Universities and Science-Based Innovation in the Private Sector

Martin Watzinger Lukas Treber Monika Schnitzer*

July 3, 2018

This paper shows that hiring a new professor increases science-based innovation in the local private sector. To establish causality we use data on newly hired professors for a German university, along with information on the runners-up for the same position. When a professor is hired, the patents of local companies start to cite her scientific articles relatively more than those of the runners-up. Furthermore, local patents become more similar to her scientific articles, and patent classes associated with the new professor's research specialization grow relatively more. At face value, our estimates imply that a new professor induces corporate science-based innovation with a value of up to half a million dollars per year. The change in local science-based innovation is driven primarily by PhD graduates working in the private sector. This suggests that universities produce not only scientific research but also the absorptive capacity necessary for corporate science-based innovation.

JEL Codes: O30, O38, I23, I28, J61, H52, R11

*Martin Watzinger and Monika Schnitzer: LMU Munich, Department of Economics, Akademiestraße 1, 80799 Munich, Germany; Lukas Treber: University of Hohenheim, Institute of Economics, 70593 Stuttgart; Watzinger (corresponding author): martin.watzinger@econ.lmu.de. We thank Pierre Azoulay, Christian Catalini, Bernd Fitzenberger, and Jeff Furman, as well as seminar participants at MIT, Harvard, the NBER Productivity Seminar, and the Munich Summer Institute for helpful comments. We gratefully acknowledge financial support of the Deutsche Forschungsgemeinschaft through SFB-TR 190.

The best way to send information is to wrap it up in a person.

J. Robert Oppenheimer

1 Introduction

What is the impact of hiring a new professor on local private sector-innovation? Answering this question is challenging because professors are not hired at random. In this paper, we identify the effect of a newly hired scientist by using information on the runners-up for the same position. In Germany, hiring committees are required by law to draw up a short-list of suitable candidates for each senior faculty appointment. Due to the incentives inherent in the institutional set-up, all candidates on the short-list are comparable. We have access to the short-lists of all appointments of one large university between 1980 and 2005. This data allows us to implement a clean and novel identification strategy to measure the effect of new hires.

In our emprical analysis, we compare the number of local patents based on the scientific articles of the hired scientist with the number of local patents based on articles of the runners-up before and after the move. We find that when a new professor is hired, corporate innovation based on her academic research increases. The likelihood of the articles of the new scientist to be cited in local patents increases by more than 100% relative to the articles of the runners-up. Taking our estimates at face value, a new professor induces an increase in patent value of up to half a million dollars per year. This is not a pure salience effect. The new scientist changes the direction of innovation around the university. The abstracts of patents become more similar to the abstracts of the scientific articles of the new professor. Furthermore, patent classes related to the research of the new professor grow relatively more.

We also provide insights into the mechanism of this effect. The induced change in science-based innovation in the private sector is driven primarily by inventors with a PhD degree from the local university. This suggests that PhD graduates working for local companies play an important role in the transfer of academic science to the private sector. Universities provide not only scientific input but by training graduate students also increase the absorptive capacity of corporations necessary for science-based innovation.

Our empirical strategy needs to address two key challenges. First, to measure the impact of a newly hired scientist we need to determine which patents are based on the articles of the new professor. Second, for identification purposes, we need to make sure that we can use the runners-up to construct a counterfactual for the hired professor.

To measure the impact of a newly hired professor we focus on "science-based patents;" defined as patents that directly cite an academic article. Science-based patents account for only around 12% of all patents but they are particularly valuable. To show this, we combine the data of Ahmadpoor and Jones (2017) on the distance of each U.S. patent to science with data on patent values from Kogan *et al.* (2017).[1] We find that U.S. science-based patents are on average more valuable than patents not based on scientific research. A patent directly citing an academic article has an average value of 18.46 million dollars, while the average value of patents completely disconnected from science is 8.33 million dollars.

The runners-up can serve as the counterfactual for the hired professor if - in the absence of the move - the number of local patents based on research of the hired professor and on research of the runners-up would have the same trend. We argue that this is a plausible assumption due to the institutional features of the German hiring process. This assumption is also supported by the observed characteristics of the hired professor and the runners-up. Almost all senior professors are civil servants and are hired in a highly regulated multi-step process that starts with the public advertisement of an open position and results in a ranked short-list of two to four suitable candidates. Offers are then made to the candidates in order of the rank on the short-list. The first candidate to accept an offer is appointed. The hiring committee has a strong incentive to weed out unsuitable candidates since it is very difficult to predict whether or not the first-ranked candidate will accept the offer. So there is a considerable likelihood that lower ranked candidates may receive and accept an offer and stay until retirement. In our data, we find similar ex-ante observables such as age, publication and citation record for the candidates that are on the same list, which suggests that incentives seem to be working for the hiring committees. We also show that the trends in article-to-article and patent-to-article citations are parallel prior to the move.

We have access to 837 ranked short-lists for tenured professorships, listing 2,227 researchers across a wide range of subjects from humanities to natural sciences and medicine for one university. For our empirical analysis, we make use of 417 short-lists from those departments that have at least

[1]We thank Mohammad Ahmadpoor and Ben Jones for sharing their data on patent-to-article links.

one scientist whose research has been cited at least once by a patent.[2] Altogether, there are 1113 scientists on these 417 short-lists. Thus, for each professor that is hired, we have 1.66 runners-up who were short-listed but not appointed. We match all researchers on the short-lists with their academic publications in Microsoft Academic along with their patent-to-article and article-to-article citations (Tang *et al.*, 2008; Sinha *et al.*, 2015). We also use the geolocated patent data of Morrison *et al.* (2017).

In our main analysis, we show that hiring new faculty increases local corporate science-based innovation. While prior to moving, candidates on the short-lists are similar in their patent-to-article citation record - an indication that they are of comparable quality - after the move citations from local patents to the articles of the moving scientiest increase relative to citations to articles of her runners-up. The effect sets in around three years after the new professor is hired and is still measureable ten years after the move. To quantify the effect we weight each patent by average dollar value by year and technology class based on the data of Kogan *et al.* (2017). If we assume that the patent citations induced by the new professor come from patents that otherwise would not have been invented, then the upper bound in value of a new hire for private sector patents is up to half a million dollars a year.

We investigate next whether an increase in patent-to-article citations simply reflects the fact that local research is more salient to inventors from local companies or whether the increase in citations captures a change in the direction and quantity of local innovation. Local inventors may hear about or meet newly hired faculty members, but it is not clear whether this makes them incorporate new research insights into their own innovation activities.

To show that we measure a real change in local innovation, we analyze whether patents change their content after the recruitment of the new professor. To do this, we calculate for all scientific articles of all scientists on the short-lists the pairwise text similarity of the abstract of the article and all abstracts of local patents. We find that after hiring a new professor, local patents become more similar to her scientific articles. The effect is concentrated in science-based patents.

We also study whether the growth of patent classes is influenced by the hiring of the new professor. For this purpose, we calculate for each scientist the ex-ante likelihood that she will be cited by a patent in a given technology class. We do this by measuring for each journal the

[2] This implies that we do not use data for the humanities and the social sciences.

probability that an article in this journal is cited in a particular technology class. Then we multiply this probability with the publication profile of each scientist and aggregate it up on the person level. In the empirical analysis, we compare the growth of patent classes that are more related to the hired scientist as measured by the ex-ante probability of being cited with the growth of patent classes that are more related to the runners-up. We find that patent classes associated with the newly hired professor grow relatively more.

To get a sense of whether our results are specific to the university whose short-lists we use for identification or hold more generally, we report the geographic pattern of science-based patents throughout Germany. We use geolocated patent data of all EPO, U.S. and WIPO patents with an inventor based in Germany and determine for each patent whether it cites scientific articles. Patents and in particular science-based patents are clustered around universities. Within 10 km of a university location, 3.53 patents are filed per 100 inhabitants, while this number falls to close to 0.54 between 90 and 100 km. Close to universities, around 20% of all patents cite a scientific article while this share falls to 5% more than a 100 km away. This cross-sectional correlation between science-based patents and universities is consistent with the idea that local companies learn from scientific research done at universities.

Finally, we investigate the mechanism for the causal effect of university hirings. For this analysis, we use data on the universe of German PhD graduates from the German National Library. German law requires that a copy of all PhD dissertations submitted at German universities are sent to the National Library. We obtained a list of the bibliographic entries for all PhD dissertations from 1965 to the present date. We show that the increase in patents based on articles of new professors is almost exclusively driven by PhD graduates working in the private sector. Around 80% of all excess citations are from PhD graduates. Furthermore, the effect is stronger for PhD graduates from the local university than from PhD graduates in general. This supports the notion that the science-specific training acquired at the local university influences the direction of the PhD inventors' future patenting activities. Our causal estimations are consistent with descriptive evidence on the universe of German PhD graduates which shows that on average local PhD graduates are twice as likely to cite research from their own university than other PhD graduates living at the same location.

The main result of our paper is to show that hiring a professor increases science-based innovation around the university. This finding contributes to the literature in two ways. First, it suggests that basic research at universities generates value for local private sector innovation. Our paper thus adds to a growing literature that aims to estimate the effect of science on private sector innovation and the local effect of universities.[3] For example, Ahmadpoor and Jones (2017) show that 21% of all U.S. patents directly link backward to a research article. Azoulay *et al.* (2015) find that increased funding from the U.S. National Institutes of Health (NIH) for basic biomedical research increases patenting by private sector companies. Belenzon and Schankerman (2013) show that scientific articles are mostly cited in close proximity to the university. Hausman (2017) finds that the Bayh-Dole Act, which increased the incentives of universities to invest in innovation, resulted in local employment and growth. In contrast, Andrews (2017) finds no positive effect of universities on local patenting that would go beyond a pure agglomeration effect.

The second implication of our main result is that new professors cause local inventors to use new information. This suggests that labor mobility can shape the rate and direction of local innovation in a regional economic cluster. The paper is thus closely related to studies on the effects of labor mobility of scientists.[4] Azoulay *et al.* (2011) study superstars in medicine moving to different universities, but find no effects on local patent-to-article citations relative to a matched sample of non-moving scientists. In a historical context, Moser *et al.* (2014) find that the forcible expulsion of Jewish scientists increased innovation of U.S. chemical companies working in the same field as émigrés. Our study shows in a contemporaneous commercial setting that the hiring policy of a university can influence local innovation.

Exploring the likely mechanism, we show that PhD graduates working in the private sector provide the absorptive capacity for the corporate sector to translate science to innovation. This observation is consistent with the argument of Florida (2014) that cities need a "creative class" of knowledge workers to innovate. Our analysis contributes to the literature studying the requirements of private companies to benefit commercially from basic science. Cohen and Levinthal (1990) have emphasized that an important reason for corporations to do R&D is to build the absorptive capacity for using scientific input in their innovation process. Cockburn and Henderson (1998) show for pharmaceutical companies how their drug discovery performance benefits from investments in

[3] For example, Aghion *et al.* (2009); Andrews (2017); Azoulay *et al.* (2015); Belenzon and Schankerman (2013); Cantoni and Yuchtman (2014); Gaetani and Bergolis (2015); Hausman (2017); Kantor and Whalley (2014); Liu (2015); Valero and Van Reenen (2016); Bikard and Marx (2018)

[4] Among others Azoulay *et al.* (2010, 2011); Borjas and Doran (2015, 2012); Moser *et al.* (2014); Waldinger (2010, 2012).

absorptive capacity through in-house basic research and coauthorships of scientific papers with publicly funded researchers. The importance of scientists as translators of scientific insights is shown also by Agrawal *et al.* (2006). They find that commercial products based on licenses from universities are more successful if the university scientists helps as a consultant. Our study suggests that by hiring scientists and fostering graduate education policy can help corporations make the best use of academic research for their innovation efforts.

Understanding the role of university scientists for private sector-innovation is important both for public policy and for the optimal organization of R&D in private companies. For governments, the size of a university is a policy relevant margin as faculty size can be scaled up relatively easily and flexibly. Thus it is important to know whether hiring more faculty benefits local companies. For companies, our results imply that co-location with universities and hiring local graduate students can give better access to scientific advances and thus lay the basis for a competitive advantage in the production of science-based products.

The rest of this paper is organized as follows. In the next section, we present our empirical strategy and the data used for our analysis. In section 3, we estimate the causal effect of a newly hired professor on private sector innovation. In section 4, we discuss the mechanism for the knowledge transfer from universities to the private sector by looking at the absorptive capacity provided by PhD graduates who work in the private sector. Section 5 concludes.

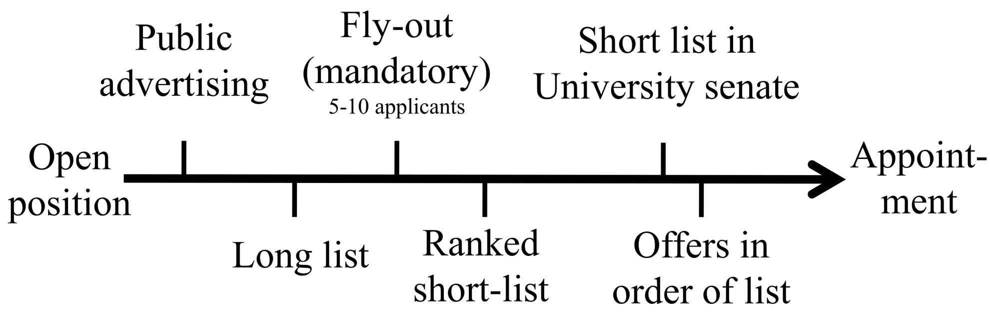
2 Institutional set-up and data

In this section, we describe the institutional features of the German hiring procedure that form the basis of our identification strategy. Then we discuss our data sources.

2.1 The German system of appointing professors

In Germany, almost all university professors are civil servants and thus are hired in a highly regulated multi-step process (Figure 1). The procedure is designed to give every qualified applicant equal access to jobs in the public service independent of personal connections. To implement equal access, every open position must be advertised in a national newspaper. The advertisement must contain a list of criteria by which the candidates are compared in the remainder of the process. These criteria

Figure 1: Procedure of appointing a professor in Germany



might include publications in international refereed journals or experience in raising third party funding. Using these criteria, the hiring committee puts together a long list of 5 to 10 candidates who are invited for fly-outs.

After the fly-outs, the hiring committee creates a ranked short-list of two to four candidates, based also on reports of at least two external referees. The ranked short-list and the external reports are submitted to the university senate for review. If the ranked short-list is approved, offers are made to the candidates in order of the rank on the short-lists. The first candidate to accept an offer is then appointed.

The hiring process contains several mechanisms with the aim to make the process objective and fair. First, internal candidates are usually not eligible to apply for tenured positions, so almost all new professors are external hires from other universities.[5] This rule prevents nepotism because it requires researchers to move at least once to get a tenured position at the senior level. Second, the composition of the hiring committee is regulated and contains several external members. The hiring committee has at least one professor in the same field but from another university, one member of the university senate from another field, a women's representative, a representative of non-tenured scientific employees and one undergraduate student representative. Third, the whole process is subject to court review. If one of the non-appointed candidates suspects that the university did not follow due process, the candidate can sue for non-appointment of the chosen candidate, compensation and invalidation of the list ("Konkurrentenklage").

On average, every open professorship in Germany attracted 41.8 applications in 2013 (Wis-

[5]Internal promotions are in theory possible but must follow much more stringent rules than external appointments. The rules are so stringent that for all practical purposes this is perceived (and labeled) as forbidden (Ban of internal promotions - "Hausberufungsverbot").

senschaftskonferenz, 2014). Around 10% of these candidates were considered suitable for the short-list, which implies that the average list had four candidates. Of all candidates, 45% received an offer for the position at one point in time. If a candidate received an offer, the probability that she accepted was around 50%.[6]

2.2 Data

We collect data on all hired professors along with their runners-up for one German university from 1980 to 2005. In total, we have access to 837 ranked short-lists for tenured professorships, listing 2,227 researchers. The university under consideration offers a wide range of subjects from humanities to natural sciences and medicine.

For our empirical study, we focus on 417 lists from those departments that have at least one scientist with a patent-to-article citation according to our data. Altogether, there are 1113 scientists on these lists. Thus, for each professor that is hired we have 1.66 runners-up who were short-listed but not appointed. Due to data availability, we use in the different specifications different subsets of these lists. In the citation analysis, we use 201 lists with 573 scientists that have at least one scientist with a positive number of patent-to-article citations. In the analysis on text similarity, we use 190 lists with 519 scientists that have at least one article with an abstract available in the data.

For each candidate, we collect the curriculum vitae. From the CVs, we know when a researcher moved to another universities. In cases where we cannot find a curriculum vitae we use data from historical course catalogs to infer whether a professor who received an offer was actually appointed.

We match all researchers on the short-lists with their academic publications and their citation record. For academic publications, we us the data from Microsoft Academic (Tang *et al.*, 2008; Sinha *et al.*, 2015).[7] The data contains for each publication the authors, year, journal, citations, keywords, the field of study, and the abstract. As source for patent data, we use PATSTAT along with information from Morrison *et al.* (2017) that provides a geo-location for all inventors and all assignees for EPO-patents (European Patent Office), U.S.-patents and WO-patents (World Intellectual Property

[6]The data is not exact as open professorships are counted in a two year interval. According to the data 1,612 professor were appointed in Germany in 2013. The universities received 67,117 applications for these positions. 6,954 researchers made it on the ranked list and 3,175 received an offer.

[7]A snapshot of this data is available for download for free at the website of the Open Academic Society. Open Academic Society website: https://www.openacademic.ai/oag/ (last accessesed 2018-02-12)

Organization).

For all articles of the short-listed researchers, we collect all citations by academic articles (article-to-article citations) and by patents (patent-to-article citations) and all available article abstracts.

To shed light on the potential mechanism driving the effects of university hirings we use data on PhD graduates. German law requires that a copy of all PhD dissertations submitted at German universities are sent to the National Library.[8] From the National Library, we obtained a list of the bibliographic entries for all PhD dissertations from 1965 to the present date. The information includes the name of the author, the field of study, the title of the dissertation, the university and the year of graduation. The data set covers approximately 1.2 million dissertations of PhD graduates.

We match each PhD graduate to patents by name, field of study and graduation period. We find 108,428 matches. In 24% of these matches, the inventors also mention their doctoral title on the patent. We use this data to classify whether a patent has at least one inventor with a PhD. Around 39% of all patents in Germany have at least one inventor with a PhD. 77% of these PhD patents have at least one inventor who directly mentions her doctoral title.

To see whether our results are consistent with the overall picture in Germany, we collect administrative data on German universities and municipalities. Germany had 396 colleges and universities with 2.7 million students in the academic year 2015/2016. In our analysis, we focus on the 39 universities that are among the Top 500 universities according to the Shanghai Ranking of 2017. Germany has four universities in the Top 100: Heidelberg, Goettingen, the LMU Munich, and the TU Munich.

3 Empirical strategy

Our aim is to measure the impact of hiring a new professor on local innovation. We do this by comparing local innovation that is based on the research of the hired professor with local innovation that is based on the research of the runners-up. To implement this strategy we need to determine in a first step which patents in the private sector are based on the articles of which research. In a second step, we need to make sure that we can use the runners-up to construct a counterfactual for the hired professor. In this section, we discuss each of this two points in turn.

[8] A bibliographic database of the German National Library is available here: `http://www.dnb.de/EN/Home/home_node.html`

3.1 Measurement: Relating private-sector innovation to university scientists

To measure the link between the scientific research produced by the newly hired professor and corporate innovation we use information on how often articles of the new professor are cited in patents of private companies. We call patents that cite an academic article science-based patents. Patents cite academic articles to give credit to the researcher who contributed to the disclosed technology of the patent.[9] Patent-to-article citations are used in many recent papers to capture the link between science and innovation, e.g. Arora *et al.* (2017) and Azoulay *et al.* (2015). Roach and Cohen (2013) suggest that patent-to-article citations reflect knowledge flows from academia to the private sector better than the commonly used patent-to-patent citation. Yet, others have voiced concerns that they might be a noisy measure and reflect salience rather than knowledge flows (Bikard and Marx, 2018).[10]

This is why we look at two more outcome variables. In Section 4.2, we investigate whether the text of patents becomes more similar to the text of the scientific articles of the hired scientist. Furthermore, in Section 4.3, we also look at the growth of patent classes. Both of these analyses are based on all patents, not only on science-based patents.

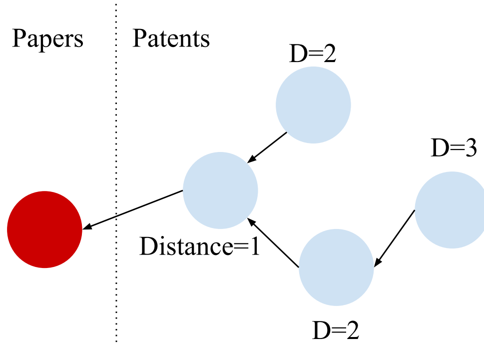
While science-based patents represent only a subset of all patents, we argue that they capture a highly relevant link between corporate innovation and university research as shown by the fact that they are more valuable than non-science-based patents. To establish this stylized fact, we classify each patent by its citation distance to science, following Ahmadpoor and Jones (2017) (upper panel of figure 2). A patent that directly cites a scientific paper has the distance of D=1, a patent that cites a (D=1)-patent but no scientific article has a distance of D=2, and so on. We then use information on patent values from Kogan *et al.* (2017) to calculate the value generated by science-based patents as compared to that of non-science-based patents for all U.S. patents.

Figure 2 shows the average patent value by distance to science. Patents are on average more valuable if they are more closely related to science. A science-based patent that directly cites an academic article (D=1) has an average value of 18.46 million dollars, while patents that are

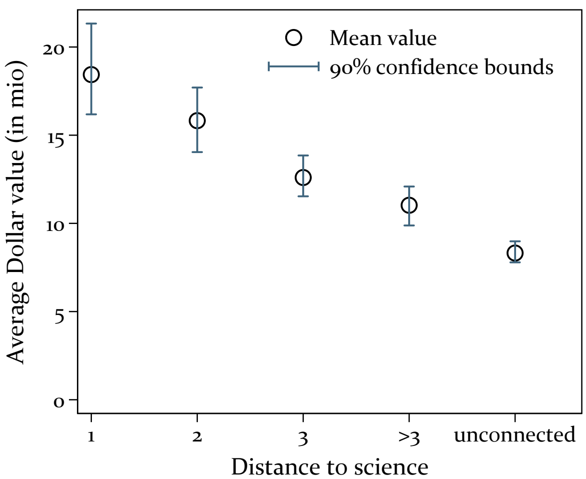
[9]Patent-to-article citations are not a mechanic result from university researcher patenting their technology as it is not possible in Europe to patent inventions that were already disclosed in an article. This requirement of "absolute novelty" is thus much more stringent than the application of a grace period in the US.

[10]Callaert *et al.* (2014) show that not all relevant articles are cited and sometime cites are only background information.

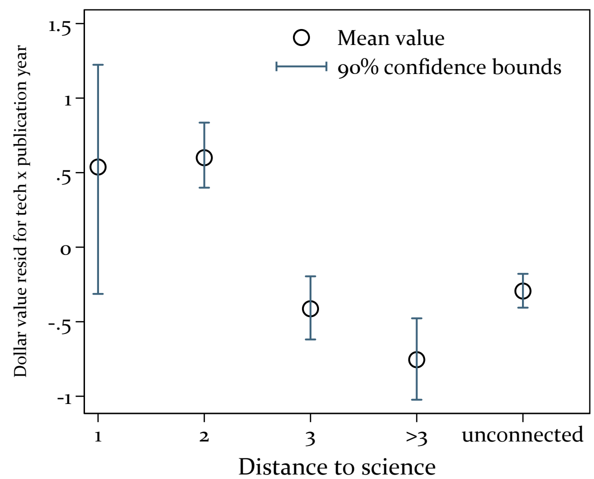
Figure 2: Value of patent by distance to science



(a) Raw



(b) Controlling for technology and year



Note: Value of U.S. patents from Kogan *et al.* (2017) with distance to science of U.S. patents from Ahmadpoor and Jones (2017). Average values along with 90% confidence bounds are displayed. The confidence bounds are based on standard errors bootstrapped by USPC technology class.

completely disconnected from science have on average a value of 8.33 million dollars (Figure 2a).[11][12]

This holds true also within technologies. In Figure 2b we subtract the mean value by USPC technology class and filing year. Science-based patents with a distance of D=1 are on average 0.53 million dollars more valuable than the average patent, while completely disconnected patents are 0.3 million dollars less valuable than the average patent.

[11]The raw data for value of patents and distance to science are shown in Figure 10 in Appendix A.

[12]Unconnected patents might include patents that are based on science but do not acknowledge that link.

3.2 Identification: Using short-lists to construct the counterfactual

To measure whether hiring a new professor increases local science-based innovation we need to compare how science-based innovation evolves after the new professor is hired with what would happen if the professor were not hired. To construct this counterfactual outcome we use the runners-up. The runners-up can serve as the counterfactual for the hired professor if in the absence of the move the number of local patents based on research of the hired professor and on research of the runners-up would have the same trend.[13]

While this assumption is untestable, we argue that it is plausible given the incentives inherent in the institutional set-up of the hiring procedure described above. It is also supported by pre-move observables and by the facts that the pre-trend is parallel and not different between non-movers, as shown below.

The candidates on the short-list are of similar quality for three institutional reasons. First, all candidates applied for the position with the advertised and fixed salary category, pension benefits and the associated teaching obligations. Thus, candidates self-selected to the position. Second, the hiring committee has a strong incentive to weed out unsuitable candidates. The reason is that it is difficult for the hiring committee to predict whether or not the first-ranked candidate will accept the offer.[14] In our estimation sample, the first-ranked candidates accept with a probability of slightly less than 80%. If the first-ranked candidate does not accept the offer, a lower-ranked candidate receives and might accept an offer. Since all senior professors in Germany are appointed for life, this lower-ranked candidate might stay until retirement. Third, hiring no one has an option value because the hiring committee can go back to the market next year. For all these reasons, the hiring committee has strong incentives to put only candidates on the short-list that they consider acceptable for the position.

This hiring procedure results in candidates on the short-lists who are similar on observables, both in levels and in trends. Table 1 shows summary statistics for the short-listed candidates - separately for movers and runners-up - covering the publications prior to the move. The differences are not economically significant and also not statistically different from zero. Up to the year prior

[13] Throughout the paper we call all non-hired candidates runners-up, independent of their rank on the short-list.

[14] Candidates can receive competing offers during the selection process and thus might have better offers on the table once the process is complete. What is more, receiving an offer from a different university usually opens the door to renegotiation at the current home university.

Table 1: Summary statistics before movement

	Mover	Runners-up	Diff	P-Value
Year of List	1992.82	1993.64	0.81	0.28
Time since first article	8.39	8.22	-0.16	0.48
Natural Sciences	0.54	0.55	0.01	0.78
Medicine	0.46	0.45	-0.01	0.78
Number of articles	17.29	17.75	0.47	0.74
Local citations to articles (<100km)	3.67	3.03	-0.64	0.29
Total citations to articles	187.30	230.36	43.06	0.16
Local patent-to-article citation (<100km)	0.05	0.05	-0.01	0.86
Total patent-to-article citation	0.62	0.67	0.05	0.87
Average article-patent similarity	0.07	0.07	-0.00	0.44
Article-patent similarity of closest patent	0.12	0.12	-0.00	0.58
Number	417	696		

Note: This table shows averages for the full sample (column 1) and separately for movers (column 2) and for runners-up (column 3). The last two columns show the result of a t-test with unequal variances between columns 2 and 3.

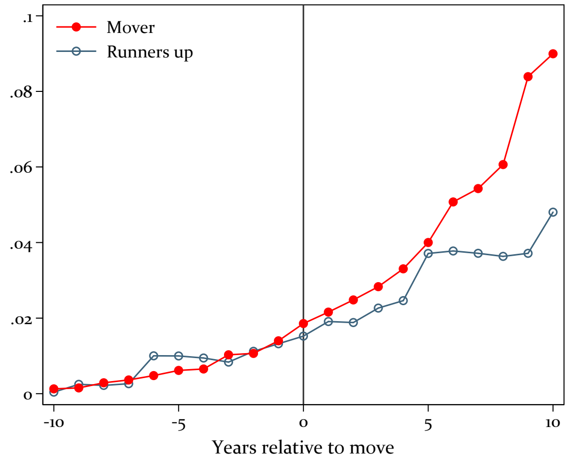
to the short-list, a short-listed and finally appointed researcher has on average 17 articles with a total of 187 article-to-article-citations, of which 3.7 are local. The runners-up are of the same age measured by the time from the first article and have marginally more articles compared to the runners-up.

The number of citations from other articles - a common measure of article quality - is similar for both groups, both in levels (Table 1) and in trends. Subfigures 3a and 3b present the mean article-to-article citations over time for the 100 km around the university and overall. Prior to the move in $t = 0$, the trends for the hired professor and the runners-up are indistinguishable. After the move, the local citations to the hired professor go up relative to the runners-up in the region close to the university, while overall the citation patterns for the two groups stay the same. In Subfigure 3c, we show the average trends for the non-moving scientists, distinguishing between higher-ranked and lower-ranked non-movers. Again, we do not see any difference up to five years after the potential hiring. This is evidence that the scientists on the short-lists are similar in terms of scientific output and diffusion of scientific ideas at the time of the hiring. As academic output is a targeted moment of the hiring committee, these summary statistics speak in favor of incentives working.

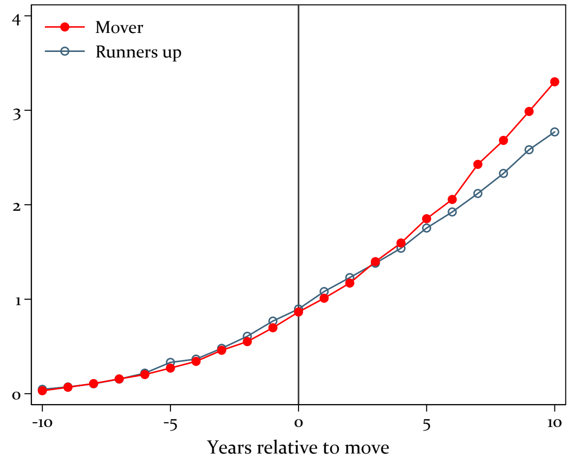
According to our identification assumption, we should see the same trend in the outcome variable

Figure 3: Average number of local article-to-article citations

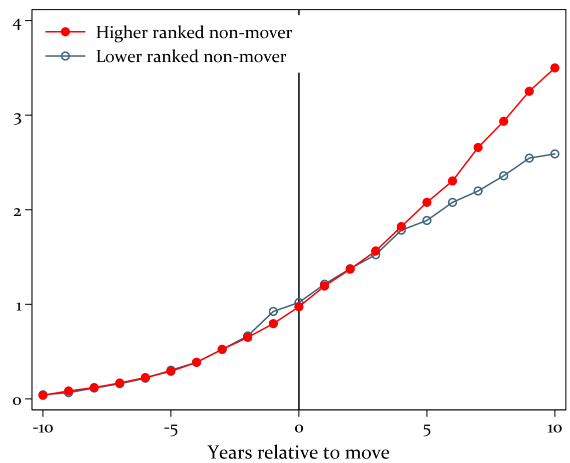
(a) Within 100 km



(b) All



(c) Non-mover- all



Note: This figure shows the evolution of article-to-article citations of the hired professor and her runners-up. Subfigure a) uses article-to-article citations within 100 km of the university as outcome. Subfigure b) uses overall article-to-article citations as outcome. In Subfigure c), we compare the higher ranked non-mover with the lower-ranked non-mover using overall article-to-article citations.

patent-to-article citations for the hired professor and the runners-up. In section 4.1, we show that the pre-trends in patent-to-article citations are indeed parallel. Furthermore, in Appendix A.2, we show that there is also a parallel trend between non-movers.

4 Results

In this section, we show that hiring a professor increases local science-based innovation. Hiring more faculty is the policy relevant margin because faculty size can be scaled up relatively easily whereas the opening of a new university requires much higher investments and hence is a much less flexible policy instrument.[15]

We first show that after a professor is hired, more patents cite her academic article relative to articles of the runners-up. We demonstrate next that the direction and quantity of local patents changes. The content of patents becomes more similar to the academic article of the hired professor, and the patent classes associated with the newly hired professor grow relatively more. This confirms that the observed increase in patent-to-article citations does not simply reflect the fact that local research is more salient to inventors from local companies. Lastly, we document that - consistent with the estimated effects - science-based patenting is clustered over proportionally close to universities all over Germany.

4.1 More patents cite academic articles

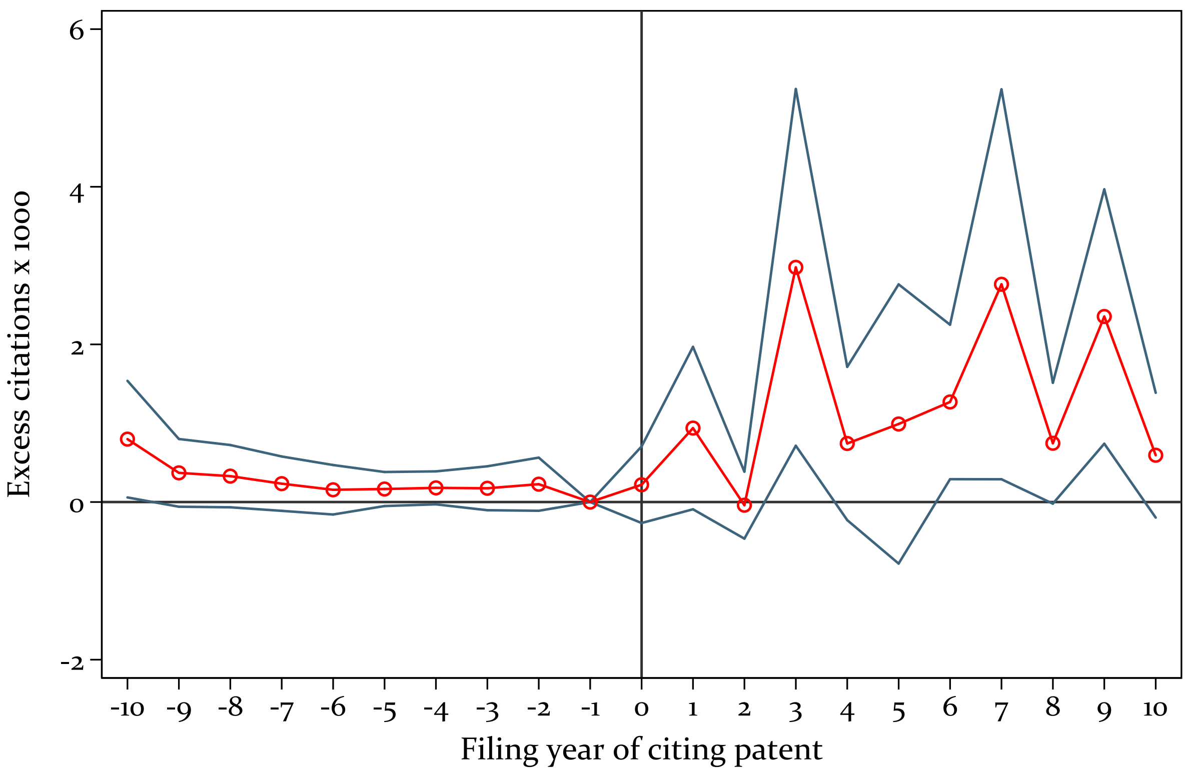
For our main result, we compare the average number of patent citations to articles of the newly hired scientists with those of their runners-up over time, using the following regression model with time-varying coefficients:

$$\#Citations_{i,k,t} = \sum_{t=-10}^{10} \beta_t \cdot Treatment_k + Controls_{tk} + \epsilon_{i,t} \tag{1}$$

The dependent variables $\#Citations_{i,k,t}$ are yearly patent-to-article citations within 100 km of

[15] From 2006 to 2016, the number of full professors increased by 21% in Germany, while no new university was opened (Figure 12b in Appendix B). Figure 12a in Appendix B shows all openings of German universities included in the Shanghai ranking. The 39 openings span the years 1386 (Heidelberg) to 1972 (Essen), resulting in one university opening every 15 years on average. Jäger (2017) studies the effects of openings of German universities in the 1970s on wages. Andrews (2017) and Cantoni and Yuchtman (2014) study the effect of university openings in a historical setting. In a contemporary setting, Pfister *et al.* (2017) study the opening of non-research colleges in Switzerland.

Figure 4: Patent-to-article citations



Note: This figure shows the average number of patent-to-article citations for articles of the hired professor relative to articles of the non-moving scientists on the same list, along with 90 percent confidence intervals. Time is counted relative to the move in $t = 0$. Standard errors are clustered on the level of the researcher.

the university to article i of author k in year t. The main independent variable, $Treatment_k$, is an indicator variable equal to one if the author k is appointed as a professor at the university, and zero otherwise. The estimate for β_t measures the average number of yearly excess citations to articles of the moving relative to citations to articles of the non-moving researchers on the same short-list. As controls, we use researcher and year fixed effects. Standard errors are clustered on the person level. For this specification, we can use 201 lists that have at least one scientist with a positive number of patent-to-article citations. These lists contain 573 scientists.

Figure 4 shows the estimates for β_t in equation (1), measuring the number of excess patent-to-article citations within a distance of 100 km. Prior to the appointment of the professor, there are no differences in citations between articles of treatment and control group scientists. This speaks in favor of parallel trends. Starting with the move, citations to articles of the moving researcher increase relative to those of the non-moving researchers. The coefficients turn significant in year

Table 2: Universities and science-based innovation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline		Weighted		Types		
	PtA	D1+D2	Dollars	Quality	Company	University	Others
Move x Post	1.1**	7.9	25.1**	1.8	0.8**	0.2	0.1
	(0.4)	(11.9)	(11.2)	(2.4)	(0.3)	(0.2)	(0.1)
Mean Dep.	0.6	15.4	14.5	2.5	0.4	0.1	0.1
R2 (within)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Obs.	959532	959532	959532	959532	959532	959532	959532

Note: This table shows the results from a difference-in-differences estimation with ten years before movement as pre-period and ten years after movement as post-period. The estimation equation is:

$$\#Citations_{i,k,t} = \beta_1 \cdot Treatment_k + \beta_2 \cdot Post_t + \beta_3 \cdot Treatment_k \cdot Post_t + Controls_{i,t} + \epsilon_{i,t}$$

where $Treatment_k$ is an indicator if professor k moved to the university and $Post_t$ is an indicator for all years after the move. As controls we use person and year fixed effects. In column (1), we use all citations of patents to academic articles within 100 km as outcome, correcting for self-citations. In column (2), we use the sum of patent-to-article citations and patent-to-patent citations to patents that cite the articles of the scientists. This shows the impact of hiring a professor on D=1 and D=2 patents. In column (3), we weight the patents with their value according to Kogan *et al.* (2017) and in column (4) with the forward citations of the citing patent as a quality measure. In columns (5) to (7), we use citations from different types of assignees. In column (5), the citations are from assignees classified as companies or individuals. In column (6), the citations are from patents of universities, and in column (7), the citations from other assignees such as hospital, government institutions or unclassified assignees. In all regressions, we use the weights suggested by Iacus *et al.* (2012) to identify the average treatment effect on the treated. Standard errors are clustered on the individual level. Coefficients are multiplied by 1000 for better readability. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

three after the move and stay elevated for the next six years. In Appendix A.2, we show that there is no such trend in patent-to-article citations of non-movers.

To calculate the average effect over ten years post move we use the following difference-in-differences specification

$$\#Citations_{i,k,t} = \beta_1 \cdot Treatment_k + \beta_2 \cdot Treatment_k \cdot Post_t + Controls_{tk} + \epsilon_{i,t} \quad (2)$$

where $Post$ is an indicator for the years after the move; i.e., for $t > 0$. We report the results in Table 2. In column (1), we use patent-to-article citations corrected for self-citations within 100 km to identify the local effect. Per article and year, the number of citations increases by 0.0011 patent-to-article citations. As each scientist has on average 17 articles in the ten years before the move, this translates to around 0.0187 additional patent citations per hired researcher. The increase in local patent-to-article citations is almost twice as large as the overall average of local patent-to-article citations. In column (2), we add local patent-to-patent citations to the patent that cites the article. We thus include also citations of patents that indirectly cite the article. The effect is larger, but not significant.

In the following two columns, we present results when we weight citing patents by quality. In column (3), each citing patent is weighted with its dollar value using imputed values based on Kogan *et al.* (2017). Per article, we find an average increase in total patent value of 0.0251 million dollars per year. As each scientist has on average 17 articles prior to the move and we consider 10 years post move, this translates into an upper bound value of 0.43 million dollars created per hired professor per year. In column (4), we weight all patents with their own forward citations. The increase in citation-weighted patents is not significantly different from zero.

In columns (5) to (7), we split the patent-to-article citations by the citing entity. In column (5), we use citations from patents of individuals and companies, in column (6) the citations from universities, and in column (7) from all others such as government entities and hospitals. We find that 72% of the effect measured in the baseline regression can be attributed to patents by companies and individuals, 18% to patents by universities, and 9% to all others.

4.2 Patents become more similar

The increase in patent-to-article citations after a professor is hired can be interpreted in two ways. On the one hand, the increase might reflect an increase in science-based innovation. On the other hand, it might reflect that the move made inventors aware of the articles of the new scientist and that they just added citations without changing the direction or the number of their patents. To investigate whether our results can be explained by a pure salience effect, we study next how local scientific research affects the content of local patents.

If a new professor has an impact on private sector innovation we would expect that patents close to the university become more similar to the scientific work of the new researcher. To calculate the similarity between the abstracts of the article and of the patent we use the "term frequency-inverse document frequency" (tf-idf) method.[16] For each term used in the abstracts of the patent and the article, tf-idf measures how often this word appears in the abstract and then standardizes this value with the probability that this term appears in general. Using the tf-idf value for each term, we

[16] We use the "gensim" implementation for our calculations (see: `https://radimrehurek.com/gensim/`).

can build a word vector for each of the abstracts. Then we determine the similarity between the abstracts of the patent and the article by calculating the correlation between the two word vectors.

To estimate the impact of new hirings on the similarity of articles and patents, we use the difference-in-differences framework of Equation 2, with the similarity measure as outcome variable. In figure 5a, we plot the coefficient of $Move \cdot Treatment$, β_3, for different patent subsamples. The result table is reported in Appendix B.1. For this specification, we use 190 lists with 519 scientists that have at least one article with an available abstract.

Overall, we find an average increase in similarity from a mean of 127.8 by around 1.1 or 0.9% for all articles (closest patent). With 17 articles per professor prior to the move, this implies that hiring a professor results in one additional patent every six to seven years whose measure of similarity to her scientific articles is twice the size than the average. The effect is stronger for patents that are close to science (D=1) and null for patents that are not connected to science at all (D>4). Overall, these findings are consistent with the interpretation that the scientific output of the hired professor affects the direction of locally produced science-based innovation.

In Figure 6, we present the evolution of similarity to the average of the closest five patents in Subfigure a) and the similarity to the closest science-based patent in Subfigure b). The graphs show that the similarity increases post move and stays elevated for almost 10 years post move.

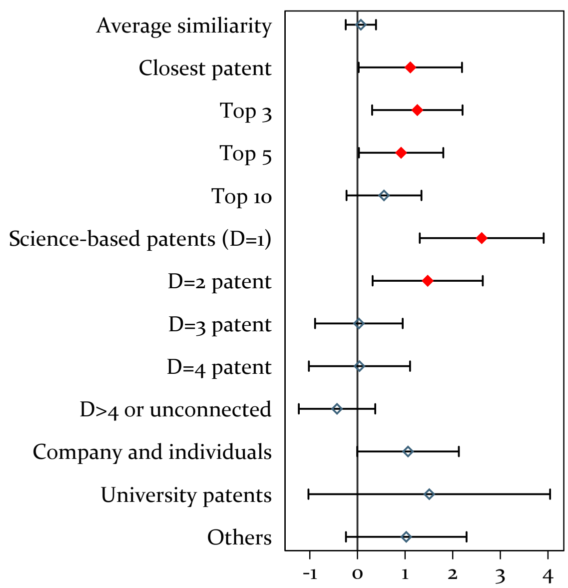
4.3 Related patent classes grow faster

We show next that hiring a new professor increases the rate of local innovation activity. Patent classes that are more related to the hired professor grow faster relative to patent classes that are more related to the other candidates on the short-list.

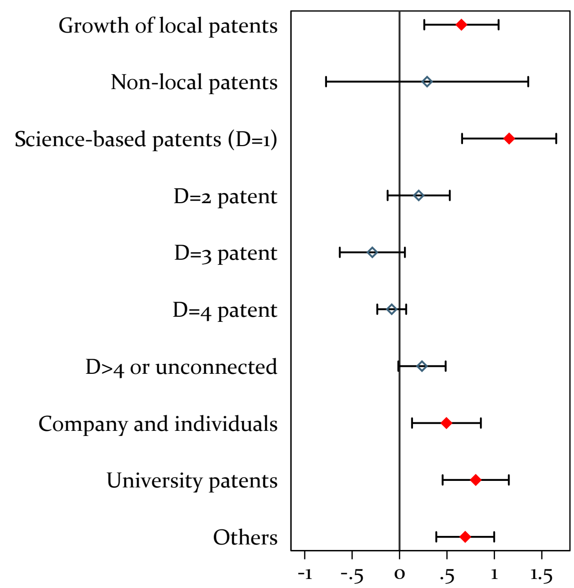
To connect patent-classes to scientific output we calculate the ex-ante probability that a professor with a given publication record is cited in a given class. We do this by measuring for each journal the probability that an article in this journal is cited in a particular technology class. Then we multiply this probability with the publication profile of each scientist and aggregate it up on the person level. The resulting probability for scientist i to be cited in patent class c is denoted by $Prob_{c,i}$ and called specialization in the following.

To quantify the effect of hiring a researcher on patent output we use the "relative specialization" of the hired professor i in class c, $Prob_{c,t,i} - Prob_{c,t,j}$; i.e., the difference between the probabilities

Figure 5: Other outcome variables



(a) Similarity for subsamples



(b) Growth of patents for subsamples

Note: Subfigure a) shows results using the following difference-in-difference specification with different measures for similarity as outcome.

$$Similarity_{i,k,t} = \beta_1 \cdot Treatment_k + \beta_2 \cdot Treatment_k \cdot Post_t + Controls_{i,t} + \epsilon_{i,t}$$

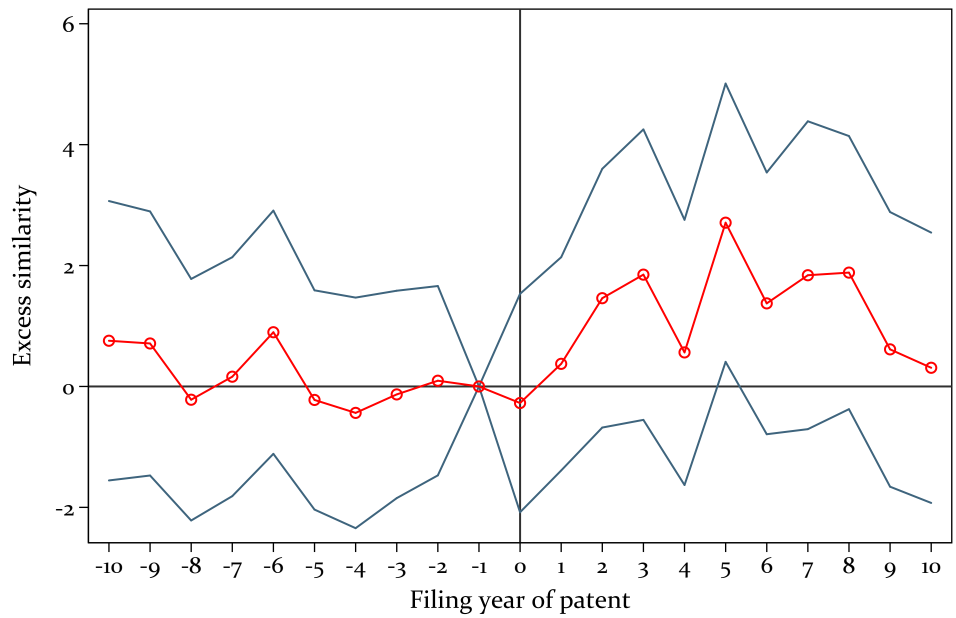
where $Similarity_{i,k,t}$ is the similarity between the abstract of article i and patents filed in year t of scientist k. $Treatment_k$ is an indicator if scientist k was hired. $Post_t$ is an indicator for the period after the move. We use list x year fixed effects as controls. As outcome we use in the first line the average similarity of all patents within 100km of the university as outcome. In line 2, the outcome is the similarity of the most similar patent. In line 3 to 5, we use the 3, 5 and 10 most similar patents. In line 6, we use the most similar patent among all patents with a distance to science. In line 7 to 9, we use the subset of patents with different closeness to science. In the last three lines we use the most similar patent in different subsets defined by being assigned to companies or individuals, universities and all other entitites.
In Subfigure b) we use the following difference in difference specification with the log number of patents in a patent class as outcome

$$ln(\#Patents_{c,t} + 1) = \beta_1 + \beta_2(Prob_{c,i} - Prob_{c,j}) \cdot Post_t + Controls_{i,t} + \varepsilon_{c,t}$$

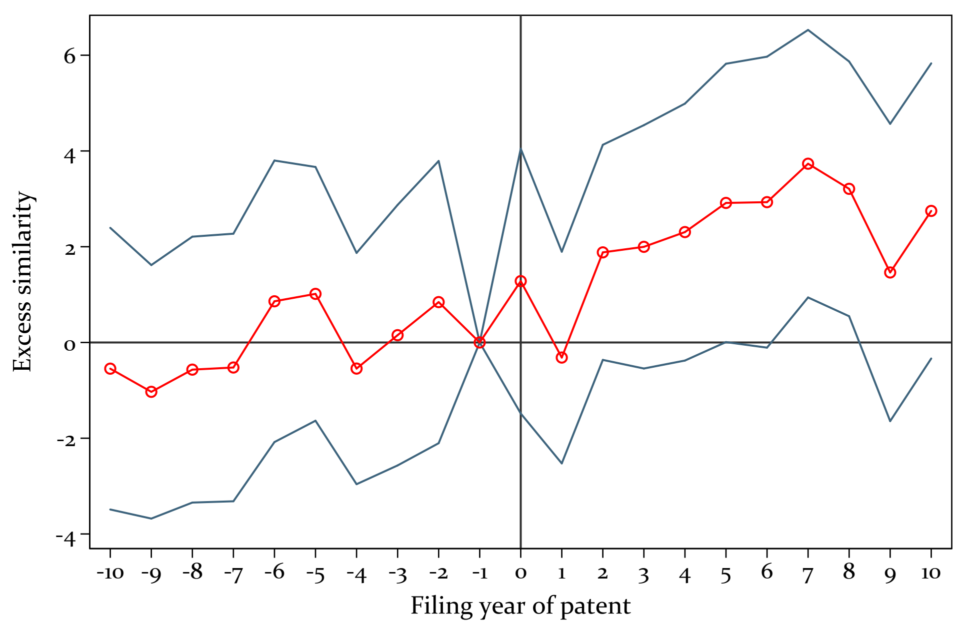
where $Prob_{c,i}$ is the ex-ante likelihood that the hired scientist i is cited in patent class c. As outcome we use in row 1, only patents with inventors within 100 km of the university. In row 2 we use all patents in the rest of Europe. In row 3 we use the number of science-based patents within 100km. In row 4 to 7 we use the number of patents with varying distance to science. In the last three rows we use patents assigned to companies and individuals, universities and other institutions, respectively.
Along with the mean coefficient β_2 we show 90% confidence intervals based on standard errors clustered on scientist level.

Figure 6: Similarity effect over time

(a) Average closest 5



(b) Closest science-based patent



Note: This figure shows the results of a difference-in-difference specification with time-varying coefficients and similarity as outcome. In Subfigure a), we use the similarity between the abstracts of the hired scientists and the runners-up with the average similarity of the five most similar patents within 100 km around the university. In Subfigure b), we use the similarity of the closest science-based patent as dependent variable. 90% confidence intervals based on standard errors clustered on individual level are shown.

of mover and non-mover to be cited in class c, to explain the number of patents within 100 km:

$$ln(\#Patents_{c,t} + 1) = \beta_0 + \beta_1(Prob_{c,i} - Prob_{c,j}) \cdot Post_t + Controls_{i,t} + \varepsilon_{c,t}$$

where $ln(\#Patents_{c,t} + 1)$ is the natural logarithm of the number of patents in patent class c in time t plus one. The relative specialization measures how related a patent class is to the hired professor relative to the runners-up. If there is more than one runner-up on the list we use the average probabilities to construct the relative specialization. We control for list, patent-class and year fixed effects. The results for various subsamples are presented in Figure 5b. The full results table is reported in Appendix B.2.

After the professor is hired, the number of local patents in patent classes in which the mover is specialized grows more strongly than in patent classes in which the mover is not specialized. For each percentage point in probability difference, a patent class grows on average by 0.7 percent. With a mean of 8.2 patents per patent class this is around 6% percent of a patent per year. As expected, the increase is driven by science-based patents; i.e., patents with distance of D=1, and not by patents that are unrelated to science. While - not surprisingly - the increase is strongest for university patents we find a significant increase also for patents filed by companies and by others.

4.4 External validity: Effects consistent with geographic pattern

To get a sense of whether our results are specific to the university on which we base our identification or hold more generally, we take a look at the geographic pattern of science-based patents. For ease of graphical presentation, we focus on the 39 universities that are among the Top 500 universities in the Shanghai Ranking of 2017. We classify all patents with respect to their distance to science, following the method of Ahmadpoor and Jones (2017) described above. For all patents with a German inventor that are filed in the U.S., we can use directly their classification. For all remaining patents (EPO and WO), we calculate the citation distance ourselves.

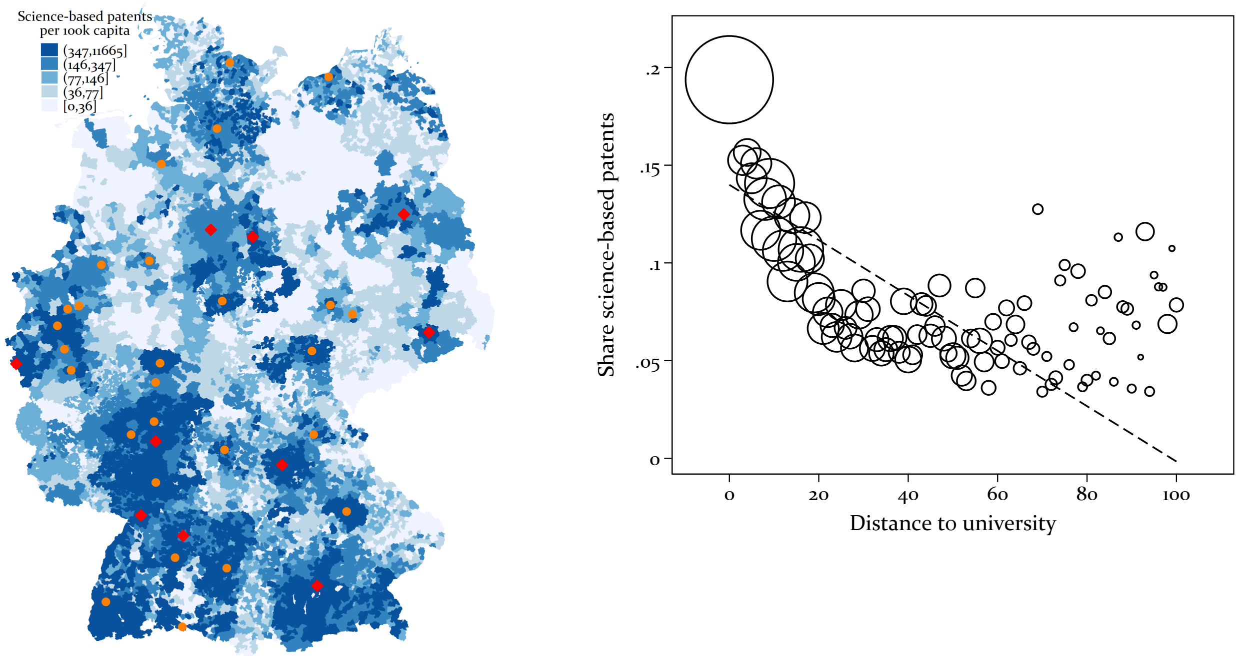
Figure 7 presents on the left hand side the spatial distribution of science-based patents per capita. It shows that science-based patents are clustered around universities. This is not a mere artifact of the spatial agglomeration of patents close to universities.[17] The right hand side of Figure

[17]We show the spatial distribution of patents in absolute and in per capita terms in Appendix B.3.

Figure 7: Science-based patents are clustered around universities

(a) Science-based patents per capita

(b) Share of science-based relative to all patents



Note: These figures show data for the whole of Germany. Subfigure a) shows the number of science-based patents per 100k capita in a municipality. Science-based patents are patents that directly cite an academic article. The geo-located patent data on municipality level is from Morrison *et al.* (2017) and universities in Shanghai ranking. Red dots are technical universities while yellow dots are full universities. If there are both types of universities in one city, the technical university is shown. Subfigure b) shows the share of science-based patents relative to all patents by distance to the closest university in kilometer. We first calculate the share by university and then take averages across universities. The size of bubbles shows the average number of underlying patents.

7 shows the share of science-based patents on all patents by distance to universities. Within 10 km of the university, around 20% of all patents directly cite at least one academic article, while in the 90 to 100 km range, the share is close to 5%. This geographic pattern throughout Germany is consistent with the causal estimates we provided before on the basis of one university.

5 Mechanism

The role of universities is not only to produce research, but also to train undergraduate and graduate students in scientific methods. In this section, we use information on the universe of German PhD graduates to show that by training doctoral students universities produce the absorptive capacity for the private sector to build on scientific research.

Using our data on the short-lists of one large university, we find that inventors with a PhD degree react more to the hiring of a new professor than inventors without a PhD degree. Furthermore, locally trained PhD graduates react more to a new hire than PhD graduates in general, which indicates that graduate education makes a difference for the direction of innovation.

This is consistent with descriptive evidence from the universe of all German PhD graduates. Inventors with a PhD produce a large share of science-based innovation in Germany and use significantly more knowledge from their alma mater than from other universities.

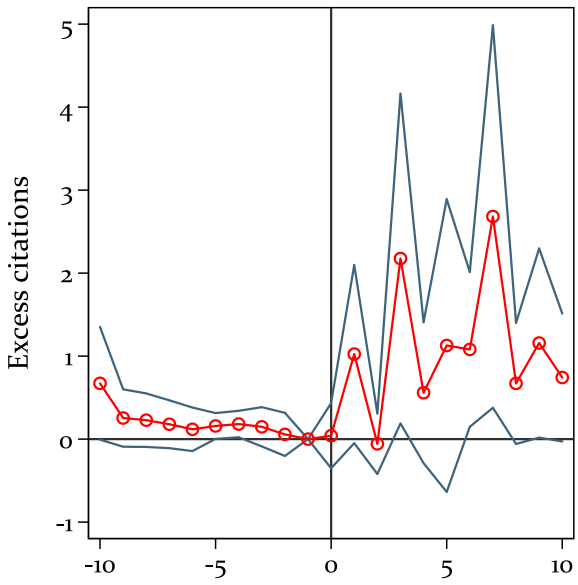
5.1 PhD inventors react to the hiring of a professor

In Figure 8, we split the patent-to-article citations into different subgroups to illustrate that graduate education influences the absorptive capacity of an inventor. In Subfigure a), the dependent variable is citations of patents with at least one inventor with a PhD degree. In Subfigure b), we use citations of patents without any PhD inventor.

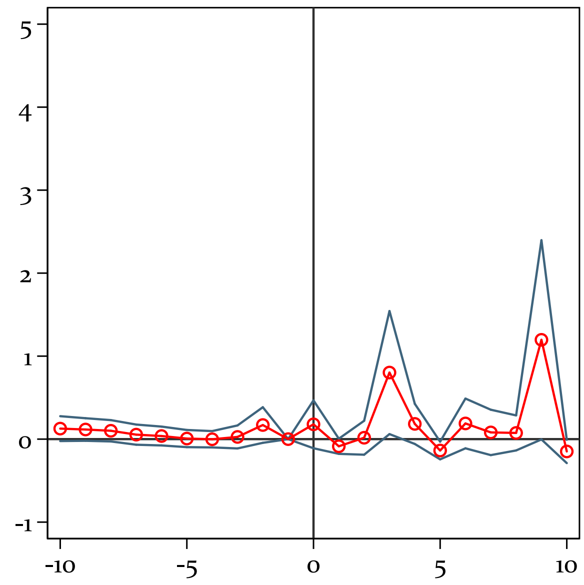
Subfigures a) and b) of Figure 8 show that PhD inventors react more to new hires than Non-PhD inventors. This is confirmed by our estimates presented in Table 3. In Panel B, we look at the patent-to-article citations for the different inventor groups. Out of the total increase of 1.1 citations, 0.9 or around 80% are attributable to patents with a PhD inventor, even though they account for only 56% of all the patents (Panel A). We also see a larger increase in similarity for PhD inventors than for Non-PhD inventors (Panel C). Finally, also the growth in patent classes is driven by an

Figure 8: Citations by different subgroups

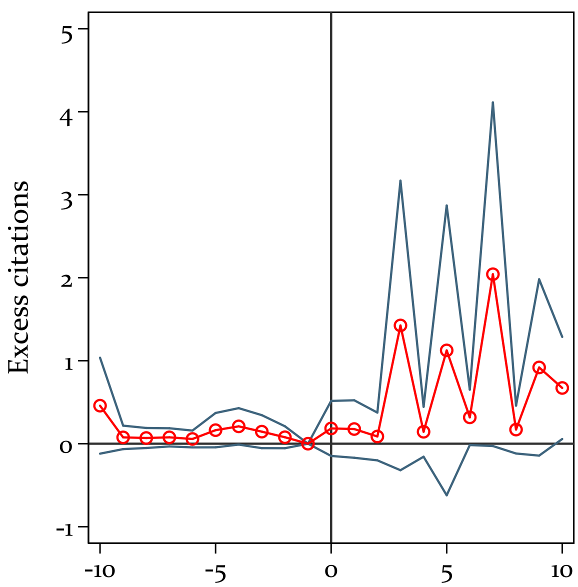
(a) PhD inventors



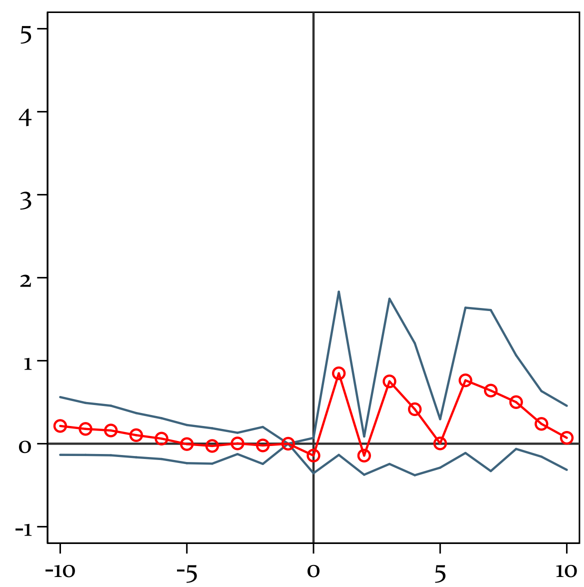
(b) Others



(c) Locally educated PhD inventors



(d) PhD inventors from other universities



Note: These figures shows the average number of patent-to-article citations for articles of the hired professor relative to articles of the non-moving scientists on the same list, along with 90 percent confidence intervals. Time is counted relative to the move in $t = 0$. Standard errors are clustered on the level of the researcher. In Subfigure a), we only use citations of patents that have at least one PhD inventor. In Subfigure b), we use citations of patents without any inventors with PhD. Subfigure c) uses citations of patents of inventors with a PhD from the local university as outcome. In Subfigure d), we use citations of inventors with a PhD from all but the local university as outcome.

Table 3: Results for the PhD subsample

	(1)	(2)	(3)	(4)	(5)
	Panel A: Descriptive				
	Baseline	PhD	No PhD	Local PhD	Non-local PhD
Share of all patents	100%	56%	44%	22%	34 %
	Panel B: Patent citations - # citations				
	Baseline	PhD	No PhD	Local PhD	Non-local PhD
Move x Post	1.1**	0.9***	0.1	0.6**	0.4
	(0.4)	(0.4)	(0.1)	(0.2)	(0.2)
Mean Dep.	0.6	0.5	0.1	0.2	0.2
R2 (within)	0.00	0.00	0.00	0.00	0.00
Obs.	959532	959532	959532	959532	959532
	Panel C: Similarity - Most similar patent				
	Baseline	PhD	No PhD	Local PhD	Non-local PhD
Move x Post	1.1*	1.3**	-0.1	2.7***	-0.4
	(0.7)	(0.6)	(0.6)	(0.7)	(0.7)
Mean Dep.	127.4	108.0	100.2	80.5	91.1
R2 (within)	0.18	0.20	0.21	0.20	0.21
Obs.	604771	604771	604771	604771	604771
	Panel D: Patent class growth- ln(# Patents+1)				
	Baseline	PhD	No PhD	Local PhD	Non-local PhD
Specialization x Post	0.7***	0.7***	0.4*	0.7***	0.5**
	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)
Mean Dep.	8.2	4.5	3.7	1.8	2.6
R2 (within)	0.88	0.83	0.84	0.76	0.78
Obs.	63424	63424	63424	63424	63424

Note: This table shows the results from the three empirical models (patent citations, similarity and patent class growth) discussed in the results sections for the PhD and not PhD subsample. In column (1), we repeat the baseline results. In column (2), we use patent citations of patents filed by inventor with PhD, the patent similarity of patents filed by PhD graduates and the the growth in patent classes by patents of PhD graduates. In column (3), we use the same outcomes for inventors without a PhD. In column (4), we only consider patents of inventors with a PhD degree who graduated from the local university. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

increase in patents with PhD inventors (Panel D). This suggests that PhD graduates react more strongly to new research resulting from an expansion of a university, an indication that they are the primary determinant of the absorptive capacity in the private sector.

5.2 Universities educate specialized PhD graduates

PhD graduates might react more to the new information brought by the newly hired professor because a) they have a higher innate ability, b) they received general science training or c) they received technology-specific training at their university. In general, PhD graduates are likely to represent a positive selection among all university graduates. Thus, the results presented in the previous section might reflect a higher innate ability - better inventors are needed for the more demanding science-based innovation. If graduate education matters over and above innate ability, we would expect that PhD inventors are specialized in the type of technology of their degree granting university.

To see this, we classify all inventors of local patents with respect to whether or not they hold a PhD degree and whether or not they graduated from the local university. Using our data on the university short-lists, Subfigures c) and d) of Figure 8 show that PhD graduates from the local university react more strongly than PhD graduates from other universities. This suggests that at least part of the stronger reaction is due to some kind of technology-specific training received at the local university.

In columns (4) and (5) of Table 3, we repeat our estimations for locally and non-locally trained PhD inventors, respectively. Local PhD graduates are responsible for two thirds of the increase of 0.9 patents that are attributable to PhD graduates in general, while they represent only around 40% of all inventors with a PhD degree. For the outcome variable similarity, we also see a larger effect for local PhD graduates than for non-local PhD graduates. This confirms that local PhD graduates are different in terms of absorptive capacity compared to PhD graduates from other universities and suggests that PhD inventors are influenced in their inventions by new information they acquire from newly hired professors at their alma mater.

5.3 Observations from the universe of German PhD graduates

Our estimates reflect a general pattern in the data for Germany. In the filing year of 2010, around 40% of all patents and 60% of all science-based patents (patents that directly cite an academic article; distance of 1) had at least one inventor with a PhD degree (figure 14 in Appendix C). Therefore, PhD inventors are responsible for the majority of science-based innovation in Germany.

We also see that PhD inventors are specialized in the type of technology of their degree granting university. Figure 9 shows the likelihood that a patent with a German based inventor cites the articles of the closest university, separately for local PhD graduates and for PhD graduates of other universities. On average, local PhD graduates are twice as likely to cite research from their own university than other PhD graduates living at the same location. This is true independent of whether these PhD graduates live close to the university or more than a 100 km away. Thus, holding the number of scientific articles constant, a university's own PhD graduates cite articles produced at this university more than other PhD graduates, independent of the location of the patent invention. This pattern is consistent with graduate education changing the direction of science-based innovation of a graduate.

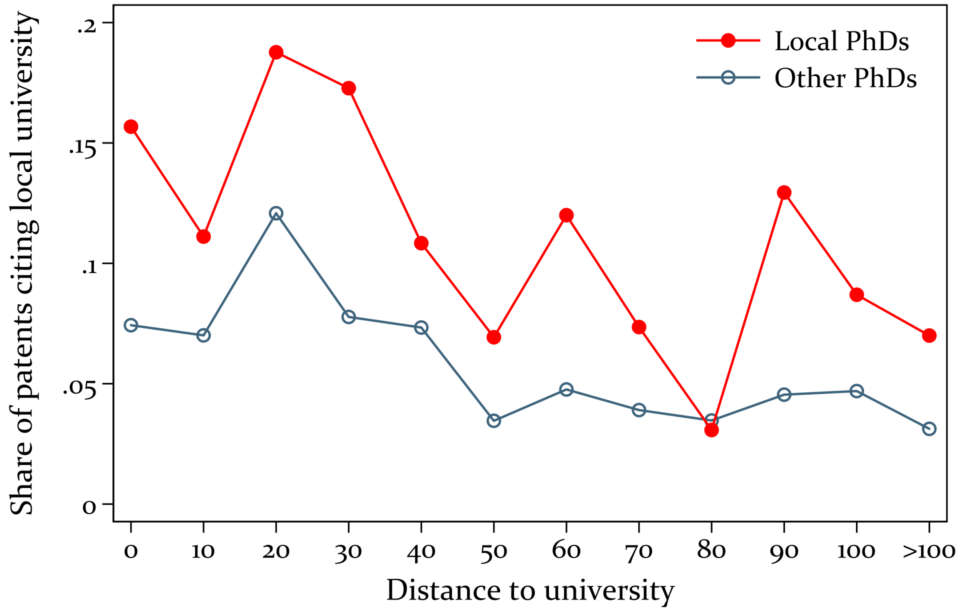
6 Conclusion

In this paper, we study the impact of hiring a new professor on local science-based innovation. Understanding this relationship is important because science-based patents are particularly valuable and because faculty size is a policy measure that can be flexibly be employed to foster the competitive advantage of local firms.

Our identification strategy leverages the legal constraints on university hiring procedures in Germany, which require universities to draw up a short-list of acceptable candidates (runners-up) for every position of a senior professorship. We show that these runners-up are very similar to the moving researchers, which enables us to use them as a close control group.

Our results show that basic research at universities generates value for local private sector innovation. Local inventors build on scientific research produced at universities, and newly hired professors cause local inventors to use new information. The effect is driven by inventors with a (local) PhD degree working in the private sector. Thus, PhD graduates create the absorptive

Figure 9: Share of patents citing local university



Note: This figure shows the share of science-based patents of inventors that cite an article from the closest university seperately for patents of inventors with a PhD from the local university (red line) and with a PhD from another university (blue line). The distance is measured in kilometer.

capacity that helps increase the transfer of technology from the university to the corporate sector.

Taken together, this paper provides evidence that local spillovers from universities to the private sector are significant. This suggests that regional policies and university policies should be determined hand-in-hand to leverage the complementarity of universities and local industry in order to maximize the benefits from university funding.

References

Aghion, P., Boustan, L., Hoxby, C. and Vandenbussche, J. (2009). The causal impact of education on economic growth: evidence from us. *Unpublished paper*.

Agrawal, A., Cockburn, I. and McHale, J. (2006). Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, **6** (5), 571–591.

Ahmadpoor, M. and Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science*, **357** (6351), 583–587.

Andrews, M. (2017). The role of universities in local invention: Evidence from the establishment of us colleges, mimeo.

Arora, A., Belenzon, S. and Sheer, L. (2017). *Back to Basics: Why do Firms Invest in Research?* Tech. rep., National Bureau of Economic Research.

Azoulay, P., Zivin, J. S. and Wang, J. (2010). Superstar extinction. *Quarterly Journal of Economics*, **125** (2), 549–589.

—, Zivin, J. S. G., Li, D. and Sampat, B. N. (2015). *Public R&D investments and private-sector patenting: evidence from NIH funding rules*. Tech. rep., National Bureau of Economic Research.

—, — and Sampat, B. N. (2011). The diffusion of scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine. In *The Rate and Direction of Inventive Activity Revisited*, University of Chicago Press, pp. 107–155.

Belenzon, S. and Schankerman, M. (2013). Spreading the word: Geography, policy, and knowledge spillovers. *Review of Economics and Statistics*, **95** (3), 884–903.

Bikard, M. and Marx, M. (2018). Hubs as lampposts: Academic location and firms' attention to science.

Borjas, G. J. and Doran, K. B. (2012). The collapse of the soviet union and the productivity of american mathematicians*. *The Quarterly Journal of Economics*, **127** (3), 1143–1203.

— and — (2015). Which peers matter? the relative impacts of collaborators, colleagues, and competitors. *Review of Economics and Statistics*, **97** (5), 1104–1117.

Callaert, J., Pellens, M. and Van Looy, B. (2014). Sources of inspiration? making sense of scientific references in patents. *Scientometrics*, **98** (3), 1617–1629.

Cantoni, D. and Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, **129** (2), 823–887.

Cockburn, I. M. and Henderson, R. M. (1998). Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *The Journal of Industrial Economics*, **46** (2), 157–182.

Cohen, W. M. and Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative science quarterly*, pp. 128–152.

Florida, R. (2014). *The Rise of the Creative Class–Revisited: Revised and Expanded*. Basic Books (AZ).

Gaetani, R. and Bergolis, M. L. (2015). The economic effects of scientific shocks.

Hausman, N. (2017). University innovation and local economic growth. *Maurice Falk Institute for Economic Research in Israel. Discussion paper series.*, (5), 0_1–63.

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

Jäger, S. (2017). The economic returns to higher education: Evidence from university openings in germany, mimeo.

Kantor, S. and Whalley, A. (2014). Knowledge spillovers from research universities: evidence from endowment value shocks. *Review of Economics and Statistics*, **96** (1), 171–188.

Kogan, L., Papanikolaou, D., Seru, A. and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, **132** (2), 665–712.

Liu, S. (2015). Spillovers from universities: Evidence from the land-grant program. *Journal of Urban Economics*, **87**, 25–41.

Morrison, G., Riccaboni, M. and Pammolli, F. (2017). Disambiguation of patent inventors and assignees using high-resolution geolocation data. *Scientific Data*, **4**.

Moser, P., Voena, A., Waldinger, F. *et al.* (2014). German jewish emigres and us invention. *American Economic Review*, **104** (10), 3222–55.

Pfister, C., Rinawi, M., Harhoff, D., Backes-Gellner, U. *et al.* (2017). *Regional Effects of Applied Research-Universities of Applied Sciences and Innovation.* Tech. rep., University of Zurich, Department of Business Administration (IBW).

Roach, M. and Cohen, W. M. (2013). Lens or prism? patent citations as a measure of knowledge flows from public research. *Management Science*, **59** (2), 504–525.

Sinha, A., Shen, Z., Song, Y., Ma, H., Eide, D., Hsu, B.-j. P. and Wang, K. (2015). An overview of microsoft academic service (mas) and applications. In *Proceedings of the 24th international conference on world wide web*, ACM, pp. 243–246.

Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L. and Su, Z. (2008). Arnetminer: extraction and mining of academic social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, pp. 990–998.

Valero, A. and Van Reenen, J. (2016). *The economic impact of universities: Evidence from across the globe.* Tech. rep., National Bureau of Economic Research.

Waldinger, F. (2010). Quality matters: The expulsion of professors and the consequences for phd student outcomes in nazi germany. *Journal of Political Economy*, **118** (4), 787–831.

— (2012). Peer effects in science: Evidence from the dismissal of scientists in nazi germany. *Review of Economic Studies*, **79**, 838–861.

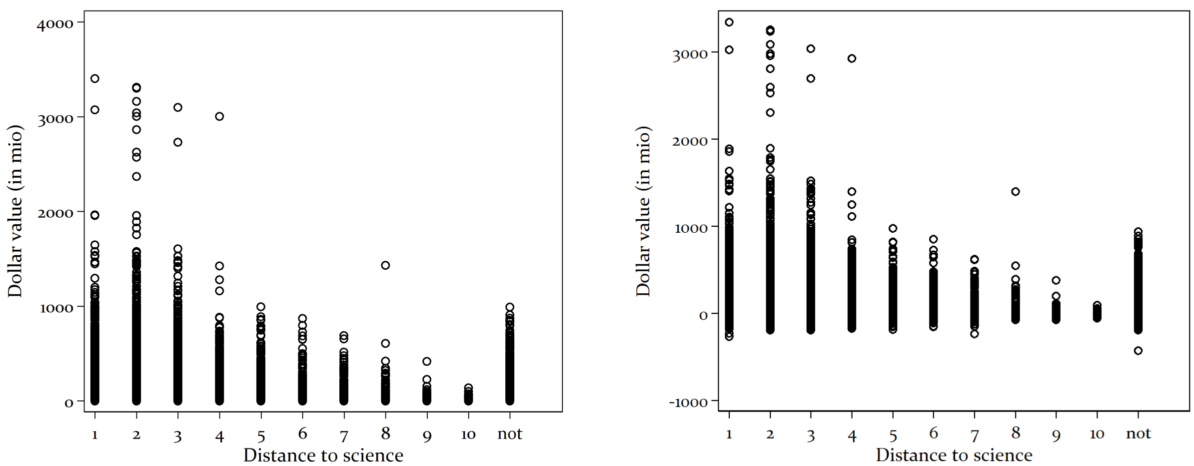
Wissenschaftskonferenz, G. (2014). *Chancengleichheit in Wissenschaft und Forschung: Vierzehnte Fortschreibung des Datenmaterials (2012/2013) zu Frauen in Hochschulen und außerhochschulischen Forschungseinrichtungen.* Gemeinsame Wissenschaftskonferenz GWK, http://www.gwk-bonn.de/fileadmin/Papers/GWK-Heft-40-Chancengleichheit.pdf.

A Appendix to Section 3

A.1 Raw data: Value of a patent and distance to science

Figure 10: Value of patents by distance to science

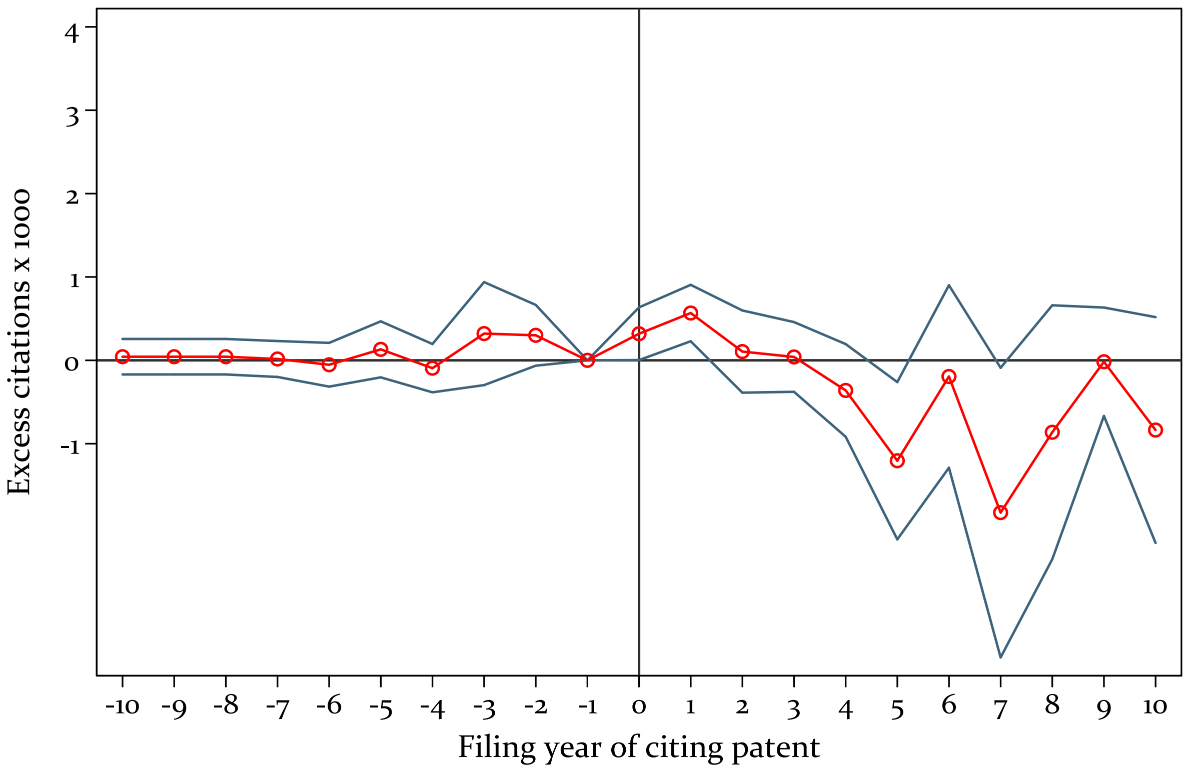
(a) Raw (b) Residualized for filing year x tech class



Note: Subfigure a) shows the raw data for patent values of Kogan *et al.* (2017) with the distance to science data of Ahmadpoor and Jones (2017). The distance to science is defined by citation links. A patent that cites directly an academic article has a distance of D=1. A patent that cites a (D=1)-patent but not an academic article has a distance of D=2. Patents are defined as "Unconnected" if there is no citation link to an academic article. In Subfigure b) we substract from each patent value the average by USPC technology class and filing year.

A.2 Parallel trend between non-movers in patent-to-article citations

Figure 11: Parallel trend between non-movers



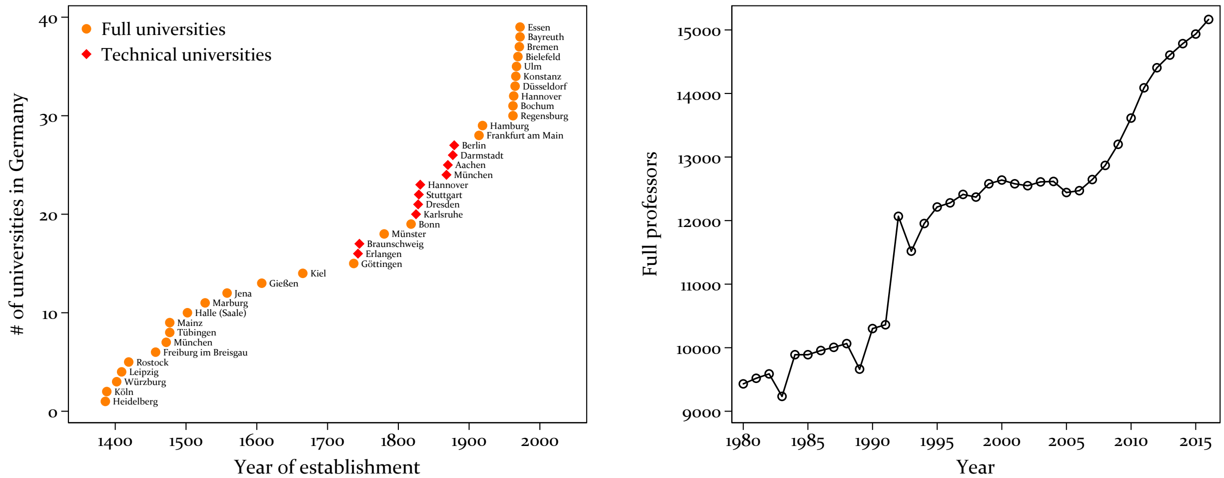
Note: This figure shows the average number of patent-to-article citations for articles of the highest ranked non-moving professor relative to articles of the lower ranked non-moving scientists on the same list. Moving scientists are excluded. Time is counted relative to the move in $t = 0$. Standard errors are clustered on the level of the researcher.

B Appendix to Section 4

Figure 12: Expansion of universities

(a) Few university openings

(b) Large increase in # of professors



Note: Subfigure a) shows the opening years of all German universities included in the Shanghai Ranking. Subfigure b) shows the number of full professorships at German universities over time. The data is from the Destatis, the German Statistical Office and from the German Education Ministery.

B.1 Similarity results

Table 4: Similarity for subsamples — regression results

	Panel A: Mean Similarities				
	(1) Mean All	(2) Mean Top 1	(3) Mean Top 3	(4) Mean Top 5	(5) Mean Top 10
Move x Post	0.1	1.1*	1.3**	0.9*	0.6
	(0.2)	(0.7)	(0.6)	(0.5)	(0.5)
Mean Dep.	73.6	127.8	114.5	106.4	95.7
R2 (within)	0.05	0.18	0.20	0.22	0.22
Obs.	604771	602951	602951	602951	602951
	Panel B: Distance to Science				
	(1) D=1	(2) D=2	(3) D=3	(4) D=4	(5) D>4
Move x Post	2.6***	1.5**	0.0	0.0	-0.4
	(0.8)	(0.7)	(0.6)	(0.6)	(0.5)
Mean Dep.	103.6	95.8	97.7	99.1	111.7
R2 (within)	0.16	0.14	0.17	0.17	0.16
Obs.	468746	357105	369253	359259	544275
	Panel C: Type of Patent				
	(1) Company	(2) University	(3) Others		
Move x Post	1.1	1.5	1.0		
	(0.7)	(1.5)	(0.8)		
Mean Dep.	126.5	85.9	89.6		
R2 (within)	0.18	0.17	0.17		
Obs.	598368	128064	249037		

Note: This table shows the results from a difference-in-differences estimation with ten years before movement as pre-period and ten years after movement as post-period. The estimation equation is:

$$Similarity_{i,k,t} = \beta_1 \cdot Treatment_k + \beta_2 \cdot Treatment_k \cdot Post_t + Year_t \, x \, List \, FE + \epsilon_{i,t}$$

where $Similarity_{i,k,t}$ is the similarity between the abstract of article i and patents filed in year t of scientist k. $Treatment_k$ is an indicator if professor k moved to the university and $Post_t$ is an indicator for all years after the move. As outcome we use in the first line the average similarity of all patents within 100km of the university as outcome. In panel A column (2), the outcome is the similarity of the most similar patent. In columns (3) to (5), we use the 3, 5 and 10 most similar patents. In Panel B column (2), we use the most similar patent among all patents with a distance to science. In columns (3) to (5), we use the subset of patents with different closesness to science. In panel C, we use the most similar patent in different subsets defined by being assigned to companies or individuals (column (1)), universities (column (2)) and all other entitites (column (3)). Standard errors are clustered on the individual level. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

B.2 Growth results

Table 5: Growth of patents for subsamples — regression results

	Panel A: Overall & Type of Patent				
	(1) Local	(2) Non-local	(3) Company	(4) University	(5) Other
Move x Post	0.7***	0.3	0.5**	0.8***	0.7***
	(0.2)	(0.6)	(0.2)	(0.2)	(0.2)
Mean Dep.	8.2	161.9	7.6	0.2	0.4
R2 (within)	0.88	0.94	0.87	0.55	0.65
Obs.	63424	63424	63424	63424	63424
	Panel B: Distance to Science				
	(1) D=1	(2) D=2	(3) D=3	(4) D=4	(5) D>4
Move x Post	1.2***	0.2	-0.3	-0.1	0.2
	(0.3)	(0.2)	(0.2)	(0.1)	(0.2)
Mean Dep.	2.1	2.0	1.5	0.5	2.1
R2 (within)	0.81	0.78	0.77	0.67	0.79
Obs.	63424	63424	63424	63424	63424

Note: This table shows the results from a difference-in-differences estimation with ten years before movement as pre-period and ten years after movement as post-period. The estimation equation is:

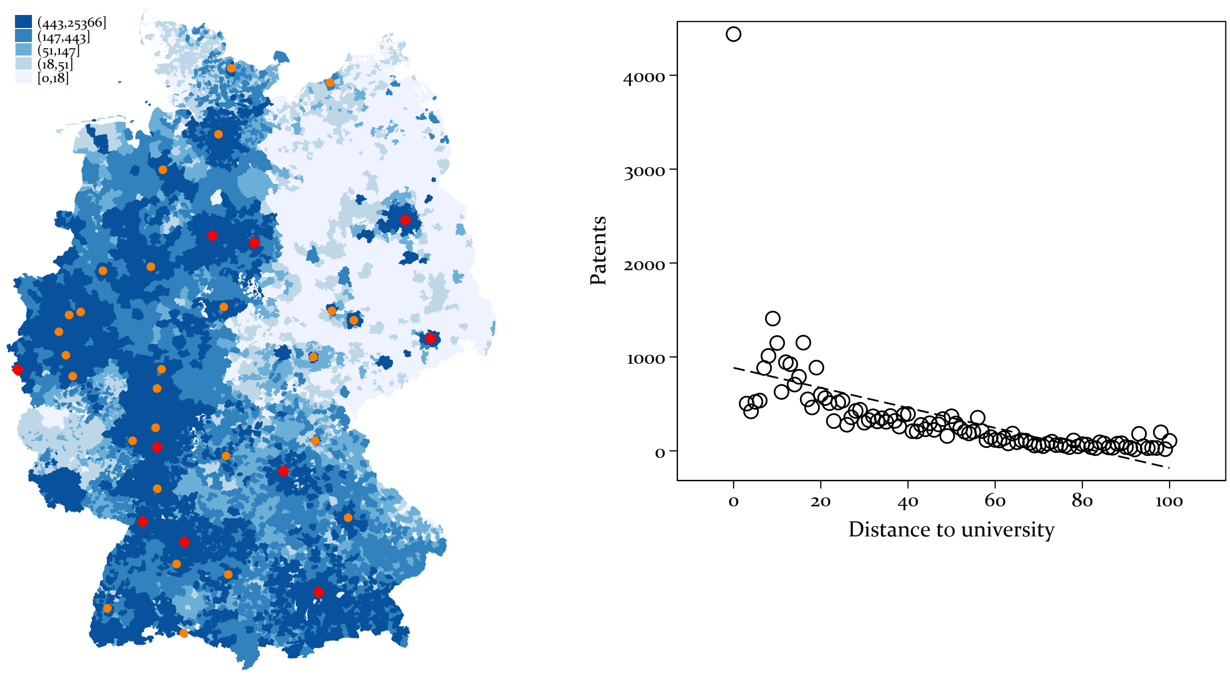
$$ln(\#Patents_{c,t} + 1) = \beta_1 + \beta_2(Prob_{c,i} - Prob_{c,j}) \cdot Post_t + Controls_{i,t} + \varepsilon_{c,t}$$

where $Prob_{c,i}$ is the ex-ante likelihood that the hired scientist i is cited in patent class c. As outcome we use in panel A column (1) only patents with inventors within 100 km of the university. In panel A column (2) we use all patents in the rest of Europe. In columns (3) to (5) of panel A we use patents assigned to companies and individuals, universities and other institutions, respectively. In panel B column (1) we use the number of science-based patents within 100km. In columns (2) to (5) of panel B we use the number of patents with varying distance to science. Standard errors are clustered on the individual level. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

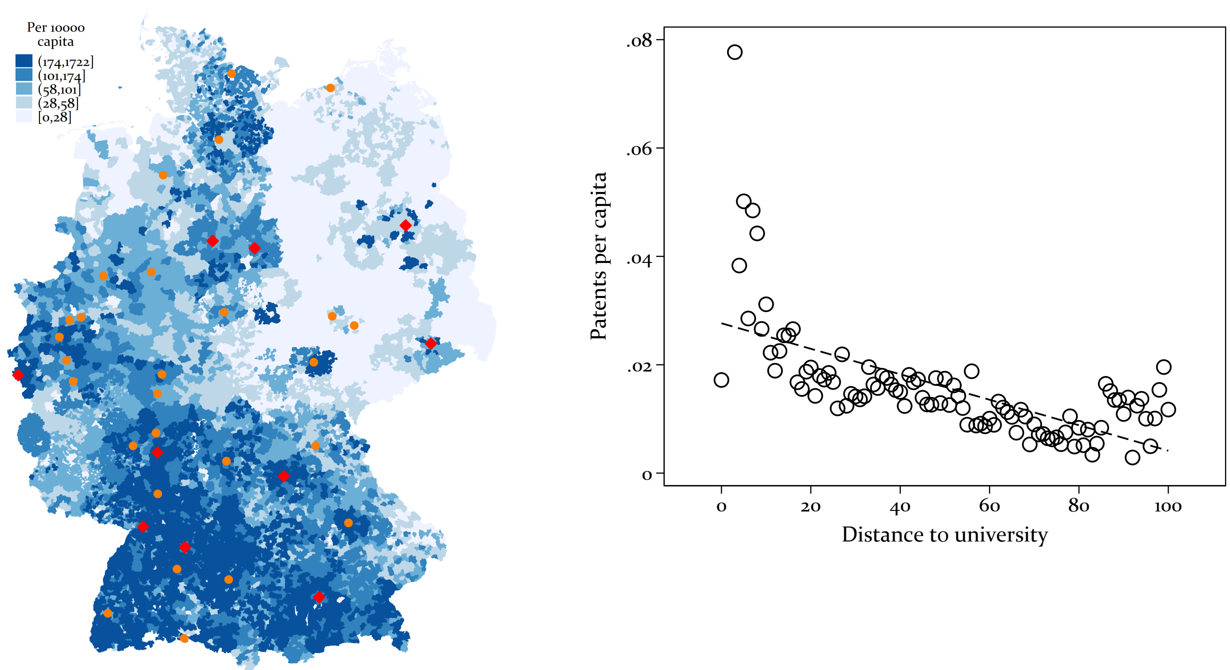
B.3 Patents by municipality

Figure 13: Patenting is clustered around universities

(a) Absolute values



(b) Per capita

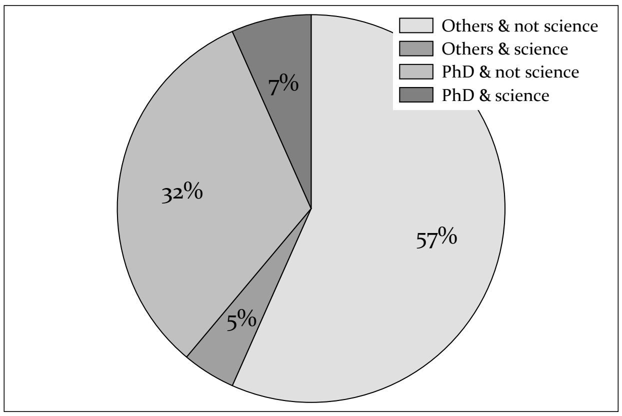


Note: Subfigure a) shows the absolute number of patents by municipality (left panel) and by distance to closest university (right panel). Subfigure b) shows the number of patents per capita by municipality (left panel) and by distance to closest university (right panel). Data is from Morrison *et al.* (2017) and German administrative data of Destatis on municipality borders and population for the year 2016. Red dots are technical universities while yellow dots are full universities. If there are both types of universities in one city, the technical university is shown.

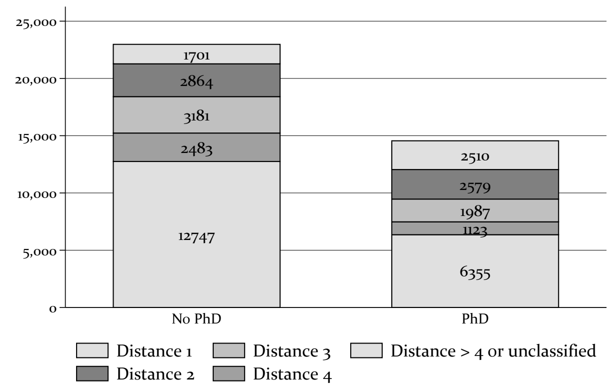
C Appendix to Section 5

Figure 14: Patents by inventors with and without PhD

(a) Shares in 2010



(b) Distance to science in 2010



Note: Subfigure a) shows the share of patents by inventors with and without a PhD. We split the patents in addition by their distance to science, i.e. science based patents that directly cite an academic article and non science-based patents. Subfigure b) splits patents further by their distance to science. The sample are all patents with the publication year 2010. Data on PhD graduates is from the German National Library and patent data is from Patstat.