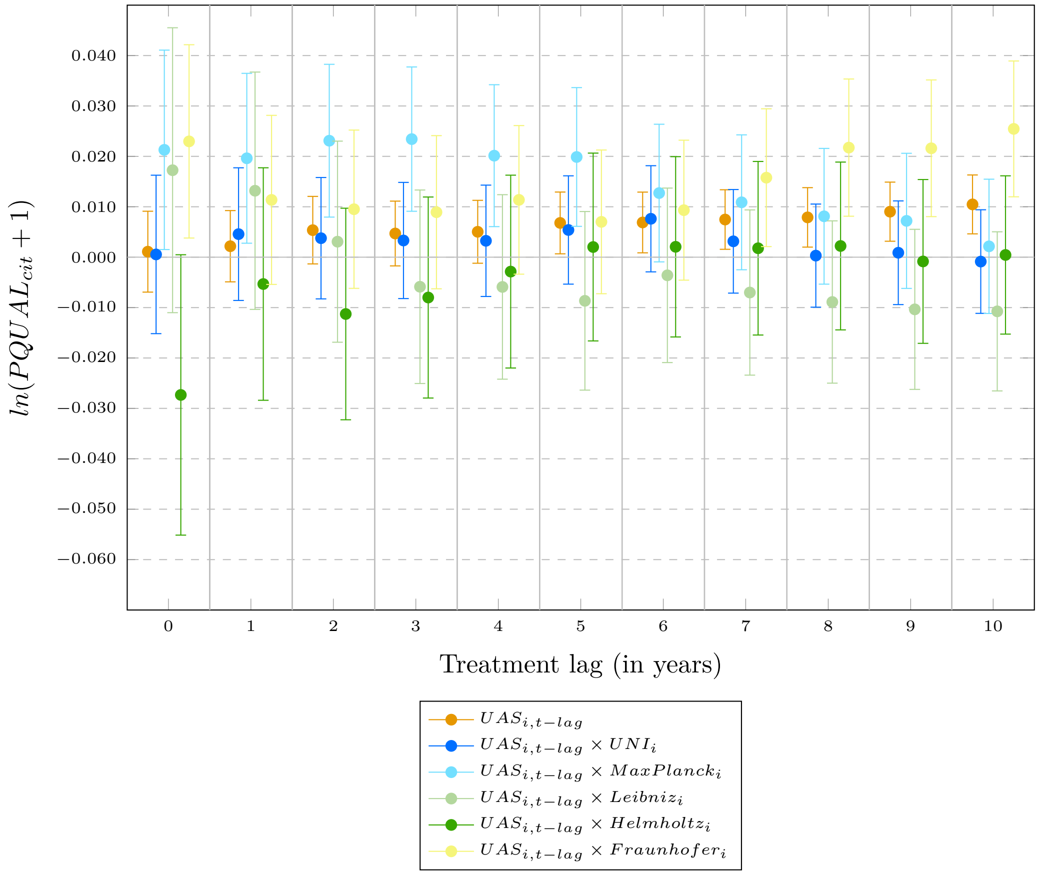
4.8.4 FE Estimation Results on Patent Quantity and Patent Quality (Citations) with Varying Treatment Lags

Figure 4.6: FE estimation results on patent quantity with varying treatment lags



Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Figure plots coefficients and their 95 percent confidence intervals from separate FE estimations for each treatment radius. With the exception of the varying treatment lag, the estimations are similar to those in column 7 of Table 4.1.

Figure 4.7: FE estimation results on patent quality (citations) with varying treatment lags



Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.

Notes: Figure plots coefficients and their 95 percent confidence intervals from separate FE estimations for each treatment radius. With the exception of the varying treatment lags, the estimations are similar to those in column 7 of Table 4.2.

4.8.5 FE Estimation Results on Patent Quantity and Patent Quality (Citations) for West German Municipalities Only

Table 4.5: FE estimation results on patent quantity (25-km travel-distance radius, 3-y. lag, West German municipalities only)

Dep. var.: $ln(PQUAN+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.060***	0.038***	0.042***	0.059***	0.055***	0.042***	0.028***
	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)	(0.008)
$UAS_{i,t-3} \times UNI_i$		0.066***					0.024*
		(0.013)					(0.014)
$UAS_{i,t-3} \times MaxPlanck_i$			0.095***				0.074***
			(0.015)				(0.017)
$UAS_{i,t-3} \times Leibniz_i$				0.015			-0.042*
				(0.023)			(0.024)
$UAS_{i,t-3} \times Helmholtz_i$					0.073***		0.015
					(0.024)		(0.025)
$UAS_{i,t-3} \times Fraunhofer_i$						0.099***	0.068***
						(0.016)	(0.018)
Municipalities	8,605	8,605	8,605	8,605	8,605	8,605	8,605
N	275,360	275,360	275,360	275,360	275,360	275,360	275,360
R^2 (within)	0.102	0.103	0.103	0.102	0.102	0.103	0.103
R^2 (between)	0.277	0.297	0.309	0.278	0.287	0.302	0.326
R^2 (overall)	0.157	0.165	0.177	0.158	0.163	0.169	0.183

Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, surface groups, and controls for $UNI_{i,t-3}$ $MaxPlanck_{i,t-3}$ $Leibniz_{i,t-3}$ $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$ $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.6: FE estimation results on patent quality (citations) (25-km travel-distance radius, 3-y. lag, West German municipalities only)

Dep. var.: $ln(PQUAL_{cit}+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.010***	0.005	0.006*	0.010***	0.011***	0.008**	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
$UAS_{i,t-3} \times UNI_i$		0.013**					0.008
		(0.006)					(0.006)
$UAS_{i,t-3} \times MaxPlanck_i$			0.022***				0.024***
			(0.007)				(0.008)
$UAS_{i,t-3} \times Leibniz_i$				-0.006			-0.016
				(0.010)			(0.010)
$UAS_{i,t-3} \times Helmholtz_i$					-0.012		-0.025**
					(0.010)		(0.011)
$UAS_{i,t-3} \times Fraunhofer_i$						0.013*	0.008
						(0.007)	(0.008)
Municipalities	8,605	8,605	8,605	8,605	8,605	8,605	8,605
N	275,360	275,360	275,360	275,360	275,360	275,360	275,360
R^2 (within)	0.026	0.026	0.026	0.026	0.026	0.026	0.026
R^2 (between)	0.147	0.169	0.182	0.146	0.137	0.160	0.176
R^2 (overall)	0.048	0.050	0.053	0.047	0.046	0.049	0.052

Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, surface groups, and controls for $UNI_{i,t-3}$, $MaxPlanck_{i,t-3}$, $Leibniz_{i,t-3}$, $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$ $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.8.6 FE Estimation Results on Patent Quantity and Patent Quality (Citations) without Surface Groups

Table 4.7: FE estimation results on patent quantity (25-km travel-distance radius, 3-y. lag, without surface groups)

Dep. var.: $ln(PQUAN+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.111***	0.079***	0.080***	0.107***	0.099***	0.081***	0.058***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
$UAS_{i,t-3} \times UNI_i$		0.095***					0.024
		(0.013)					(0.015)
$UAS_{i,t-3} \times MaxPlanck_i$			0.150***				0.111***
			(0.016)				(0.017)
$UAS_{i,t-3} \times Leibniz_i$				0.051**			-0.049**
				(0.023)			(0.023)
$UAS_{i,t-3} \times Helmholtz_i$					0.151***		0.064***
					(0.024)		(0.025)
$UAS_{i,t-3} \times Fraunhofer_i$						0.165***	0.112***
						(0.017)	(0.019)
Municipalities	11,266	11,266	11,266	11,266	11,266	11,266	11,266
N	376,305	376,305	376,305	376,305	376,305	376,305	376,305
R^2 (within)	0.123	0.124	0.125	0.123	0.124	0.125	0.126
R^2 (between)	0.034	0.045	0.075	0.038	0.046	0.063	0.090
R^2 (overall)	0.053	0.058	0.072	0.054	0.059	0.066	0.079

Source: Authors' calculations based on patent data from the EPO, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.

Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, and controls for $UNI_{i,t-3}$ $MaxPlanck_{i,t-3}$ $Leibniz_{i,t-3}$ $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$ $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.8: FE estimation results on patent quality (citations) (25-km travel-distance radius, 3-y. lag, without surface groups)

Dep. var.: $ln(PQUAL_{cit}+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.023***	0.017***	0.016***	0.023***	0.022***	0.019***	0.014***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$UAS_{i,t-3} \times UNI_i$		0.017***					0.004
		(0.005)					(0.005)
$UAS_{i,t-3} \times MaxPlanck_i$			0.036***				0.035***
			(0.006)				(0.006)
$UAS_{i,t-3} \times Leibniz_i$				0.004			-0.017**
				(0.008)			(0.008)
$UAS_{i,t-3} \times Helmholtz_i$					0.015*		-0.003
					(0.008)		(0.009)
$UAS_{i,t-3} \times Fraunhofer_i$						0.025***	0.015**
						(0.006)	(0.007)
Municipalities	11,266	11,266	11,266	11,266	11,266	11,266	11,266
N	376,305	376,305	376,305	376,305	376,305	376,305	376,305
R^2 (within)	0.034	0.035	0.035	0.034	0.035	0.035	0.035
R^2 (between)	0.003	0.001	0.002	0.003	0.002	0.000	0.004
R^2 (overall)	0.022	0.023	0.027	0.022	0.023	0.024	0.028

Source: Authors' calculations based on patent data from the EPO, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, and controls for $UNI_{i,t-3}$ $MaxPlanck_{i,t-3}$ $Leibniz_{i,t-3}$ $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$ $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.8.7 FE Estimation Results on Additional Patent Quality Measures

Table 4.9: FE estimation results on patent quality (priority claims) (25-km travel-distance radius, 3-y. lag)

Dep. var.: $ln(PQUAL_{claims}+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.010***	0.016***	0.014***	0.014***	0.011***	0.012***	0.019***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)
$UAS_{i,t-3} \times UNI_i$		-0.019***					-0.015**
		(0.006)					(0.007)
$UAS_{i,t-3} \times MaxPlanck_i$			-0.023***				-0.005
			(0.008)				(0.009)
$UAS_{i,t-3} \times Leibniz_i$				-0.055***			-0.049***
				(0.010)			(0.011)
$UAS_{i,t-3} \times Helmholtz_i$					-0.025**		-0.007
					(0.010)		(0.011)
$UAS_{i,t-3} \times Fraunhofer_i$						-0.012	0.004
						(0.008)	(0.009)
Municipalities	11,266	11,266	11,266	11,266	11,266	11,266	11,266
N	341,885	341,885	341,885	341,885	341,885	341,885	341,885
R^2 (within)	0.036	0.036	0.036	0.036	0.036	0.036	0.036
R^2 (between)	0.116	0.107	0.098	0.108	0.108	0.110	0.099
R^2 (overall)	0.067	0.065	0.062	0.065	0.065	0.066	0.063

Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, surface groups, and controls for $UNI_{i,t-3}$ $MaxPlanck_{i,t-3}$ $Leibniz_{i,t-3}$ $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$ $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.10: FE estimation results on patent quality (family size) (25-km travel-distance radius, 3-y. lag)

Dep. var.: $ln(PQUAL_{size}+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.035***	0.036***	0.034***	0.039***	0.035***	0.030***	0.035***
	(0.008)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)
$UAS_{i,t-3} \times UNI_i$		-0.002					-0.016
		(0.015)					(0.017)
$UAS_{i,t-3} \times MaxPlanck_i$			0.007				0.018
			(0.018)				(0.021)
$UAS_{i,t-3} \times Leibniz_i$				-0.054**			-0.071***
				(0.024)			(0.027)
$UAS_{i,t-3} \times Helmholtz_i$					0.005		0.005
					(0.022)		(0.025)
$UAS_{i,t-3} \times Fraunhofer_i$						0.030	0.040*
						(0.018)	(0.021)
Municipalities	11,266	11,266	11,266	11,266	11,266	11,266	11,266
N	341,885	341,885	341,885	341,885	341,885	341,885	341,885
R^2 (within)	0.064	0.064	0.064	0.064	0.064	0.064	0.064
R^2 (between)	0.157	0.157	0.159	0.155	0.158	0.161	0.163
R^2 (overall)	0.102	0.102	0.103	0.102	0.103	0.104	0.105

Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, surface groups, and controls for $UNI_{i,t-3}$, $MaxPlanck_{i,t-3}$, $Leibniz_{i,t-3}$, $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.8.8 FE Estimation Results for Pre-Bologna Period

Table 4.11: FE estimation results on patent quantity for pre-2005 period (25-km travel-distance radius, 3-y. lag)

Dep. var.: $ln(PQUAN+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.069***	0.045***	0.046***	0.065***	0.060***	0.046***	0.030***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
$UAS_{i,t-3} \times UNI_i$		0.070***					0.015
		(0.012)					(0.012)
$UAS_{i,t-3} \times MaxPlanck_i$			0.113***				0.080***
			(0.014)				(0.015)
$UAS_{i,t-3} \times Leibniz_i$				0.049**			-0.031
				(0.021)			(0.022)
$UAS_{i,t-3} \times Helmholtz_i$					0.118***		0.047**
					(0.021)		(0.022)
$UAS_{i,t-3} \times Fraunhofer_i$						0.130***	0.092***
						(0.015)	(0.017)
Municipalities	11,266	11,266	11,266	11,266	11,266	11,266	11,266
N	217,959	217,959	217,959	341,885	217,959	217,959	217,959
R^2 (within)	0.091	0.092	0.092	0.091	0.092	0.093	0.093
R^2 (between)	0.138	0.153	0.186	0.141	0.143	0.175	0.202
R^2 (overall)	0.111	0.118	0.137	0.113	0.116	0.132	0.147

Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, surface groups, and controls for $UNI_{i,t-3}$ $MaxPlanck_{i,t-3}$ $Leibniz_{i,t-3}$ $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$ $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4.12: FE estimation results on patent quality (citations) for pre-2005 period (25-km travel-distance radius, 3-y. lag)

Dep. var.: $ln(PQUAL_{cit}+1)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$UAS_{i,t-3}$	0.018***	0.013***	0.012***	0.018***	0.018***	0.016***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
$UAS_{i,t-3} \times UNI_i$		0.016***					0.007
		(0.006)					(0.007)
$UAS_{i,t-3} \times MaxPlanck_i$			0.032***				0.032***
			(0.007)				(0.008)
$UAS_{i,t-3} \times Leibniz_i$				0.005			-0.012
				(0.009)			(0.010)
$UAS_{i,t-3} \times Helmholtz_i$					0.007		-0.008
					(0.010)		(0.011)
$UAS_{i,t-3} \times Fraunhofer_i$						0.015**	0.006
						(0.008)	(0.009)
Municipalities	11,266	11,266	11,266	11,266	11,266	11,266	11,266
N	217,959	217,959	217,959	341,885	217,959	217,959	217,959
R^2 (within)	0.026	0.026	0.026	0.026	0.026	0.026	0.026
R^2 (between)	0.083	0.092	0.115	0.083	0.084	0.093	0.121
R^2 (overall)	0.046	0.048	0.053	0.046	0.046	0.048	0.055

Source: Authors' calculations based on patent data from the EPO, data on surface groups from Chapter 3, data on HEIs from the German Rectors' Conference, data on PROs from BMBF (2018b), and self-collected data on HEIs and PROs.
Notes: Robust standard errors in parentheses. All models include intercept, municipality FE, year FE, surface groups, and controls for $UNI_{i,t-3}$ $MaxPlanck_{i,t-3}$ $Leibniz_{i,t-3}$ $Helmholtz_{i,t-3}$ and $Fraunhofer_{i,t-3}$ $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Chapter 5

Conclusion

Although many European countries have introduced UASs—HEIs that focus on vocational knowledge and applied research—since the 1960s (Lepori and Kyvik, 2010), these UASs remain widely unexplored in the literature investigating the innovation effects of HEIs. In their focus on vocational knowledge and applied research, UASs differ from UNIs, which focus on academic knowledge and basic research. Therefore, the question arises as to how UASs and UNIs differ in their direct and indirect effects on innovation. In theory, through their applied research activities, UASs might contribute directly to regional innovation by transferring basic research knowledge to technological applications. Moreover, UASs might indirectly influence regional innovation by producing human capital that can play a key role in firms' innovation activities. Consequently, UASs potentially differ from UNIs in their impacts on regional economies.

In this dissertation, I investigate the impact of UASs in Switzerland and Germany on regional innovation. Despite institutional differences in, for example, admission requirements for studying at a UAS (Lackner, 2019; Nikolai and Ebner, 2013; SCCRE, 2018), these two countries maintained the distinct profiles of UASs and UNIs when they orginally introduced UASs to their higher education systems (Lepori and Kyvik, 2010). I analyze two potential aspects of the UAS effect on regional innovation. First, I examine whether firms employ the human capital that UASs produce for innovation activities (chapter 2). Second, I study the role of regional research infrastructures for the UAS effect on innovation (chapters 3 and 4). Through these analyses, I extend Pfister et al.'s (2021) pioneering work on the UAS innovation effect. They find a significant increase in regional patenting activities after the introduction of UASs in Switzerland. More generally, I contribute to the literature on how HEIs affect regional innovation by investigating UASs, a type of HEI that remains widely under-researched despite its central role in the education and innovation systems of many European countries.

In Chapter 2, to examine the first aspect of the UAS effect on innovation—whether firms use the skills of UAS graduates for innovation activities—I study the effects of Swiss UASs on the employment of R&D personnel in local firms. Following Pfister et al. (2021), I exploit quasi-random variation in the location and staggered timing of UAS campus openings to identify the causal effect of UASs on R&D employment at the establishment level. I find that treated establishments (i.e., those within 25 kilometers of a UAS campus) employ more personnel for R&D tasks. Moreover, these firms pay a larger percentage of their total wage sum to R&D employees. Analyses of effect heterogeneity suggest that small firms, including potential start-ups, particularly profit from the introduction of UASs.

The analysis in Chapter 2 shows that firms' intensifying their R&D efforts by employing more personnel for R&D constitutes one important mechanism underlying the UAS innovation effect. UASs thus provide local firms with a type of human capital that fosters innovation activities in firms. The increasing wage sum paid to R&D employees implies that firms value the skills of UAS graduates for R&D tasks. The introduction of UASs in Switzerland thus enhanced regional skill resources by combining vocational and applied research knowledge. For local firms, these newly available skills constitute a previously missing resource crucial for promoting innovation activities. By providing local labor markets with graduates who possess these skills, UASs indirectly contribute to regional economies through their human capital formation.

In Chapters 3 and 4, I examine the second aspect of the UAS effect on innovation—knowledge complementarities between UASs and other research institutions coexisting within regional research infrastructures. I analyze whether the innovation effect of opening a

UAS campus in German regions differs between regions where no other research institution existed before the UAS opening and regions with preexisting research institutions. In Germany, a variety of other public research institutions potentially interact with UASs, including UNIs and PROs. These institutions cover wide ranges of research activities along the entire spectrum of basic to applied research. To measure innovation, I follow Pfister et al. (2021) by examining the development of patenting activities.

Unfortunately, unlike in Switzerland, the spatial and temporal variation in the locations of UAS campuses is not quasi-random in Germany. Consequently, I cannot identify the causal effects of the introduction of UASs in Germany without accounting in my empirical framework for factors determining this endogenous variation. However, German data on such factors does not exist for meeting two necessary data property conditions: a time series dating back to 1985 (when German UASs began to conduct applied research) and sufficiently small regional units. This absence of data presents a challenge for empirically identifying the innovation effects of German UASs.

Therefore, in Chapter 3, I develop a proxy measure for regional economic activity—a factor that determines (at least a part of) the spatial and temporal variation of German UAS campuses—from a long time series of remotely-sensed daytime satellite imagery. From 1984 through 2020, this measure contains land cover (i.e., the observable terrestrial features of the earth's surface) in six categories that I call "surface groups" (e.g., built-up surfaces, cropland). In computing the measure, I apply machine-learning techniques from the remote-sensing literature that uses daytime satellite imagery to classify land cover (e.g., Dewan and Yamaguchi, 2009; Goldblatt et al., 2016; Liu et al., 2018). Using regionally disaggregated external validation data available for a limited time series (part of the daytime satellite imagery's observation period), I show that the surface groups explain a significant percentage of the variation in economic activity. Therefore, the surface groups constitute a suitable proxy for regional economic factors determining the openings of UAS campuses.

In addition to the application for identifying the causal effects of opening UAS campuses in Germany, the proxy measure that I develop in Chapter 3 shows further potential for economic research. In Germany, the measure indicates economic activity in East German regions (those in the former German Democratic Republic), for which no reliable data on regional economic activity exists before reunification. Thus the measure can provide useful information for studies comparing the development of East and West German regions. Such studies could exploit the measure for matching East and West German regions with similar economic developments before the reunification, thereby making possible the identification of causal effects. Consequently, the measure can help understand why, for example, differences in technological performance and innovation between East and West German regions with otherwise similar economic preconditions still persist 30 years after the reunification (EFI, 2020). Furthermore, Chapter 3 makes a more general methodological contribution to the literature evaluating the use of satellite-based measures for economic research by showing that daytime satellite imagery constitutes a valuable data source for analyzing long time series and small regional units.

Incorporating the measure developed in Chapter 3 into the analysis of UASs in Chapter 4, I find that strong knowledge complementarities arise between UASs and other research institutions. While UASs increase innovation in regions with no other research institutions, this effect is substantially larger in regions with coexisting research institutions, whether basic or applied. Knowledge complementarities between UASs and other research institutions thus further enhance regional innovation.

The analysis in Chapter 4 shows that knowledge complementarities between UASs and other research institutions constitute a second aspect of the UAS effect on innovation. In light of the findings of Chapter 2—that firms value the type of human capital that UASs produce for R&D activities—specific types of human capital formation evolve in regions with UASs and other research institutions (i.e., human capital that combines applied research skills from UASs and basic research skills from, e.g., Max Planck institutes). Such human capital formation is thus one mechanism through which knowledge complementarities indirectly affect regional innovation in local firms. In addition, Schultheiss et al. (2021) argue that workers with the new type of human capital acquired at UASs also positively affect the innovation activities of workers who completed a dual apprenticeship program (i.e., workers with vocational skills but largely without research skills), thereby enhancing regional innovation. Schlegel et al.'s (2022) finding that regional labor market strength is an important precondition for UASs increasing innovation further supports the argument that the regional combination of different types of human capital formation constitutes a crucial mechanism underlying the UAS effect on innovation.

In addition to human capital formation, further direct (e.g., cooperative research projects) and indirect (e.g., firms' drawing on different types of regionally available research knowledge) linkages between UASs and other research institutions likely lead to knowledge complementarities. The results in Chapter 4 show that the regional research infrastructure surrounding new UASs contributes significantly to the UAS effect on innovation. The availability of different types of research knowledge for UASs to draw upon can thus play a crucial role for UASs driving regional innovation.

In essence, this dissertation shows that the type of knowledge that UASs produce contributes significantly to regional economies by promoting innovation. By studying UASs in Switzerland and Germany, I shed light on two aspects of the UAS effect on innovation. First, one determinant of the success of Swiss UASs as a driver of innovation is that the combination of vocational and applied research skills that UASs impart to their students is a key input factor for local firms' R&D activities. This finding supports Backes-Gellner and Pfister's (2019) and Schultheiss et al.'s (2021) argument that the skills of UAS graduates complement the innovation activities of local firms rather than replacing other skills. Second, the availability of complementary research knowledge within a region boosts the innovation effect of German UASs. This finding implies that UASs complement other research institutions rather than replacing them for example, by helping translate basic research results into applicable technologies. Therefore, UASs can contribute to regional economies by providing research knowledge complementary to the research knowledge that other actors in regional economies produce.

This dissertation analyzes UASs in two countries, Switzerland and Germany. Although UASs share their focus on vocational knowledge and applied research in both countries, institutional differences between the two countries' education and innovation systems limit the direct transferability of my findings from one country to the other. For example, in Switzerland students need to have successfully completed a dual VET program (in combination with a professional baccalaureate) for admission to a UAS, whereas in Germany students with a general baccalaureate can study at a UAS without further admission requirements. Moreover, particularly in Germany, UASs and UNIs have become more similar over time, including in the focus of their teaching (Enders, 2010). Consequently, after graduation, Swiss UAS students possess deeper practical and vocational knowledge than German UAS students, a combination that could be the key factor for the value of Swiss UAS students to local firms. If so, the findings for Switzerland in Chapter 2 might not hold true for Germany (or other countries), or hold only to a lesser extent. However, independent of country-specific institutional settings, the findings of this dissertation show that the availability of applied research knowledge within regional research infrastructures can be a valuable resource for countries with a strong VET sector or with intentions to strengthen that sector.

Furthermore, Switzerland and Germany differ in their regional research infrastructures. The German innovation system comprises the large sector of PROs, which provide complementary knowledge to German UASs (chapter 4). Comparable institutions do not exist in Switzerland on a nationwide scale. Other institutions, such as the Swiss Federal Institutes of Technology, the UNIs, and other research institutions, might provide complementary knowledge to UASs in Switzerland. Although the findings for German UASs in Chapter 4 are generalizable in the sense that regional research infrastructures matter for the innovation effect of UASs, future research needs to determine precisely which institutions might provide complementary research knowledge to UASs in Switzerland (or other countries).

The findings of this dissertation also contribute to a current political debate on the role of UASs in both Switzerland's and Germany's education systems. Different opinions exist about the binary divide, the clear distinction between teaching-intensive UASs that focus on vocational knowledge and applied research and research-intensive UNIs that focus on academic knowledge and basic research (e.g., Imboden, 2018; Kreckel, 2011; Ziegele et al., 2019). One controversial issue in this debate is whether UASs should be awarded the right to grant doctoral degrees (Imboden, 2018; Meurer, 2018). In this dissertation, I study the innovation effects of UASs from a longitudinal perspective. Therefore, substantial parts of the effects I find, particularly the effects of German UASs, originate from a time when the distinction between UASs and UNIs was far more pronounced than it is today. This dissertation thus provides evidence that, at that time, the clearly differentiated UASs positively affected regional innovation. Whether the now less distinct UASs still exert this effect today, or will in the future, remains an open question.

From an international perspective, the analyses for Switzerland and Germany in this dissertation are generalizable in the sense that institutions combining vocational and applied research knowledge—the knowledge combination that UASs emphasize—can be a crucial element of national education and innovation systems. These system can profit from institutions that build bridges between (a) the vocational and academic tiers and (b) different types of public research institutions. The design of such institutions (and thus how closely they resemble the Swiss and German UASs) depends strongly on the characteristics of a country's education and innovation system. More specifically, these institutions must offer an attractive educational pathway that is appropriately embedded in existing pathways and complement existing public research through a distinct research profile.

The analyses of two aspects of the positive UAS effect on regional innovation, including the new approach for identifying this effect, set the stage for a wide range of future research projects. While I can already investigate firms' reaction to UAS openings in terms of R&D employment, one potential direction for future research is to investigate the impacts of UASs from a broader labor market perspective. Such research could, for example, focus on the implications of UASs for individuals with other educational degrees. For R&D, Schultheiss et al. (2021) have found an upgrading of the tasks of workers with VET degrees after the UAS openings. However, other groups of individuals or individuals with other job tasks might also experience changes to their labor market situation. Moreover, the labor market outcomes of UAS graduates themselves, such as employment opportunities or remuneration, need further exploration.

Another potential direction for future research is to analyze the exact types of linkages between UASs and other actors in regional education and innovation systems. Various channels can lead to knowledge complementarities between UASs and other research institutions. For example, UAS researchers could work jointly on projects with researchers from other institutions, UAS researchers and other researchers could engage in a scientific exchange at locally organized workshops or conferences, UAS graduates and UNI graduates could choose to work in the same firm and combine their skills in firm-level innovation activities, and firms could cooperate with researchers from both UASs and other research institutions. A deeper understanding of these or other potential channels leading to knowledge complementarities between UASs and other research institutions would provide useful information for policies aimed at fostering cooperation between research institutions.

Finally, the methodological contribution of this dissertation—the application of *surface groups* derived from daytime satellite imagery to control for endogeneity in the temporal and spatial distribution of UASs—offers great potential for future research. Beyond my application of studying German UASs, the surface groups provide a proxy for pre-reunification economic activity in East Germany and other former Soviet Union bloc states. This proxy thus can facilitate important analyses that previously suffered from the absence of (reliable) economic data. Such analyses include comparisons and policy evaluations of democratic and former communist regions with similar economic developments before the fall of the iron curtain.

Bibliography

Abramovsky, L., Simpson, H., 2011. Geographic proximity and firm–university innovation linkages: Evidence from Great Britain. Journal of Economic Geography 11, 949–977. doi:10.1093/jeg/lbq052.

Adams, J.D., Chiang, E.P., Jensen, J.L., 2003. The influence of federal laboratory R&D on industrial research. The Review of Economics and Statistics 85, 1003–1020. doi:10.1162/003465303772815899.

Adams, J.D., Chiang, E.P., Starkey, K., 2001. Industry-university cooperative research centers. The Journal of Technology Transfer 26, 73–86. doi:10.1023/A:1007836328722.

Andersson, R., Quigley, J.M., Wilhelmsson, M., 2009. Urbanization, productivity, and innovation: Evidence from investment in higher education. Journal of Urban Economics 66, 2–15. doi:10.1016/j.jue.2009.02.004.

Andrews, M., 2019. How do institutions of higher education affect local invention? Evidence from the establishment of U.S. colleges.

Andrews, M., 2020. How do institutions of higher education affect local invention? Evidence from the establishment of U.S. colleges. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3072565.

Anselin, L., Varga, A., Acs, Z., 1997. Local geographic spillovers between university research and high technology innovations. Journal of Urban Economics 42, 422–448. doi:10.1006/juec.1997.2032.

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. Statistics Surveys 4, 40–79. doi:10.1214/09-SS054.

Arvanitis, S., Kubli, U., Woerter, M., 2008. University-industry knowledge and technology transfer in Switzerland: What university scientists think about co-operation with private enterprises. Research Policy 37, 1865–1883. doi:10.1016/j.respol.2008.07.005.

Athey, S., Imbens, G.W., 2019. Machine learning methods that economists should know about. Annual Review of Economics 11, 685–725. doi:10.1146/annurev-economics-080217-053433.

Audretsch, D.B., Feldman, M.P., 1996. R&D spillovers and the geography of innovation and production. American Economic Review 86, 630–640.

Autant-Bernard, C., 2001. Science and knowledge flows: Evidence from the French case. Research Policy 30, 1069–1078. doi:10.1016/S0048-7333(00)00131-1.

Autor, D., 2003. Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. Journal of Labor Economics 21, 1–42. doi:10.1086/344122.

Backes-Gellner, U., Pfister, C., 2019. The contribution of vocational education and training to innovation: The case of Switzerland. Study elaborated as part of the report "Research and innovation in Switzerland 2020" Part C, Study 1. SERI. Bern.

Backes-Gellner, U., Rupietta, C., Tuor Sartore, S.N., 2017. Reverse educational spillovers at the firm level. Evidence-based HRM: a Global Forum for Empirical Scholarship 5, 80–106. doi:10.1108/EBHRM-03-2015-0007.

Balland, P.A., Rigby, D., 2017. The geography of complex knowledge. Economic Geography 93, 1–23. doi:10.1080/00130095.2016.1205947.

Balzter, H., Cole, B., Thiel, C., Schmullius, C., 2015. Mapping CORINE land cover from Sentinel-1A SAR and SRTM digital elevation model data using random forests. Remote Sensing 7, 14876–14898. doi:10.3390/rs71114876.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data 5, 180214. doi:10.1038/sdata.2018.214.

Becker, W., Dietz, J., 2004. R&D cooperation and innovation activities of firms: Evidence for the German manufacturing industry. Research Policy 33, 209–223. doi:10.1016/j.respol.2003.07.003.

Beise, M., Stahl, H., 1999. Public research and industrial innovations in Germany. Research Policy 28, 397–422. doi:10.1016/S0048-7333(98)00126-7.

Bentley, P.J., Gulbrandsen, M., Kyvik, S., 2015. The relationship between basic and applied research in universities. Higher Education 70, 689–709. doi:10.1007/s10734-015-9861-2.

Bianchi, N., Giorcelli, M., 2020. Scientific education and innovation: From technical diplomas to university STEM degrees. Journal of the European Economic Association 18, 2608–2646. doi:10.1093/jeea/jvz049.

Bindschadler, R., 2003. Landsat coverage of the earth at high latitudes. Photogrammetric Engineering & Remote Sensing 69, 1333–1339.

Bolli, T., Rageth, L., Renold, U., 2018. Der soziale Status der Berufsbildung in der Schweiz. KOF Studien Nr. 101. ETH Zurich – KOF Konjunkturforschungsstelle. Zurich. doi:10.3929/ethz-b-000266722.

Boresjö Bronge, L., Näslund-Landenmark, B., 2002. Wetland classification for Swedish CORINE Land Cover adopting a semi-automatic interactive approach. Canadian Journal of Remote Sensing 28, 139–155. doi:10.5589/m02-011.

Borusyak, K., Jaravel, X., 2017. Revisiting event study designs, with an application to the estimation of the marginal propensity to consume. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2826228.

Breidenbach, P., Eilers, L., 2018. RWI-GEO-GRID: Socio-economic data on grid level. Journal of Economics and Statistics 238, 609–616. doi:10.1515/jbnst-2017-0171.

Brill, A., 2017. Von der "Blauen Liste" zur gesamtdeutschen Wissenschaftsorganisation. Dokserver des Zentrums für Zeithistorische Forschung Potsdam, Potsdam. doi:10.14765/zzf.dok.1.1115.

Broekel, T., Graf, H., 2012. Public research intensity and the structure of German R&D networks: A comparison of 10 technologies. Economics of Innovation and New Technology 21, 345–372. doi:10.1080/10438599.2011.582704.

Brunow, S., Birkeneder, A., Rodríguez-Pose, A., 2018. Creative and science oriented employees and firm-level innovation. Cities 78, 27–38. doi:10.1016/j.cities.2018.02.002.

Bundesgesetz über die Fachhochschulen (Fachhochschulgesetz), 1995. SR/RS 414.71 (2011).

Burchfield, M., Overman, H.G., Puga, D., Turner, M.A., 2006. Causes of sprawl: A portrait from space. The Quarterly Journal of Economics 121, 587–633. doi:10.1162/qjec.2006.121.2.587.

Buschle, N., Haider, C., 2016. Private Hochschulen in Deutschland. WISTA 1. Statistisches Bundesamt. Wiesbaden.

Büttner, G., Feranec, J., Jaffrain, G., 2002. Corine land cover update 2002: Technical guidelines. Technical report 89. European Environment Agency. Copenhagen.

Büttner, G., Kosztra, B., 2017. CLC2018 technical guidelines. Technical Report Service Contract No 3436/R0-Copernicus/EEA.56665. Environment Agency Austria. Vienna.

Cabrer-Borrás, B., Serrano-Domingo, G., 2007. Innovation and R&D spillover effects in Spanish regions: A spatial approch. Research Policy 36, 1357–1371. doi:10.1016/j.respol.2007.04.012.

Castelló-Climent, A., Chaudhary, L., Mukhopadhyay, A., 2018. Higher education and prosperity: From catholic missionaries to luminosity in India. The Economic Journal 128, 3039–3075. doi:10.1111/ecoj.12551.

Chen, X., Nordhaus, W.D., 2011. Using luminosity data as a proxy for economic statistics. Proceedings of the National Academy of Sciences 108, 8589–8594. doi:10.1073/pnas.1017031108.

Cihlar, J., Jansen, L.J.M., 2001. From land cover to land use: A methodology for efficient land use mapping over large areas. The Professional Geographer 53, 275–289. doi:10.1080/00330124.2001.9628460.

Cinnirella, F., Streb, J., 2017. The role of human capital and innovation in economic development: Evidence from post-Malthusian Prussia. Journal of Economic Growth 22, 193–227. doi:10.1007/s10887-017-9141-3.

Comber, A.J., Wadsworth, R., Fisher, P., 2008. Using semantics to clarify the conceptual confusion between land cover and land use: The example of 'forest'. Journal of Land Use Science 3, 185–198. doi:10.1080/17474230802434187.

Comin, D., Licht, G., Pellens, M., Schubert, T., 2019. Do companies benefit from public research organizations? The impact of the Fraunhofer Society in Germany. ZEW Discussion Paper No. 19-006. ZEW – Leibniz Centre for European Economic Research. Mannheim.

Commission of Experts for Research and Innovation (EFI), 2010. Research, innovation and technological performance in Germany – EFI report 2010. EFI, Berlin.

Commission of Experts for Research and Innovation (EFI), 2018. Research, innovation and technological performance in Germany – EFI report 2018. EFI, Berlin.

Commission of Experts for Research and Innovation (EFI), 2020. Research, innovation and technological performance in Germany – EFI report 2020. EFI, Berlin.

Copernicus Land Monitoring Service, 2020. CLC seamless data coverage. V2020_v20u1. Copernicus Land Monitoring Service. URL: https://land.copernicus.eu/user-corner/technical-library/clc-country-coverage-1990-2018-v20u1.pdf.

Cowan, R., Zinovyeva, N., 2013. University effects on regional innovation. Research Policy 42, 788–800. doi:10.1016/j.respol.2012.10.001.

Davis, M.A., Fisher, J.D.M., Whited, T.M., 2014. Macroeconomic implications of agglomeration. Econometrica 82, 731–764. doi:10.3982/ECTA9029.

Demircioglu, M.A., Audretsch, D.B., 2019. Public sector innovation: The effect of universities. The Journal of Technology Transfer 44, 596–614. doi:10.1007/s10961-017-9636-2.

Dettori, B., Marrocu, E., Paci, R., 2012. Total factor productivity, intangible assets and spatial dependence in the European regions. Regional Studies 46, 1401–1416. doi:10.1080/00343404.2010.529288.

Dewan, A.M., Yamaguchi, Y., 2009. Land use and land cover in Greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization. Applied Geography 29, 390–401. doi:10.1016/j.apgeog.2008.12.005.

Dingel, J.I., Miscio, A., Davis, D.R., 2021. Cities, lights, and skills in developing economies. Journal of Urban Economics 125, 103174. doi:10.1016/j.jue.2019.05.005.

Dingle Robertson, L., King, D.J., 2011. Comparison of pixel- and object-based classification in land cover change mapping. International Journal of Remote Sensing 32, 1505–1529. doi:10.1080/01431160903571791.

Donaldson, D., Storeygard, A., 2016. The view from above: Applications of satellite data in economics. Journal of Economic Perspectives 30, 171–198. doi:10.1257/jep.30.4.171.

Duro, D.C., Franklin, S.E., Dubé, M.G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sensing of Environment 118, 259–272. doi:10.1016/j.rse.2011.11.020.

Eggenberger, C., Rinawi, M., Backes-Gellner, U., 2018. Occupational specificity: A new measurement based on training curricula and its effect on labor market outcomes. Labour Economics 51, 97–107. doi:10.1016/j.labeco.2017.11.010.

Egger, P.H., Koethenbuerger, M., Loumeau, G., 2022. Local border reforms and economic activity. Journal of Economic Geography 22, 81–102. doi:10.1093/jeg/lbab030.

Elvidge, C.D., Baugh, K.E., Kihn, E.A., Kroehl, H.W., Davis, E.R., Davis, C.W., 1997. Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. International Journal of Remote Sensing 18, 1373–1379. doi:10.1080/014311697218485.

Elvidge, C.D., Zhizhin, M., Ghosh, T., Hsu, F.C., Taneja, J., 2021. Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. Remote Sensing 13, 922. doi:10.3390/rs13050922.

Enders, J., 2010. Hochschulen und Fachhochschulen, in: Simon, D., Knie, A., Hornbostel, S. (Eds.), Handbuch Wissenschaftspolitik. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 443–456.

Esri, Maxar, Earthstar Geographics, GIS User Community, 2009. World Imagery (updated January 14, 2022. Dataset. Esri. Redlands, CA. URL: https://services.arcgisonline.com/ArcGIS/rest/services/World_Imagery/MapServer.

European Environment Agency (EEA), 2017. Copernicus Land Service – Pan-European component: CORINE Land Cover. EEA. Copenhagen.

European Environment Agency (EEA), 2020. Copernicus Land Monitoring Service: Corine Land Cover. Dataset. European Union. URL: https://www.eea.europe.eu/ds_resolveuid/DAT-195-en.

Eyles, A., Machin, S., 2019. The introduction of academy schools to England's education. Journal of the European Economic Association 17, 1107–1146. doi:10.1093/jeea/jvy021.

Federal Ministry of Education and Research (BMBF), 2004. Die Fachhochschulen in Deutschland. BMBF, Bonn, Berlin.

Federal Ministry of Education and Research (BMBF), 2018a. Bundesbericht Forschung und Innovation 2018: Forschungs- und innovationspolitische Ziele und Maßnahmen. BMBF, Berlin.

Federal Ministry of Education and Research (BMBF), 2018b. Organisationen und Einrichtungen in Forschung und Wissenschaft: Organisationenband Bundesbericht Forschung und Innovation 2018. BMBF, Berlin.

Feldman, M.P., Florida, R., 1994. The geographic sources of innovation: Technological infrastructure and product innovation in the United States. Annals of the Association of American Geographers 84, 210–219. doi:10.1111/j.1467-8306.1994.tb01735.x.

Feranec, J., Hazeu, G., Christensen, S., Jaffrain, G., 2007. Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia). Land Use Policy 24, 234–247. doi:10.1016/j.landusepol.2006.02.002.

Fisher, P., Comber, A.J., Wadsworth, R., 2005. Land use and land cover: Contradiction or complement, in: Fisher, P., Unwin, D.J. (Eds.), Re-presenting GIS. John Wiley & Sons Ltd, West Sussex, pp. 85–98.

Foster, A.D., Rosenzweig, M.R., 2003. Economic growth and the rise of forests. The Quarterly Journal of Economics 118, 601–637. doi:10.1162/00335530332l675464.

Freitag, M., Bühlmann, M., 2003. Die Bildungsfinanzen der Schweizer Kantone: Der Einfluss sozioökonomischer Bedingungen, organisierter Interessen und politischer Institutionen auf die Bildungsausgaben im kantonalen Vergleich. Swiss Political Science Review 9, 139–168. doi:10.1002/j.1662-6370.2003.tb00403.x.

Fritsch, M., Schwirten, C., 1999. Enterprise-university co-operation and the role of public research institutions in regional innovation systems. Industry and Innovation 6, 69–83. doi:10.1080/13662719900000005.

Fritsch, M., Slavtchev, V., 2007. Universities and innovation in space. Industry and Innovation 14, 201–218. doi:10.1080/13662710701253466.

Gallego, J., Bamps, C., 2008. Using CORINE land cover and the point survey LUCAS for area estimation. International Journal of Applied Earth Observation and Geoinformation 10, 467–475. doi:10.1016/j.jag.2007.11.001.

Gao, B.C., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment 58, 257–266. doi:10.1016/S0034-4257(96)00067-3.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random Forests for land cover classification. Pattern Recognition Letters 27, 294–300. doi:10.1016/j.patrec.2005.08.011.

Goldblatt, R., Heilmann, K., Vaizman, Y., 2020. Can medium-resolution satellite imagery measure economic activity at small geographies? Evidence from Landsat in Vietnam. The World Bank Economic Review 34, 635–653. doi:10.1093/wber/lhz00.

Goldblatt, R., Stuhlmacher, M.F., Tellman, B., Clinton, N., Hanson, G., Georgescu, M., Wang, C., Serrano-Candela, F., Khandelwal, A.K., Cheng, W.H., Balling Jr, R.C., 2018. Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. Remote Sensing of Environment 205, 253–275. doi:10.1016/j.rse.2017.11.026.

Goldblatt, R., You, W., Hanson, G., Khandelwal, A.K., 2016. Detecting the boundaries of urban areas in India: A dataset for pixel-based image classification in Google Earth Engine. Remote Sensing 8, 634. doi:10.3390/rs8080634.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment 202, 18–27. doi:10.1016/j.rse.2017.06.031.

Graf, H., 2011. Gatekeepers in regional networks of innovators. Cambridge Journal of Economics 35, 173–198. doi:10.1093/cje/beq001.

Griffiths, P., van der Linden, S., Kuemmerle, T., Hostert, P., 2013. A pixel-based Landsat compositing algorithm for large area land cover mapping. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6, 2088–2101. doi:10.1109/JSTARS.2012.2228167.

Griliches, Z., 1992. The search for R&D spillovers. The Scandinavian Journal of Economics 94, S29–S47. doi:10.2307/3440244.

Hachmeister, C.D., Duong, S., Roessler, I., 2015. Forschung an Fachhochschulen aus der Innen- und Außenperspektive: Rolle der Forschung, Art und Umfang. CHE Arbeitspapier Nr. 181. Centrum für Hochschulentwicklung (CHE). Gütersloh.

Han, K.S., Champeaux, J.L., Roujean, J.L., 2004. A land cover classification product over France at 1 km resolution using SPOT4/VEGETATION data. Remote Sensing of Environment 92, 52–66. doi:10.1016/j.rse.2004.05.005.

Helmers, C., Overman, H.G., 2017. My precious! The location and diffusion of scientific research: Evidence from the synchrotron diamond light source. The Economic Journal 127, 2006–2040. doi:10.1111/ecoj.12387.

Henderson, J.V., Squires, T., Storeygard, A., Weil, D., 2018. The global distribution of economic activity: Nature, history, and the role of trade. The Quarterly Journal of Economics 133, 357–406. doi:10.1093/qje/qjx030.

Henderson, J.V., Storeygard, A., Weil, D.N., 2012. Measuring economic growth from outer space. American Economic Review 102, 994–1028. doi:10.1257/aer.102.2.994.

Heymann, Y., Steenmans, C., Croisille, G., Bossard, M., 1994. CORINE land cover: Technical guide. Technical Report EUR 12585 EN. Office for Official Publications of the European Communities. Luxembourg.

Hinz, M., Gothe, K., von Polier, X., Wagner, B., Wimmer, M., Falke, U., 2016. Forschung an Fachhochschulen: Wie aus praxisorientierter Forschung Produkte und Dienstleistungen werden. BMBF, Bonn.

Hodler, R., Raschky, P.A., 2014. Regional favoritism. The Quarterly Journal of Economics 129, 995–1033. doi:10.1093/qje/qju004.

Hoffmann, D., Trischler, H., 2015. Die Helmholtz-Gemeinschaft in historischer Perspektive, in: Mlynek, J., Bittner, A. (Eds.), 20 Jahre Helmholtz-Gemeinschaft. Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V., Bonn, pp. 9–47.

Hohn, H.W., 2010. Außeruniversitäre Forschungseinrichtungen, in: Simon, D., Knie, A., Hornbostel, S. (Eds.), Handbuch Wissenschaftspolitik. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 457–477.

Holl, A., 2004. Manufacturing location and impacts of road transport infrastructure: Empirical evidence from Spain. Regional Science and Urban Economics 34, 341–363. doi:10.1016/S0166-0462(03)00059-0.

Holl, A., Peters, B., Rammer, C., 2022. Local knowledge spillovers and innovation persistence in firms. Economics of Innovation and New Technology forthcoming. doi:10.1080/10438599.2022.2036609.

Huang, Q., Yang, X., Gao, B., Yang, Y., Zhao, Y., 2014. Application of DMSP/OLS nighttime light images: A meta-analysis and a systematic literature review. Remote Sensing 6, 6844–6866. doi:10.3390/rs6086844.

Imboden, D.M., 2018. Wider das Kastenwesen: Für mehr institutionelle Durchlässigkeit zwischen Universität and Fachhochschule. GAIA – Ecological Perspectives for Science and Society 27, 353–356. doi:10.14512/gaia.27.4.6.

Intarakumnerd, P., Goto, A., 2018. Role of public research institutes in national innovation systems in industrialized countries: The cases of Fraunhofer, NIST, CSIRO, AIST, and ITRI. Research Policy 47, 1309–1320. doi:10.1016/j.respol.2018.04.011.

Jaffe, A.B., 1989. Real effects of academic research. American Economic Review 79, 957–970.

Jäger, S., 2013. The returns to higher education: Evidence from university openings in Germany. Mimeo.

Janssen, S., Tuor Sartore, S.N., Backes-Gellner, U., 2016. Discriminatory social attitudes and varying gender pay gaps within firms. Industrial and Labor Relations Review 69, 253–279. doi:10.1177/0019793915601633.

Jones, C.I., 1995. Time series tests of endogenous growth models. The Quarterly Journal of Economics 110, 495–525. doi:10.2307/2118448.

Kamhöfer, D.A., Schmitz, H., Westphal, M., 2019. Heterogeneity in marginal non-monetary returns to higher education. Journal of the European Economic Association 17, 205–244. doi:10.1093/jeea/jvx058.

Keola, S., Andersson, M., Hall, O., 2015. Monitoring economic development from space: Using nighttime light and land cover data to measure economic growth. World Development 66, 322–334. doi:10.1016/j.worlddev.2014.08.017.

Key, O., Seeßelberg, C., 2012. Diplom war gestern: Das neue Studium an den Fachhochschulen. Beiträge zur Hochschulforschung 34, 42–65.

Kiener, U., 2013. Die Fachhochschule als Missverständnis: Reform, Identität, Selbstbeschreibung. Swiss Journal of Sociology 39, 341–360.

Köppen, W., 1884. Die Wärmezonen der Erde, nach der Dauer der heissen, gemässigten und kalten Zeit und nach der Wirkung der Wärme auf die organische Welt betrachtet. Meteorologische Zeitschrift 1, 215–226.

Kosztra, B., Büttner, G., Hazeu, G., Arnold, S., 2019. Updated CLC illustrated nomenclature guidelines. Technical Report Service Contract No 3436/R0-Copernicus/EEA.57441 Task 3, D3.1 – Part 1. Environment Agency Austria. Vienna.

Kreckel, R., 2011. Zwischen Spitzenforschung und Breitenausbildung: Strukturelle Differenzierungen an deutschen Hochschulen im internationalen Vergleich, in: Krüger, H.H., Rabe-Kleberg, U., Kramer, R.T., Budde, J. (Eds.), Bildungsungleichheit revisited. 2 ed.. VS Verlang fü Sozialwissenschaften, Wiesbaden, pp. 237–258.

Kulicke, M., Stahlecker, T., 2004. Forschungslandkarte Fachhochschulen: Potenzialstudie. BMBF, Bonn, Berlin.

Kulkarni, R., Haynes, K., Stough, R., Riggle, J., 2011. Light based growth indicator (LGBI): Exploratory analysis of developing a proxy for local economic growth based on night lights. Regional Science Policy & Practice 3, 101–113. doi:10.1111/j.1757-7802.2011.01032.x.

Kyui, N., 2016. Expansion of higher education, employment and wages: Evidence from the Russian Transition. Labour Economics 39, 68–87. doi:10.1016/j.labeco.2016.01.001.

Lackner, H., 2019. Die Stellung der Fachhochschulen im deutschen Hochschulsystem, in: Cai, J., Lackner, H. (Eds.), Jahrbuch Angewandte Hochschulbildung 2016: Deutsch-chinesische Perspektiven und Diskurse. Springer VS, Wiesbaden, pp. 133–157. doi:10.1007/978-3-658-22422-6.

Lechner, M., 2010. The estimation of causal effects by difference-in-difference methods. Foundations and Trends in Econometrics 4, 165–224. doi:10.1561/0800000014.

Leibniz Institute for Economic Research (RWI), Micromarketing-Systeme and Consult GmbH (microm), 2019. RWI-GEO-GRID: Socio-economic data on grid level – Scientific use file (wave 8). Version: 1. Dataset. RWI. Essen. doi:10.7807/microm:suf:V8.

Lepori, B., Kyvik, S., 2010. The research mission of universities of applied sciences and the future configuration of higher education systems in Europe. Higher Education Policy 23, 295–316. doi:10.1057/hep.2010.11.

Lepori, B., Müller, C., 2016. Fachhochschulen als Akteure im schweizerischen Forschungs- und Innovationssystem. Studie im Rahmen des Berichts "Forschung und Innovation in der Schweiz 2016" Teil C, Studie 4. SERI. Bern.

Leten, B., Landoni, P., Van Looy, B., 2014. Science or graduates: How do firms benefit from the proximity of universities? Research Policy 43, 1398–1412. doi:10.1016/j.respol.2014.03.005.

Leyden, D.P., Menter, M., 2018. The legacy and promise of Vannevar Bush: Rethinking the model of innovation and the role of public policy. Economics of Innovation and New Technology 27, 225–242. doi:10.1080/10438599.2017.1329189.

Liu, X., Hu, G., Chen, Y., Li, X., Xu, X., Li, S., Pei, F., 2018. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. Remote Sensing of Environment 209, 227–239. doi:10.1016/j.rse.2018.02.055.

Loveland, T.R., Dwyer, J.L., 2012. Landsat: Building a strong future. Remote Sensing of Environment 122, 22–29. doi:10.1016/j.rse.2011.09.022.

Lulla, K., Nellis, M.D., Rundquist, B., Srivastava, P.K., Szabo, S., 2021. Mission to earth: LANDSAT 9 will continue to view the world. Geocarto International 36, 2261–2263. doi:10.1080/10106049.2021.1991634.

Lyzenga, D.R., 1981. Remote sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and Landsat data. International Journal of Remote Sensing 2, 71–82. doi:10.1080/01431168108948342.

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., Liu, Y., 2017. A review of supervised object-based land-cover image classification. ISPRS Journal of Photogrammetry and Remote Sensing 130, 277–293. doi:10.1016/j.isprsjprs.2017.06.001.

Markham, B.L., Helder, D.L., 2012. Forty-year calibrated record of earth-reflected radiance from Landsat: A review. Remote Sensing of Environment 122, 30–40. doi:10.1016/j.rse.2011.06.026.

Markham, B.L., Storey, J.C., Williams, D.L., Irons, J.R., 2004. Landsat sensor performance: History and current status. IEEE Transactions on Geoscience and Remote Sensing 42, 2691–2694. doi:10.1109/TGRS.2004.840720.

Masek, J.G., Wulder, M.A., Markham, B., McCorkel, J., Crawford, C.J., Storey, J., Jenstrom, D.T., 2020. Landsat 9: Empowering open science and applications through continuity. Remote Sensing of Environment 248, 111968. doi:10.1016/j.rse.2020.111968.

Matejicek, L., Kopackova, V., 2010. Changes in croplands as a result of large scale mining and the associated impact on food security studied using time-series Landsat images. Remote Sensing 2, 1463–1480. doi:10.3390/rs2061463.

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing 17, 1425–1432. doi:10.1080/01431169608948714.

Mellander, C., Lobo, J., Stolarick, K., Matheson, Z., 2015. Night-time light data: A good proxy measure for economic activity? PLoS One 10, e0139779. doi:10.1371/journal.pone.0139779.

Mellor, A., Boukir, S., Haywood, A., Jones, S., 2015. Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. ISPRS Journal of Photogrammetry and Remote Sensing 105, 155–168. doi:10.1016/j.isprsjprs.2015.03.014.

Meuer, J., Rupietta, C., Backes-Gellner, U., 2015. Layers of co-existing innovation systems. Research Policy 44, 888–910. doi:10.1016/j.respol.2015.01.013.

Meurer, P., 2018. Zugang von FH-/HAW-Absolventinnen und -Absolventen zur Promotion, kooperative Promotionen und Promotionsrecht. Studien zum deutschen Innovationssystem No. 16-2018. Commission of Experts for Research and Innovation. Berlin.

Michalopoulos, S., Papaioannou, E., 2013. Pre-colonial ethnic institutions and contemporary African development. Econometrica 81, 113–152. doi:10.3982/ECTA9613.

Millard, K., Richardson, M., 2015. On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. Remote Sensing 7, 8489–8515. doi:10.3390/rs70708489.

Morain, S.A., 1998. A brief history of remote sensing applications, with emphasis on Landsat, in: Liverman, D., Moran, E.F., Rindfuss, R.R., Stern, P.C. (Eds.), People and pixels: Linking remote sensing and social science. The National Academies Press, Washington, DC, pp. 28–50. doi:10.17226/5963.

Moretti, E., 2004. Workers' education, spillovers, and productivity: Evidence from plant-level production functions. American Economic Review 94, 656–690. doi:10.1257/0002828041464623.

Mowery, D.C., Sampat, B.N., 2005. Universities in national innovation systems, in: Fagerberg, J., Mowery, D.C., Nelson, R.R. (Eds.), The Oxford handbook of innovation. Oxford University Press, Oxford, pp. 1–26.

Myint, S.W., Gober, P., Brazel, A., Grossmann-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sensing of Environment 115, 1145–1161. doi:10.1016/j.rse.2010.12.017.

Neumann, K., Herold, M., Hartley, A., Schmullius, C., 2007. Comparative assessment of CORINE2000 and GLC2000: Spatial analysis of land cover data for Europe. International Journal of Applied Earth Observation and Geoinformation 9, 425–437. doi:10.1016/j.jag.2007.02.004.

Nikolai, R., Ebner, C., 2013. Zur Verknüpfung von Berufsbildung und Hochschulbildung: Reformen in Deutschland, Österreich und der Schweiz, in: Amos, K., Schmid, J., Schrader, J., Thiel, A. (Eds.), Kultur – Ökonomie – Globalisierung: Eine Erkundung von Rekalibrierungsprozessen in der Bildungspolitik. Nomos Verlagsgesellschaft, Baden-Baden, pp. 63–79. doi:10.5771/9783845244983-1.

Organization for Economic Co-operation and Development (OECD), 2010. The OECD innovation strategy: Getting a head start on tomorrow. OECD Publishing, Paris. doi:10.1787/9789264083479-en.

Organization for Economic Co-operation and Development (OECD), 2015. Frascati Manual 2015: Guidelines for collecting and reporting data on research and experimental development. The measurement of scientific, technological and innovation activities, OECD Publishing, Paris. doi:10.1787/9789264239012-en.

Organization for Economic Co-operation and Development (OECD), 2017. Education at a glance 2017: OECD indicators. OECD Publishing, Paris. doi:10.1787/eag-2017-en.

Peiró-Palomino, J., 2016. European regional convergence revisited: The role of intangible assets. The Annals of Regional Science 57, 165–194. doi:10.1007/s00168-016-0767-4.

Pekkarinen, A., Reithmaier, L., Strobl, P., 2009. Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data. ISPRS Journal of Photogrammetry and Remote Sensing 64, 171–183. doi:10.1016/j.isprsjprs.2008.09.004.

Pérez-Hoyos, A., García-Haro, F.J., San-Miguel-Ayanz, J., 2012. A methodology to generate a synergetic land-cover map by fusion of different land-cover products. International Journal of Applied Earth Observation and Geoinformation 19, 72–87. doi:10.1016/j.jag.2012.04.011.

Peri, G., Shih, K., Sparber, C., 2015. STEM workers, H-1B visas, and productivity in US cities. Journal of Labor Economics 33, S225–S255. doi:10.1086/679061.

Petzina, D., 2005. Der Bologna-Prozess in Deutschland: Stand und Perspektiven. Beiträge zur Hochschulforschung 27, 18–36.

Pfister, C., 2017. Different educational structures and their economic impact on individuals and the economy. Dissertation. University of Zurich. doi:10.5167/uzh-150061.

Pfister, C., Koomen, M., Harhoff, D., Backes-Gellner, U., 2021. Regional innovation effects of applied research institutions. Research Policy 50, 104197. doi:10.1016/j.respol.2021.104197.

Pfister, C., Rinawi, M., Harhoff, D., Backes-Gellner, U., 2018. Regional innovation effects of applied research institutions. Swiss Leading House "Economics of Education" Working Paper No. 117 (first version 2016). Leading House VPET-ECON. Zurich.

Pinkovskiy, M., Sala-i-Martin, X., 2016. Lights, camera ... income! Illuminating the national accounts–household surveys debate. The Quarterly Journal of Economics 131, 579–631. doi:10.1093/qje/qjw003.

Ponds, R., van Oort, F., Frenken, K., 2007. The geographical and institutional proximity of research collaboration. Papers in Regional Science 86, 423–444. doi:10.1111/j.1435-5957.2007.00126.x.

Popp, D., 2017. From science to technology: The value of knowledge from different energy research institutions. Research Policy 46, 1580–1594. doi:10.1016/j.respol.2017.07.011.

Projektgruppe Bund–Kantone Hochschullandschaft 2008, 2004. Bericht über die Neuordnung der schweizerischen Hochschullandschaft. SERI, Bern.

Robin, S., Schubert, T., 2013. Cooperation with public research institutions and success in innovation: Evidence from France and Germany. Research Policy 42, 149–166. doi:10.1016/j.respol.2012.06.002.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing 67, 93–104. doi:10.1016/j.isprsjprs.2011.11.002.

Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C., Roberts, D., 2008. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. Remote Sensing of Environment 112, 2272–2283. doi:10.1016/j.rse.2007.10.004.

Rouse Jr, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. Progress Report RSC 1978-1. Texas A&M University Remote Sensing Center. College Station.

Roy, D.P., Wulder, M.A., Loveland, T.R., Woodcock, C.E., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., Zhu, Z., 2014. Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment 145, 154–172. doi:10.1016/j.rse.2014.02.001.

Rumerman, J.A., 1999. NASA historical data books (SP-4012) volume VI: NASA space applications, aeronautics and space research and technology, tracking and data acquisition/Support operations, commercial programs, and resources, 1979–1988. NASA History Division, Washington, DC. Updated October 15, 2010.

Rupietta, C., Backes-Gellner, U., 2019. How firms' participation in apprenticeship training fosters knowledge diffusion and innovation. Journal of Business Economics 89, 569–597. doi:10.1007/s11573-018-0924-6.

Schlegel, T., Pfister, C., Harhoff, D., Backes-Gellner, U., 2022. Innovation effects of universities of applied sciences: An assessment of regional heterogeneity. The Journal of Technology Transfer 47, 63–118. doi:10.1007/s10961-020-09839-w.

Schneider, A., 2012. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. Remote Sensing of Environment 124, 689–704. doi:10.1016/j.rse.2012.06.006.

Schuler, M., Dessemontet, P., Joye, D., 2005. Eidgenössische Volkszählung 2000: Die Raumgliederungen der Schweiz. Statistik der Schweiz. Swiss Federal Statistical Office. Neuchâtel.

Schultheiss, T., Pfister, C., Backes-Gellner, U., Gnehm, A.S., 2019. Tertiary education expansion and task demand: Does a rising tide lift all boats? Swiss Leading House "Economics of Education" Working Paper No. 154. Leading House VPET-ECON. Zurich.

Schultheiss, T., Pfister, C., Backes-Gellner, U., Gnehm, A.S., 2021. Tertiary education expansion and task demand: Does a rising tide lift all boats? Swiss Leading House "Economics of Education" Working Paper No. 154 (first version 2019). Leading House VPET-ECON. Zurich.

Squicciarini, M., Dernis, H., Criscuolo, C., 2013. Measuring patent quality: Indicators of technological and economic value. OECD Science, Technology and Industry Working Papers 2013/03. Organisation for Economic Co-operation and Development. Paris. doi:10.1787/5k4522wkw1r8-en.

State Secretariat for Education, Research and Innovation (SERI), 2015. Hochschulen und Forschung in der Schweiz. SERI, Bern.

State Secretariat for Education, Research and Innovation (SERI), 2019. Higher education and research in Switzerland. SERI, Bern.

Sterlacchini, A., 2008. R&D, higher education and regional growth: Uneven linkages among European regions. Research Policy 37, 1096–1107. doi:10.1016/j.respol.2008.04.009.

Sutton, P.C., Costanza, R., 2002. Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. Ecological Economics 41, 509–527. doi:10.1016/S0921-8009(02)00097-6.

Swiss Coordination Centre for Research in Education (SCCRE), 2018. Swiss Education Report 2018. SCCRE, Aarau.

Swiss Federal Statistical Office (SFSO), 2002. NOGA: General Classification of Economic Activities: Title. SFSO, Bern.

Swiss Federal Statistical Office (SFSO), 2016. Schweizerische Lohnstrukturerhebung (LSE) 2014: Codebook. SFSO, Bern.

Toivanen, O., Väänänen, L., 2016. Education and invention. The Review of Economics and Statistics 98, 382–396. doi:10.1162/REST_a_00520.

Toner, P., 2010. Innovation and vocational education. The Economic and Labour Relations Review 21, 75–98. doi:10.1177/103530461002100206.

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Marcantonio, M., Ricotta, C., Tonon, G., 2019. Estimating tree species diversity from space in an alpine conifer forest: The Rao's Q diversity index meets the spectral variation hypothesis. Ecological Informatics 52, 26–34. doi:10.1016/j.ecoinf.2019.04.001.

Traustason, B., Snorrason, A., 2008. Spatial distribution of forests and woodlands in Iceland in accordance with the CORINE land cover classification. Icelandic Agricultural Sciences 21, 39–47.

Trianni, G., Lisini, G., Angiuli, E., Moreno, E.A., Dondi, P., Gaggia, A., Gamba, P., 2015. Scaling up to national/regional urban extent mapping using Landsat data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8, 3710–3719. doi:10.1109/JSTARS.2015.2398032.

U.S. Geological Survey (USGS), 2018. Landsat Collections. Fact Sheet 2018-3049. USGS. Sioux Falls. doi:10.3133/fs20183049.

Valero, A., Van Reenen, J., 2019. The economic impact of universities: Evidence from across the globe. Economics of Education Review 68, 53–67. doi:10.1016/j.econedurev.2018.09.001.

Vandenbussche, J., Aghion, P., Meghir, C., 2006. Growth, distance to frontier and composition of human capital. Journal of Economic Growth 11, 97–127. doi:10.1007/s10887-006-9002-y.

Wang, E.C., 2010. Determinants of R&D investment: The Extreme Bounds-Analysis approach applied to 26 OECD countries. Research Policy 39, 103–116. doi:10.1016/j.respol.2009.11.010.

Warnecke, C., 2019. Welche Rolle nehmen Universitäten und Fachhochschulen in regionalen Innovationssystemen ein?, in: Cai, J., Lackner, H. (Eds.), Jahrbuch Angewandte Hochschulbildung 2017: Deutsch-chinesische Perspektiven und Diskurse. Springer VS, Wiesbaden, pp. 207–230. doi:10.1007/978-3-658-22103-4.

Waser, L.T., Schwarz, M., 2006. Comparison of large-area land cover products with national forest inventories and CORINE land cover in the European Alps. International Journal of Applied Earth Observation and Geoinformation 8, 196–207. doi:10.1016/j.jag.2005.10.001.

Whiteside, T.G., Boggs, G.S., Maier, S.W., 2011. Comparing object-based and pixel-based classifications for mapping savannas. International Journal of Applied Earth Observations and Geoinformation 13, 884–893. doi:10.1016/j.jag.2011.06.008.

Williams, D.L., Goward, S., Arvidson, T., 2006. Landsat: Yesterday, today, and tomorrow. Photogrammetric Engineering & Remote Sensing 72, 1171–1178. doi:10.14358/PERS.72.10.1171.

Winters, J.V., 2014. Foreign and native-born STEM graduates and innovation intensity in the United States. IZA Discussion Paper No. 8575. IZA – Institute of Labor Economics. Bonn.

Wissenschaftsrat, 2002. Empfehlungen zur Entwicklung der Fachhochschulen. Drs. 5102/02, Wissenschaftsrat, Berlin.

Wong, T.T., 2015. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognition 48, 2839–2846. doi:10.1016/j.patcog.2015.03.009.

Working Committee of the Surveying Authorities of the Laender of the Federal Republic of Germany (AdV), 1991. Verzeichnis der flächenbezogenen Nutzungsarten im Liegenschaftskataster und ihrer Begriffsbestimmungen. Nutzungsartenverzeichnis. AdV. Mainz.

Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sensing of Environment 122, 2–10. doi:10.1016/j.rse.2012.01.010.

Wulder, M.A., White, J.C., Goward, S.N., Masek, J.G., Irons, J.R., Herold, M., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2008. Landsat continuity: Issues and opportunities for land cover monitoring. Remote Sensing of Environment 112, 955–969. doi:10.1016/j.rse.2007.07.004.

Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global Landsat archive: Status, consolidation, and direction. Remote Sensing of Environment 185, 271–283. doi:10.1016/j.rse.2015.11.032.

Xue, J., Su, B., 2017. Significant remote sensing vegetation indices: A review of developments and applications. Journal of Sensors 2017. doi:10.1155/2017/1353691.

Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., Ermon, S., Burke, M., 2020. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. Nature Communications 11, 2583. doi:10.1038/s41467-020-16185-w.

Yu, W., Zang, S., Wu, C., Liu, W., Na, X., 2011. Analyzing and modeling land use land cover change (LUCC) in the Daqing City, China. Applied Geography 31, 600–608. doi:10.1016/j.apgeog.2010.11.019.

Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International Journal of Remote Sensing 24, 583–594. doi:10.1080/01431160304987.

Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment 118, 83–94. doi:10.1016/j.rse.2011.10.028.

Ziegele, F., Roessler, I., Mordhorst, L., 2019. Hochschultyp im Wandel?, in: Cai, J., Lackner, H. (Eds.), Jahrbuch Angewandte Hochschulbildung 2016: Deutsch-chinesische Perspektiven und Diskurse. Springer VS, Wiesbaden, pp. 159–174. doi:10.1007/978-3-658-22422-6_11.

Curriculum Vitae

Personal Details

Name	Patrick Lehnert
Date of Birth	December 4, 1990

Education

08/2016–07/2020	PhD studies in Management & Economics, University of Zurich
10/2013–03/2016	Master of Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, Labor Market and Human Resources
10/2010–08/2013	Bachelor of Arts, Friedrich-Alexander-Universität Erlangen-Nürnberg, Economics (Focus: Business Education)

Professional Experience

since 07/2016	Research Assistant at the Chair for Business and Personnel Economics, University of Zurich
04/2016–06/2016	Research Assistant at the Institute for Employment Research, Nuremberg
08/2013–09/2014 & 02/2015–03/2016	Student Assistant at the Institute for Employment Research, Nuremberg
10/2011–07/2013	Student Assistant at the School of Business (Business Mathematics), Friedrich-Alexander-Universität Erlangen-Nürnberg
03/2011–04/2011	Internship at Netigate Deutschland GmbH, Wiesbaden
07/2010–09/2010	Internship at GI Group Deutschland GmbH, Wiesbaden