

# **A Microgeography of Innovation in the City: Location Patterns of Innovative Firms in Berlin**

Christian Rammer<sup>a\*</sup>, Jan Kinne<sup>a</sup> and Knut Blind<sup>b</sup>

a Centre for European Economic Research (ZEW), Dpt. Economics of Innovation and Industrial Dynamics,  
Mannheim, Germany

b Technical University Berlin, Faculty of Economics and Management, Chair of Innovation Economics, and  
Fraunhofer Institute for Open Communication Systems FOKUS, Innovation and Technology Transfer,  
Berlin, Germany

November 2016

## **Abstract**

This paper investigates the micro-location pattern of innovative and non-innovative firms in Berlin using detailed information on the firms' addresses and their local environment. The study employs a unique, representative panel data set of Berlin-based firms from manufacturing and services covering a five-year period (2011-2015) and applying the standard concepts and measurement approaches used in the Community Innovation Surveys. While controlling for firm size, age and sector, we find product innovators and R&D performing firms located closer to research infrastructures, start-ups and other firms from the same industry. They tend to prefer more dynamic neighbourhoods and avoid very densely populated areas. For process innovators, no significant differences from non-process innovators are found. Firms are more likely to introduce new-to-market innovations if other firms in their direct neighbourhood had introduced such innovations in the previous period, but also if firms with such innovations have moved out of their neighbourhood. The 'creative environment' of a firm in terms of bars, cafes, clubs, leisure facilities or cultural locations does not seem to be linked to the innovative activity of firms.

JEL-Classification: O31, O32, O33, R12, R39

Key Words: Microgeography, Innovation, Location Decision, Berlin, Knowledge Spillovers

\* Corresponding author: ZEW, L 7, 1, D-68161 Mannheim  
Phone: +49 621 1235 184  
E-mail: rammer@zew.de

Acknowledgements: We are grateful to the Technical University Berlin and the *Technologiestiftung Berlin* for funding the data collection activity upon which this paper rests.

# 1. Introduction

Innovation activities strongly rely on external knowledge. Innovative ideas of users, advanced technology from suppliers, new knowledge generated in science and research, innovations of competitors, or support from consultants and other service providers are often crucial inputs to innovation (Cassiman and Veugelers 2006, West and Bogers 2014). In order to access this knowledge, firms not only need absorptive capacities and adequate search strategies, but also need to interact with these external sources for innovation. Geographical proximity can certainly facilitate the exchange of knowledge (Jaffe et al. 1993, Thompson 2006, Singh and Marx 2013); particularly if knowledge is tacit and absorbing it requires learning from others (Audretsch and Feldman 1996, Howells 2002).

Innovation also strongly rests on the creativity of a firm's employees. For attracting creative employees and for exploiting their creative potentials, a stimulating local environment may be crucial (Cohendet et al. 2010). The literature on creative industries emphasises the role of dynamic urban spaces and localisation economies (Florida and King 2016). Research on clusters and regional cooperation in innovation has shown that an innovative regional environment stimulates innovation in firms (Saxenian 1994). Proximity to other innovative firms, and to innovative start-ups, can form an innovative milieu in which firms mutually provide inputs for innovation (from learning, observing and interacting).

The role of geographical proximity between actors of an innovation system and regional knowledge spillovers has mostly been analysed on a rather broad geographical scale such as metropolitan areas or administratively defined regional units (see Cooke 2001, 2002). Only recently, more attention has been paid to the local environment of a firm and how the configuration of a firm's neighbourhood affects knowledge exchange and innovation results. A critical issue is the role of neighbourhood configuration and the very local situation on collaboration and knowledge exchange, and resulting inventive and innovative activity (see Catalini 2016, Kabo et al. 2014). Opportunities for meeting and contacting other innovative actors and knowledge sources tend to have a significant impact on innovation performance.

This article aims to contribute to this literature by investigating the microgeography of innovative firms in the German capital of Berlin. Based on a comprehensive firm-level panel data set including exact address information, we explore the local environment of innovative and non-innovative firms in terms of research infrastructure, entrepreneurial activities, cultural and entertainment facilities, and transport infrastructure at a very fine-grid geographical scale. Using the panel nature of our data, we also analyse the impact of innovative and non-innovative firms moving in or out a firm's neighbourhood, or changing their innovative activity, as well as the role of start-ups and closure on a firm's innovation activities. When controlling for size, age and sector, we find that innovative firms are located

in urban environments with a much higher number of other firms and start-ups in close vicinity (less than 250 meters) and with a higher inflow of other firms. Close proximity to research institutes and universities is another distinctive feature of innovative firms. We do not find significant differences in location patterns of innovative opposed to non-innovative firms with respect to proximity to cultural, leisure and entertainment facilities (including bars and restaurants) or public transport. What we do find is a higher price level of office space for innovative firms. If a firm has introduced a product innovation with a high degree of novelty, average rents are 4 to 5% higher. A firm's probability of introducing a new-to-market innovation increases significantly if neighbouring firms had introduced such innovations in the previous period, but also if firms with such innovations have moved out of their neighbourhood.

The innovation concepts and measurement approaches used in this paper rely on the Oslo Manual (OECD and Eurostat 2005) and are fully in line with standard innovation surveys such as the Community Innovation Surveys (CIS) of the European Commission. The database we use is actually an extension of the German contribution to the CIS, including all Berlin-based firms with 5 or more employees in manufacturing and knowledge-intensive services. Address data allow us to link firm-level information with very detailed geographical information on activities and infrastructures located around a firm, applying a scale of geographical detail of less than 10 meters.

The next section of the paper briefly discusses the role of microgeography for firm innovation. Section 3 presents our empirical approach and describes the database. Characteristics of the urban locations of innovative firms derived from a matching analysis with non-innovative firms are shown in Section 4. Section 5 presents the findings of the role of innovation dynamics in a firm's local environment as revealed from changes in innovation activities of neighbouring firms as well as moving in and out of innovative and non-innovative firms. Section 6 concludes.

## **2. The Role of Microgeography in Innovation**

The local environment provides opportunities and limitations to a firm's innovation activities. While modern information and communication technologies and the global integration of economic activities ease the firms' ability to spread their innovation processes beyond geographical boundaries, there are still many local factors that affect innovation performance. Since Marshall (1891) it is well known that the geographic concentration of activities provides positive externalities to firms. These localisation and agglomeration economies are particularly relevant for innovation activities. A pool of highly skilled workers, specialised suppliers, sophisticated demand requiring new solutions, and a well-developed technical, social and research infrastructure are among the localised factors that can drive innovation.

Transport infrastructure and how locations are interlinked with each other is another important factor (Agrawal et al. 2016). The literature on innovative clusters has demonstrated the role of these drivers and how they can mutually reinforce each other to form dynamic regional concentrations of innovative activities (see Saxenian 1994, Glaeser 2000, Porter 1996, Forman et al. 2016, Florida 1995, Feldman and Audretsch 1999, Audretsch 2003). Most of these studies apply a rather broad geographical scale and conceptualise localised innovation enhancing factors at a regional level such as a metropolitan area, a city or a larger region such as the Bay Area or Silicon Valley. Rather few studies took a closer geographic look and stressed the very local dimension of geographical proximity as a driver of innovation.

One tradition of such a ‘microgeography of innovation’ is related to the role of creative urban locations as places of innovation. Jacobs (1961, 1970) has identified dense, diverse, mixed-use and walkable urban areas as a prime generator of innovation. Florida (2002) has stressed the role of such urban environments as the place where a new group of innovative actors — the ‘creative class’ — emerges. Dense, dynamic and diverse urban spaces stimulate innovation by enabling and promoting contacts among innovative actors with different backgrounds and ideas, challenging established approaches and routines and providing ground for new departures. The exact shape of an urban space can hence play a major role, including the co-location of different activities, the proximity to creativity-stimulating places, the availability of sites for communication among different groups of actors, and opportunities for accidentally meeting each other.

Another strand of microgeographic research has been looking at the specific role of local infrastructure devoted to spur innovation and to deliberately provide a place for exchange in innovation processes, such as science parks, technology centres and incubators (Löfsten and Lindelöf 2002, Phan et al. 2005, Sternberg 1990). These localised hubs of innovation bring together innovative firms, researchers and dedicated infrastructure for innovation at one place, providing opportunity for exchanging knowledge and cooperation. Related to this research is the concept of localised knowledge spillovers (Breschi and Lissoni 2001). Access to external knowledge and exchange with other actors are critical elements in the innovation process. The more tacit knowledge is, the more important is face-to-face communication, mutual understanding, a common background and trust. Geographic proximity to knowledge sources can be highly important to effectively source and absorb external knowledge. The role of geographic proximity for knowledge spillovers may also be an indirect one, however, mediated by social networks based on common working experience and mutual trust. As members of such social networks tend to be, but not necessarily need to be, located in close geographic proximity, microgeography may appear to play a critical role (Breschi and Lissoni 2003, Gertler 2003).

The link between local environments and innovation in firms is not a unidirectional one. Firms are not only seeking for innovative environments, their innovation activities will also shape their local environment. They will attract other innovative actors and change institutions and infrastructures in the region. Local environments may hence adapt to the needs of innovative firms. Feldman (2014) stresses that entrepreneurs and other innovative firms are also pivotal as agents of change that can transform local communities. “The character of place—a spirit of authenticity, engagement, and common purpose is the particular feature that differentiates successful places” (Feldman 2014: 10).

While there is ample evidence that geographical proximity can support innovation, the focus on collaboration among local actors and the use of localised knowledge may also lead to lock-in (Boschma 2005). A lack of openness and flexibility as well as a restriction of search strategies within local, well-known and well-established may hinder firms to identify new upcoming trends and routes for innovation that differ from the paths they have taken so far.

The approaches discussed above are using a geographical scale that allows looking into details of places such as neighbourhood configuration, architecture and the walking distance to certain facilities or other actors. However, they usually do not consider exact distances or the layouts of streets and buildings. In recent years, research emerged that is looking in the very microgeographic configurations that may drive innovation. Kabo et al. (2015) use path overlap within an academic research building as a measure of proximity and examine how physical space is shaping the formation and success of scientific collaborations. They find that when two investigators traverse paths with greater overlap, both their propensity to form new collaborations and to win grant funding for their joint work increase. Catalini (2016) shows that researcher co-location within matters for the rate, quality and direction of scientific collaboration. Using data on research labs that were forced to move within a Paris university campus without being able to choose their new location, he found that collaboration between two labs increases significantly if the labs have moved to the same place, as long as the type of research done in both labs is sufficiently similar. Other studies looked at the impact of spatial distance on work performance and found significant effects (Allen 1977, Olson et al. 2002).

While this research mostly focuses on collaboration and performance of researchers and other individuals within the same organisation, there is little microgeographic research on the impact of distance and the spatial configuration of the environment at the firm level. This study attempts to fill this gap by investigating the role of the local environment of innovative and non-innovative firms in an urban space, using Berlin as the place for our empirical analysis. Though we are not able to conduct analysis at the level of individual buildings or offices, detailed address data still allow geographical analysis at a scale of 50-meter distances and below. Hence we can identify the very local environment of a firm, including exact distances to other firms, facilities and infrastructures as well as how the urban space around a

firm is shaped. At the same time, our data include information on innovation activities of a very large sample of firms in manufacturing and knowledge-intensive services in Berlin (covering almost two-thirds of the entire firm population) for a five-year time period. This provides the opportunity to examine how a firm's innovation activities are affected by changes in innovation activities of surrounding firms, including relocations, start-ups and closures. Using these data, we address two research questions:

- Does the configuration of a firm's neighbourhood in terms of diversity and density of activities, and proximity to innovation-related infrastructure relate to innovation performance of firms?
- Does the direct geographical neighbourhood to other innovative firms increase innovation activity?

### 3. Methodology and Data

#### Research methods

When investigating the link between a firm's local environment and the firm's innovation activities, endogeneity problems immediately emerge. Since geography matters for innovation firms will choose those locations that best fit to their innovative strategy and their innovation activities. Innovative firms may hence deliberately locate in close proximity to other innovative firms and will favour 'creative' environments. At the same time, locations may change their characteristics if they host many (or few) innovative firms as infrastructures may adapt to the need of firms, and other firms move in or out, strengthening or weakening a location as a host for innovative firms.

We consider this endogeneity in two ways. First, we refrain from analysing a causal link between location and firm innovation but rather investigate significant differences in the local environment of innovative firms as compared to non-innovative ones. As location patterns will be influenced by some basic firm characteristics such as size, age and sector, we apply a matching approach (Heckman et al. 1998) to ensure that only firms with the same basic characteristics are compared. While matching is usually employed to identify treatment effects of policy intervention, the method is also useful for our purpose. Matching each innovative firm  $i$  in our sample ( $IN_i = 1$ ) with a non-innovative firm ( $IN_i = 0$ ) showing the same basic characteristics  $\mathbf{X}$  allows us to identify difference in location patterns (based on a set  $m$  of location variables  $LOC$ ) and whether innovative firms tend to chose certain places over others. The average difference  $\delta$  for a location variable  $m$  between innovative and non-innovative firms (distinguishing different types  $k$  of innovative activity) is given by ( $t$  indicating the time dimension of our data):

$${}^{km}\delta = E({}^{km}\delta_i) = E({}^mLOC_{it} | \mathbf{X}_{it}, {}^kIN_{it} = 1) - E({}^mLOC_{it} | \mathbf{X}_{it}, {}^kIN_{it} = 0)$$

for  $k \in \{1, \dots, K\}$ , for  $m \in \{1, \dots, M\}$  (1)

The basic firm characteristics  $\mathbf{X}$  used for the matching include firm size, firm age and sector (at the 2-digit level).

The second approach to tackle endogeneity is to analyse whether changes in innovation activities of firms  $j$  in period  $t-1$  located in neighbourhood  $n$  of firm  $i$  ( $i \neq j$ ) affect the focus firm  $i$ 's innovation activities in  $t$ . Such changes would include five types of dynamics:

- transition of firm  $j$  from innovative to non-innovative status between  $t-2$  and  $t-1$  (and vice versa) while firm  $j$  remains located in location  $n$  in  $t-2$ ,  $t-1$  and  $t$ ;
- moving in of an innovative (or non-innovative) firm  $j$  in  $t-1$  into neighbourhood  $n$  (and staying in  $n$  in  $t$ );
- moving out of an innovative (or non-innovative) firm  $j$  in  $t-1$  from neighbourhood  $n$ ;
- foundation of a new (innovative or non-innovative) firm  $j$  in  $t-1$  in neighbourhood  $n$  (and remaining located in  $n$  in  $t$ );
- closure of an innovative (or non-innovative) firm  $j$  in  $t-1$  in neighbourhood  $n$ .

We assume that past innovation dynamics in other firms are exogenous to a focus firm  $i$ 's later innovation activities (i.e. we do not assume that firms  $j$  anticipate future innovation activities of firm  $i$  and react on these expectations by changing their innovation behaviour or their location before firm  $i$  had performed the anticipated innovation activities). At the same time, innovation dynamics in a firm  $i$ 's local environment may alter the firm's opportunities for innovating. If neighbouring firms have newly introduced innovations, innovative firms moved into firm  $i$ 's neighbourhood, or innovative start-ups opened their business in firm  $i$ 's neighbourhood, firm  $i$  may be stimulated by these activities and may learn for its own innovative efforts. At the same time, the loss of an innovative environment due to closure or moving out of innovative firms or stopping of innovative activities in neighbouring firms might discourage innovation in firm  $i$ . In order to test the likely impacts of past innovation dynamics in the local environment on current innovation in firm  $i$ , we run the following regression model:

$${}^kIN_{int} = \alpha + \sum_d {}^k\beta_{1d} {}^kIND1_{djn(t-1)} + \sum_d {}^k\beta_{0d} {}^kIND0_{djn(t-1)} + \sum_l \chi_l CTR_{lnt} + \varepsilon_{it}$$

for  $k \in \{1, \dots, K\}$ ,  $i \neq j$  (2)

$IND$  represents the different types  $d$  of innovation dynamics variables (transition of status, moving in and out, start-up and closure of firms) based on the number of firms reporting a respective dynamics.  $IND1$  represents changes into innovation, or moving in, moving out, start-up and closure of innovative firms, whereas  $IND0$  represents changes out of innovation, or moving in, moving out, start-up and closure of non-innovative firms.  $CTR$  represents control variables  $l$  that may affect firm  $i$ 's innovation decision in  $t$ , such as size, age and

sector.  $\alpha$  is a constant,  $\beta$  and  $\chi$  are parameters to be estimated, and  $\varepsilon$  is the error term. We estimate model (2) both for a firm  $i$ 's innovation status in  $t$  and for a change in innovation status between  $t-1$  and  $t$ .

In both the matching and the regression model approach, we consider different types of innovation indicators  $k$  as we are interested in whether the impact of local innovation dynamics holds for various indicators or is specific to certain types of innovation activities. Relying on well-established indicators for innovation activities and performance at the firm level that are regularly applied in innovation surveys such as the CIS, we distinguish the following indicators:

- introduction of product innovation (either good or service), differentiating by novelty: new-to-market innovation, product-line innovation<sup>1</sup>;
- introduction of process innovation, differentiating by the focus: cost-reducing, quality improving;
- performance of innovation activities, differentiating by the type of R&D activity involved: in-house R&D continuously or occasionally, contracting-out R&D to others (external R&D)

All indicators are binary and capture whether a firm has conducted the respective innovation activity in the reference period of the survey. Such indicators are well suited for measuring innovation in small firms. In small firms, innovation is mostly a yes/no decision whether to carry out an innovation project or not, and whether a firm is able to complete the project by introducing an innovation on the market or within the firm. The chosen indicators are also well suited for capturing innovation both in manufacturing and services. Using quantitative indicators such as R&D expenditure or sales with new products is often less useful as they can be subject to extreme values in small firms and may overrate differences in innovation performance (see Rammer et al. 2009). As about 80% of the firms in our empirical study are small firms with fewer than 50 employees, and about 70% are from service sectors, we believe that the choice of binary indicators is adequate for our study.

## Data

We employ a unique panel data set on innovation activities of Berlin-based firms, the 'Berlin Innovation Panel'. This panel survey has been initiated in 2012 by the Technical University of Berlin and has received funding from the *Technologiestiftung* Berlin. The survey covers all legally independent enterprises with 5 or more employees in manufacturing and B2B services that are headquartered in Berlin. The survey is conducted as part of the German Innovation

---

<sup>1</sup> New-to-market product innovation is an indicator commonly used in innovation surveys. Product-line innovation is less frequently used (and sometimes referred to as "new-to-the-firm" innovation). It is defined as a product innovation that had no predecessor product in the innovating firm, i.e. the innovation opens up a new product line. Product-line innovations may be new-to-the-market, but they may also represent products that are already offered in the market by other firms (see Rammer et al. 2009 for more details).

Survey, which is the German contribution to the Community Innovation Surveys (CIS), and shares all methodological features, including questionnaire design, quality control and data processing routines (see Peters and Rammer 2013 for more details on the Germany Innovation Survey). The Berlin Innovation Panel as well as the German Innovation Survey are conducted by the Centre for European Economic Research (ZEW) as a voluntary mail survey (including an online version of the questionnaire) on an annual base. This paper uses the first five waves of the Berlin Innovation Panel conducted in the years 2012 to 2016 and covering the reference years 2011 to 2015.

The gross sample of the survey includes basically all firms of the target population of the survey (which is around 5,600 firms according to the official business register). In the first year of the survey, the gross sample was 4,927. The sample has been refreshed in 2013 and 2015 in order to compensate for firm closure and firms moving out of Berlin. Panel mortality is substantially high in the Berlin panel, reflecting high dynamics in the urban firm sector. In each survey wave between 14 and 19% of the gross sample represents not utilisable addresses because of closure and relocation. The response rate of the survey is around 20%, which is somewhat lower than the response rate in the German Innovation Survey (which is about 25%), reflecting a lower propensity of survey participation among smaller firms. Following the German Innovation Survey, the Berlin Innovation Panel includes a comprehensive non-response survey which collects information on the presence of product or process innovation as well as in-house R&D activities, applying the same definitions as in the paper/online questionnaire. The non-response survey is based on a stratified sample of non-responding firms and conducted by telephone. The number of firms covered by the non-response survey exceeds the number of responding firms, leading to a high share of firms from the gross sample for which innovation-related information has been collected (between 42 and 47%). Table 1 provides details on the sample size of the Berlin Innovation Panel.

Table 1: Sample size of the Berlin Innovation Panel

Survey year	Gross sample (#)	Not utilisable* (#)	Responses (#)	Non-response survey (#)	Response rate (%) <sup>a</sup>	Response rate (incl. NR) (%) <sup>a</sup>
2012	4,927	908	770	909	19.2	41.8
2013	5,275	914	806	1,101	18.5	43.7
2014	4,886	782	752	997	18.3	42.6
2015	4,810	918	791	1,048	20.3	47.3
2016	4,002	554	707	901	20.5	46.6

\* Firm closure, moving out of Berlin, wrong address due to relocation etc.

a as a percentage of gross sample net of not utilisable addresses.

The high dynamics in the firm sector of Berlin reduces the panel nature of the survey data. In the first five survey years, 7,936 different firms had been surveyed. Only 2,092 firms were

part of the sample in each wave while 1,755 firms were included in only one wave.<sup>2</sup> Consequently, the Berlin Innovation Panel is quite unbalanced. Just 70 firms responded in each wave. This number increases, however, if both responses and information from the non-response survey are taken together (see Table 2). The total number of firm-year observations with information on a firm's innovation activities is 8,782.

Table 2: Responses in the Berlin Innovation Panel by number of panel participations

No. of panel participations	Gross sample (#)	Responses (#)	Non-response survey (#)	Responses or non-response survey (#)
0	-	5,914	5,026	4,171
1	1,755	1,018	1,599	1,359
2	1,626	498	776	917
3	1,366	282	367	691
4	1,097	154	136	474
5	2,092	70	32	324

\* Based on the 3-year reference period of innovation data.

As the unbalanced nature of the survey limits our analysis particularly with respect to measuring innovation activities that take place in a firm's neighbourhood, we try to extend the data set towards a more balanced panel by interpolating and extrapolating observations. For this purpose, we exploit additional data from the Mannheim Enterprise Panel (MEP), which is the sampling pool for the survey.<sup>3</sup> Information on firm foundation and closure is used to exclude firms from the panel if they have not entered the market yet in that year, or if they exited the market in earlier years. Annual address, employment and sector data are used to fill in geographical, employment and sector information for those years a firm has not participated in the survey, but was active in the market. For innovation indicators, inter- and extrapolation is facilitated by the fact that each indicator refers to a three-year reference period, following the common practice of the CIS and the recommendations of the Oslo Manual (OECD and Eurostat 2005). This enables us to fill in innovation data for some years for which the firm did not provide information through the survey.<sup>4</sup> Note that the intra- and extrapolation procedures applied for innovation data do not affect the measurement of our

<sup>2</sup> The last group includes firms which were classified as not utilisable in the first wave they were part of the sample and which were excluded in the next survey wave.

<sup>3</sup> The MEP is a kind of business register based on information collected by Germany's largest credit rating agency, *Creditreform*, and maintained by ZEW. These data also serve as the base for the German firm data in the Bureau-van-Dijk company databases Amadeus and Orbis. The MEP is also the sampling pool of the German Innovation Survey (see Bersch et al. 2014 for details on the MEP).

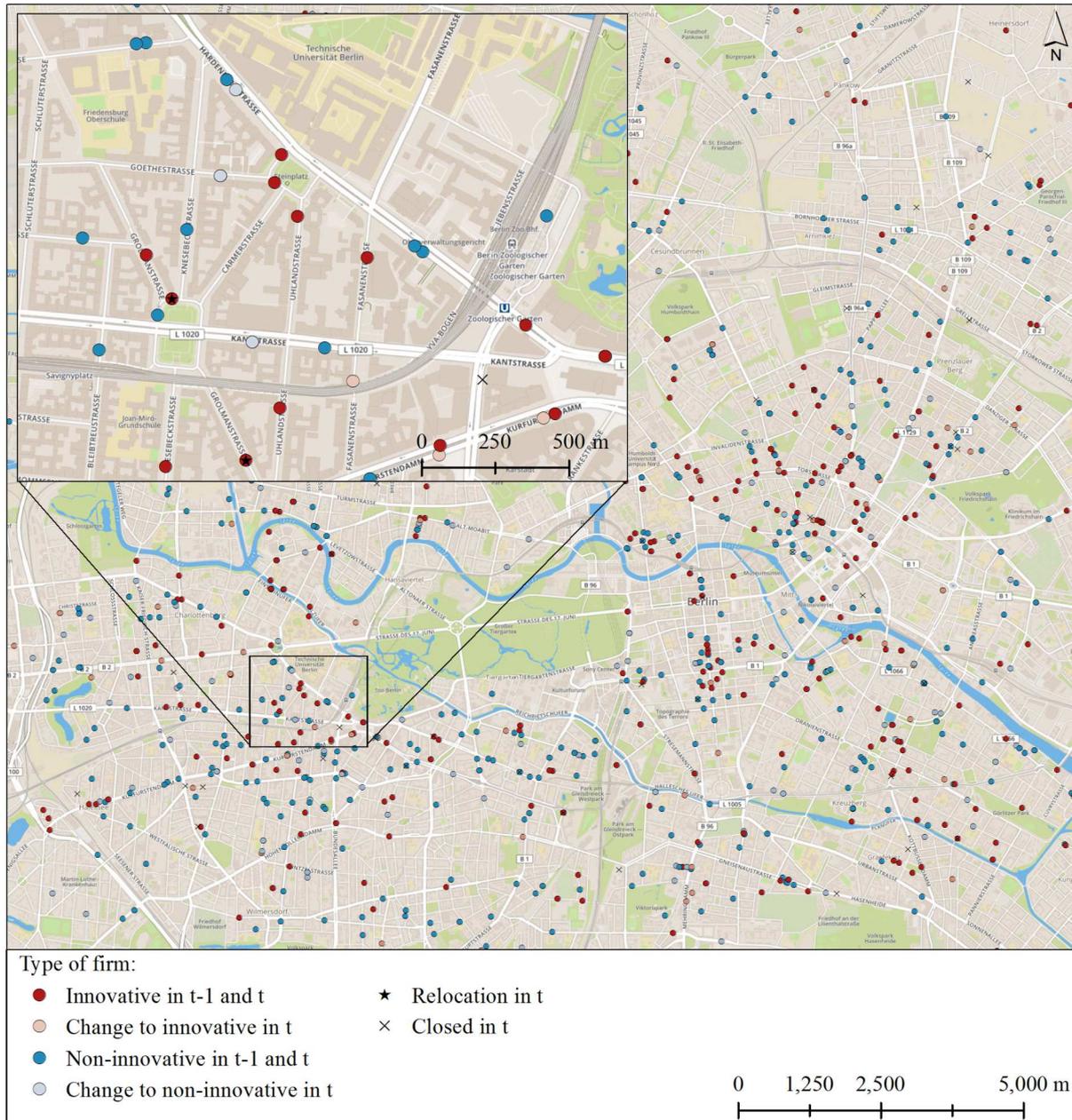
<sup>4</sup> This is particularly true for non-innovating firms, since a 'not innovating' response in year t also refers to years t-1 and t-2, but also for innovating firms if a firm reported innovation in t-1, but no innovation in t, and has missing data in t-2 and t-3. The following table illustrates the intra- and extrapolation procedures applied:

year	original response	intra-/extrapolation									
t	0	0	t	0	0	t	1	1	t	0	0
t-1	.	0	t-1	1	1	t-1	.	1	t-1	.	0
t-2	.	0	t-2	.	1	t-2	.	1	t-2	1	1
t-3	.	.	t-3	.	1	t-3	1	1	t-3	.	.

innovation dynamics variables since a change in the innovation status of a firm can only be observed based on original information.

In that way, we extended the total number of year-firm observations by 53% to 13,405. The more balanced panel contains 3,723 different firms. The average number of observations per firm is 3.6. For 1,511 firms, full information for each of the five years is available. Figure 1 illustrates the geographical distribution of firms in the balanced sample by innovation status for the central area of Berlin.

Figure 1: Example for the geographic distribution of firms in Berlin by innovation status



Basemap: MapBox/OSM

For each firm in the Berlin Innovation Panel, exact address information is available for each year. In order to describe the local environment of a firm, the following indicators are used. Details of the location indicators are provided in Table 3.

- Science & education: number of students in universities and number of researchers in research institutes. For both universities and research institutes, data have been collected for individual locations (buildings);
- Entertainment: number of bars, restaurants, cultural and leisure facilities;
- Transport: walking distance to the nearest public transport stop (underground, buses, regional trains), driving distance to the nearest international airport and main-line railway station;
- Business: number of firms active in the same 2-digit sector as firm  $i$ , number of firms that moved into firm  $i$ 's neighbourhood in  $t-1$ , number of firms newly started in firm  $i$ 's neighbourhood in  $t-1$ ;
- Density: population per unit of area;
- Real estate market: prices and price dynamics for commercial space and for residential space;
- Innovation dynamics: the five indicators described above.

Table 3: Location indicators

<i>Group</i>	<i>Indicator</i>	<i>Short name</i>	<i>Measured by</i>	<i>Unit</i>
Science & education	Universities	uni_[d]	Distance threshold	No. of students
	Research institutes	ins_[d]	Distance threshold	No. of researchers
Entertainment	Bars and restaurants	bars_[d]	Distance threshold	No. of bars/restaur.
	Cultural facilities	cult_[d]	Distance threshold	No. of facilities
	Leisure facilities	leis_[d]	Distance threshold	No. of facilities
Transport	Public transport	pubtr	Nearest feature	Distance in meters
	International airport	air	Nearest feature	Distance in meters
	Main-line railway station	train	Nearest feature	Distance in meters
Business activity	Stock of firms in same sector	fst_[d]	Distance threshold	No. of firms
	Firms moving in	fin_[d]	Distance threshold	No. of firms
	Start-up activity	st_[d]	Distance threshold	No. of firms
Density	Population density	popden	Local value	Inhabitants per km <sup>2</sup>
Real estate market	Commercial rent	comm	Local value	€ per m <sup>2</sup>
	Commercial rent change	comm_ch	Local value	Delta € per m <sup>2</sup> $t-1/t$
	Residential rent	resdt	Local value	€ per m <sup>2</sup>
	Residential rent change	resdt_ch	Local value	Delta € per m <sup>2</sup> $t-1/t$
Innovation dynamics	Incoming innovators/non-inn.	in[k]1_[d]/in[k]0_[d]	Distance threshold*	No. of firms
	Outgoing innovators/non-inn.	ot[k]1_[d]/ot[k]0_[d]	Distance threshold*	No. of firms
	Transition into/out of innov.	ch[k]1_[d]/ch[k]0_[d]	Distance threshold*	No. of firms
	Innovative/non-inn. start-ups	nf[k]1_[d]/nf[k]0_[d]	Distance threshold*	No. of firms
	Closing innovators/non-inn.	cl[k]1_[d]/cl[k]0_[d]	Distance threshold*	No. of firms

[d]: alternative distance thresholds: 50m, 100m, 250m 500m, 1000m, 2500m; [k]: type  $k$  of innovation activity

\* for innovation dynamics, only the 250m threshold is used in model estimations presented in this paper as this threshold provided the most robust results.

Data sources and measurement details for each location indicator are presented in the Appendix. For the indicators on science & education, diversity, business and innovation dynamics we use alternative distance thresholds for defining a firm  $i$ 's neighbourhood. The

thresholds are 50, 100, 250, 500, 1000 and 2500 meters, measured from the entry door of the building in which a firm is located. For some indicator groups (science & education, transport, density, real estate market), data beyond the border of the City of Berlin have been included. For other groups (entertainment, business activity, innovation dynamics), location indicators refer to the territory of Berlin. For the City of Berlin this territorial confinement is of little harm since Berlin—in contrast to most other large cities—has very few urbanised areas outside its administrative borders. Commercial activities in Berlin are very much concentrated in the more central parts of the city, and suburbanisation is a process that has not touched Berlin to a larger extent yet. This idiosyncratic development reflects the special situation of Berlin until 1990 as the city was divided, and the Western part was an isolated territory with no connection to its surroundings. After the unification, real estate prices remained very low even in the most central parts of the city, contributing to a concentration process rather than suburbanisation until the recent years.

Table 7 in the Appendix shows descriptive statistics for all variables used in our analysis.

#### **4. Location Characteristics of Innovative Firms**

The results of the matching analysis reveal significant differences in the local environment of innovative and non-innovative firms in Berlin. These differences vary by type of innovation activity. Table 4 presents the results for three rather broad definitions of innovation: whether a firm has introduced any new product or new process ('innovator'), whether a firm has introduced a new product ('product innovator') and whether a firm has introduced a new process ('process innovator'). For all three types of innovators we find statistically significant positive differences to non-innovators for the proximity to research institutes. The largest relative differences are found for the number of researchers in research institutes located very close to the firm (less than 50 m). An innovator on average has 3 researchers from public research institutes located within a 50-meter radius, which is 79 per cent more than for non-innovators. The relative difference is monotonously declining with increasing distance. The proximity to research institutes is significant until a distance of about 1 km (except process innovators: 0.5 km). For universities, we also find a significant positive difference. This result does not hold for process innovators, however. Interestingly, only for a distance of 0.5 to 1 km the number of university students in a firm's neighbourhood differs significantly between innovators and non-innovators. We would explain this result by the fact that most university buildings are rather huge and often located next to each other, leaving little space for other firms to locate nearby, except for firms that provide direct services to the university or to students (e.g. printing shops, facility management). Firms looking for proximity to universities will have to choose locations that are quite distant from the main university buildings.

Table 4: Matching results for product/process innovators

Variable	Innovator			Product innovator			Process innovator		
	mean	$\delta$	t value	mean	$\delta$	t value	mean	$\delta$	t value
uni_50	11	-21	-0.28	11	22	0.38	17	52	1.09
uni_100	21	-1	-0.01	22	24	0.56	21	4	0.07
uni_250	147	16	0.75	164	19	0.93	159	16	0.70
uni_500	982	25	2.99 *	1068	27	3.22 *	897	8	0.74
uni_1000	2760	14	2.51 *	2928	17	3.03 *	2622	4	0.55
uni_2500	12188	-2	-0.53	12437	2	0.48	12086	-3	-0.66
ins_50	3.3	79	4.86 *	3.7	76	4.36 *	3.1	70	3.65 *
ins_100	8.8	55	4.49 *	9.3	52	3.76 *	7.8	30	1.78
ins_250	48	40	4.35 *	53	46	5.00 *	47	28	2.68 *
ins_500	117	31	4.73 *	128	38	5.76 *	114	18	2.35 *
ins_1000	286	17	3.06 *	305	22	4.00 *	280	8	1.22
ins_2500	1192	0	0.10	1219	5	1.45	1158	-7	-1.72
bars_50	0.7	1	0.14	0.7	4	0.70	0.7	-1	-0.20
bars_100	2.1	-2	-0.38	2.1	2	0.33	2.1	-4	-0.65
bars_250	11	-3	-0.63	11	0	-0.02	11	-1	-0.30
bars_500	41	-3	-0.86	41	0	0.01	41	-5	-1.36
bars_1000	143	-3	-1.05	144	-1	-0.30	143	-6	-1.66
bars_2500	710	-2	-0.78	715	0	-0.10	713	-4	-1.29
cult_50	0.0	-26	-1.37	0.0	-36	-1.91	0.0	-28	-1.32
cult_100	0.1	-14	-1.11	0.1	-23	-1.73	0.1	-11	-0.84
cult_250	0.7	-2	-0.34	0.7	-2	-0.34	0.7	3	0.40
cult_500	2.6	-3	-0.53	2.6	-1	-0.26	2.6	0	0.03
cult_1000	9.2	-4	-0.98	9.3	-2	-0.52	9.3	-3	-0.68
cult_2500	47	-3	-1.05	47	-3	-0.82	47	-4	-1.06
leis_50	0.0	0	0.00	0.0	-34	-1.33	0.0	39	2.01 *
leis_100	0.1	5	0.48	0.1	1	0.08	0.1	9	0.80
leis_250	0.8	-5	-0.95	0.8	-8	-1.52	0.8	-4	-0.75
leis_500	3.4	-2	-0.72	3.5	-1	-0.28	3.4	-6	-1.58
leis_1000	14	-1	-0.36	14	-1	-0.28	14	-2	-0.85
leis_2500	76	-2	-0.87	76	-2	-0.94	77	-2	-1.01
pubtr	155	-3	-1.08	158	0	0.07	153	-4	-1.44
air	9080	-1	-1.27	9009	-3	-2.49 *	9164	1	1.00
train	3887	2	1.18	3911	4	2.07 *	3925	6	2.75 *
fst_50	0.5	-1	-0.16	0.5	11	1.39	0.4	-17	-1.71
fst_100	0.7	1	0.12	0.7	12	1.77	0.6	-11	-1.30
fst_250	1.7	7	1.51	1.8	14	2.94 *	1.7	4	0.67
fst_500	4.0	3	0.63	4.2	6	1.42	3.9	3	0.50
fst_1000	11	0	-0.07	11	3	0.64	11	-2	-0.44
fst_2500	43	-8	-2.10 *	44	-9	-2.08 *	42	-8	-1.89
fin_50	2.6	17	1.81	2.8	19	2.00 *	2.5	-9	-0.77
fin_100	4.6	13	1.96 *	4.8	15	2.25 *	4.4	-5	-0.67
fin_250	15	4	0.74	16	9	1.67	15	0	0.02
fin_500	44	-7	-1.34	45	1	0.15	44	-6	-1.16
fin_1000	140	-8	-1.98 *	144	0	-0.02	139	-9	-2.03 *
fin_2500	646	-5	-1.56	657	-1	-0.19	644	-6	-1.89
st_50	0.2	12	1.28	0.2	13	1.35	0.2	1	0.05
st_100	0.4	14	2.12 *	0.5	14	1.98 *	0.4	4	0.53
st_250	1.7	5	0.89	1.7	9	1.62	1.7	5	0.83
st_500	5.2	-1	-0.11	5.3	1	0.28	5.3	3	0.51
st_1000	17	0	-0.06	18	1	0.18	18	0	-0.09
st_2500	80	-2	-0.52	80	-1	-0.25	80	-1	-0.19
popden	6396	-5	-2.01 *	6365	-5	-1.86	6368	-6	-2.21 *
comm	11	0	0.15	11	1	0.96	11	1	0.67
comm_ch	0.3	-12	-0.62	0.3	-11	-0.53	0.3	-14	-0.68
resdt	8.5	-1	-1.68	8.5	-1	-1.19	8.6	-1	-0.93
resdt_ch	0.4	-12	-1.68	0.4	-11	-1.51	0.4	-14	-1.75
# obs. <sup>a</sup>	4,119 / 5,924			3,305 / 6,778			2,558 / 7,584		

\* p<0.05.  $\delta$ : difference to control group in %. a: number of innovators / no. of non-innovators with common support.

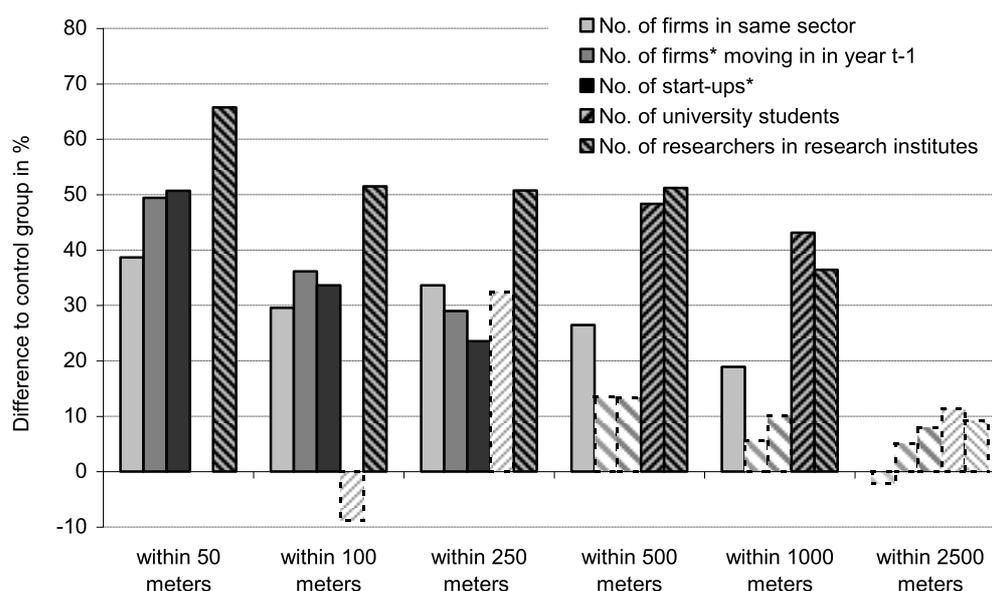
For our measures of entertainment facilities (bars, restaurants, cultural and leisure facilities), we find virtually no significant differences between innovators and non-innovators. The same holds for public transport and for our real estate price indicators. What we do find is a positive difference for the proximity to main line railway stations. Both product innovators and process innovators are located closer to such interregional transport hubs, though the relative difference is small (just 4 to 6 per cent less distance on average).

Another significant difference relates to the proximity to other firms (from manufacturing and knowledge-intensive services) that recently have moved into a firm's neighbourhood from another location (either in Berlin or elsewhere). A product innovator has on average about 2.8 new neighbours in a 50m radius, and 4.8 in a 100m radius. This is 19 and 15 per cent more, respectively, than for firms with no product innovations. Product innovators also show a larger number of recent start-ups in their neighbourhood, exciding the number of start-ups of firms without product innovation by 14 per cent for a 100m radius. We do not find higher local firm dynamics for process innovators. In addition, we find some indication of localisation economies for product innovators as the number of firms from the same sector of the innovating firm within a 250m radius is significantly higher than for non-innovating firms. For smaller radii differences are not highly significant. Again, we do not find similar results for process innovators.

Finally, innovators tend to be located in slightly less densely populated areas. The average population density of an innovator in the 100m quadrant around the firm's location is about 6,400 inhabitants per km<sup>2</sup>, which is 5 per cent lower than for non-innovators. One source for the difference may relate to the fact that innovators tend to locate closer to research institutes and other firms.

When looking at the novelty of product innovations (market novelties, product-line novelties) and the type of impact of process innovation (cost-reducing, quality-improving), the results for product and process innovators are largely confirmed (see Table 8 in the Appendix). A main difference relates to localisation economies, however. Both firms with market novelties and product-line novelties are located in much closer proximity to other firms from their sector as compared to similar firms without these types of product innovation. Figure 2 illustrates these differences with respect to market novelties. The number of firms in the same sector located within a 50m radius is 39 per cent higher. The number of firms that have moved into the direct neighbourhood of a firm with market novelties is 49 per cent higher for a 50m radius and 36 per cent higher for a 100m radius. The same holds for the number of start-ups. Significant positive differences can be observed up to 250m for firms moving in and for start-ups and up to 1 km for firms in the same sector. For the proximity to research institutes and universities we see similar findings as for all product innovators.

Figure 2: Difference in location variables between firms with and without market novelties, by distance thresholds



\* in manufacturing and knowledge-intensive services

Bars shown in grey hatching and dotted frame represent insignificant results.

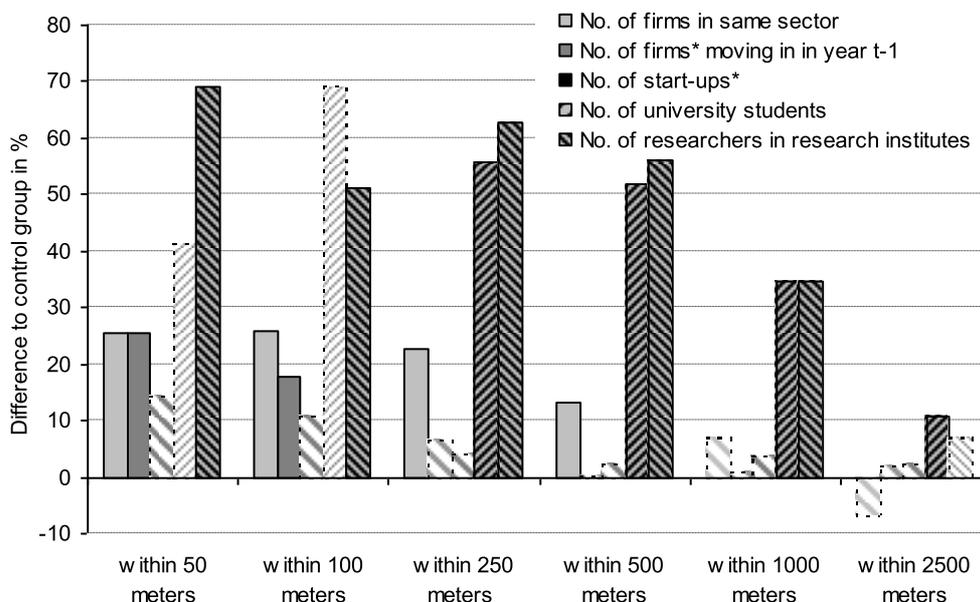
Another interesting result for firms with market novelties (which also holds for firms with product-line novelties) is the higher average rent for commercial space that is paid in their neighbourhood. The difference to the rent level of other firms is 5 per cent (4 per cent for firms with product-line novelties). The mark-up is equal to 0.6 € per m<sup>2</sup>. This may either indicate that firms heading for real novelties in their innovation strategy look for more prestigious and modern places, which imply higher rents. On the other hand, landlords may anticipate the higher innovation performance of their renters and try to appropriate a part of the (expected) higher returns usually resulting from such innovations.

The two types of process innovators—cost reducing and quality improving—show similar location patterns as all process innovators. In line with the findings for other innovative firms, they are located closer to research institutes and universities, though proximity to university seems to be less a location feature than for product innovators.

For firms conducting in-house R&D on a continuous basis we find similar location patterns as for firms with market novelties. However, firms conducting R&D on an occasional basis show a different location pattern as we do not find any differences in the proximity to firms from the same sector or in incoming firms or start-ups (see Table 9 in the Appendix). A firm's choice for the way it organises its in-house R&D activities hence seems to relate quite strongly to the local environment. We find a significant positive difference for continuously R&D performing firms with respect to the number of firms in the same sector up to a 500m radius (Figure 3). For the number of ingoing firms, significant effects are confined up to a 100m radius. In addition, proximity to research institutes is much higher for continuously R&D performing firms. They also tend to be located closer to university campuses than any

other type of innovation active firm as the number of students in a 250m radius is significantly higher while we find significant differences for university students only from a 500m distance threshold for other types of innovation activities. Proximity to start-ups is not higher for firms with continuous R&D.

Figure 3: Difference in location variables between firms with and without continuous in-house R&D activities, by distance thresholds



\* in manufacturing and knowledge-intensive services

Bars shown in grey hatching and dotted frame represent insignificant results.

For firms contracting-out R&D to other firms or organisations, a similar location pattern as for continuously R&D performing firms emerges, though proximity to firms from the same sector is less distinctive. This similarity may relate to the fact that in-house R&D and external R&D tend to be complementary rather than substitutive (Cassiman and Veugelers 2006). For firms contracting-out R&D we also find some significant positive differences with respect to entertainment variables. These firms tend to be located closer to leisure facilities than other firms when looking at a 100m and 250m radius.

## 5. Effects of Innovation Dynamics in a Firm's Neighbourhood

The second part of our empirical analyses investigates the role of changes in the innovation activities of other firms located close to a focal firm. These changes refer to starting or stopping a certain innovation activity by firms already located in the area or moving in or out of innovative or non-innovative firms (including firm closure and start-ups). We assume that a firm's innovation activities of type  $k$  are particularly affected if changes in the same type of innovation activity take place. We therefore measure the innovation dynamics variables

specific to each type  $k$  of innovation activity. In order to limit the variety of results, we concentrate on six types: introduction of any innovation (product or process), introduction of product innovation, introduction of process innovation, introduction of market novelties, introduction of product-line novelties, and performance of in-house R&D. We also restrict the definition of a firm's neighbourhood to a 250m radius and refrain from applying alternative thresholds since the analysis in the previous section has shown that the 250m distance threshold is often a demarcation between significant and insignificant differences in location patterns.

We employ three dependent variables when estimating model (2): (a) the probability of a firm  $i$  to conduct an innovation activity of type  $k$  in year  $t$  (irrespective whether firm  $i$  has conducted the same innovation activity in year  $t-1$ ), (b) the probability of a firm  $i$  to enter into innovation activity  $k$  in year  $t$  (i.e. firm  $i$  did not conduct this activity in  $t-1$ ), and (c) the probability of a firm  $i$  to stop innovation activity  $k$  in year  $t$  (while having conducted this activity in  $t-1$ ). All independent variables are measured for period  $t-1$  (i.e. a change in the innovation status of a neighbouring firm has occurred between  $t-2$  and  $t-1$ ). As (b) and (c) are rather rare events, which also applies to some of our innovation dynamics variables (closure, start-ups, moving in and moving out of firms with a particular innovation activity  $k$ ), we are facing data problems when estimating model (2) due to perfect correlation between the dependent and some of our innovation dynamics variables. To avoid this problem, we aggregate the ten innovation dynamics variables ( $in1$ ,  $in0$ ,  $ot1$ ,  $ot0$ ,  $ch1$ ,  $ch0$ ,  $cl1$ ,  $cl0$ ,  $nf1$ ,  $nf0$ ) to four variables:  $into1$  represents the number of firms that entered into innovation activity  $k$  in neighbourhood  $n$  in  $t-1$  (sum of  $in1$ ,  $ch1$ ,  $nf1$ ),  $into0$  represents the respective number for non-innovative firms (sum of  $in0$ ,  $ch0$ ,  $nf0$ ),  $outo1$  represents the number for firms exiting innovation activity  $k$  in neighbourhood  $n$  in  $t-1$  (sum of  $ot1$ ,  $cl1$ ), and  $outo0$  represents the respective number for non-innovative firms (sum of  $ot0$ ,  $cl0$ ).

Table 5 contains the results of probit regression models for six types of innovation activities. The first part of the table reports results for the model with all ten innovation dynamics variables; the second part reports the results when using the four aggregate variables only. The results for the disaggregate innovation dynamics variables show a higher propensity to introduce product innovation if other firms in the neighbourhood have changed their product innovation activity in the year before. The positive influence of local innovation dynamics holds both for starting and stopping product innovation in neighbouring firms. The marginal effect is rather low: 1.2 percentage points for 'positive' innovation dynamics and 2.1 percentage points for 'negative' dynamics, compared to an average share of product innovators in the sample of 31.6%. The results for the aggregated innovation dynamics variables reveal that moving in of firms with no product innovation or stopping of product innovation activities by existing firms stimulates product innovation in the following period. The marginal effect is again low (1.7 percentage points increase).

Table 5: Estimation results of probit models on the impact of innovation dynamics in a firm's neighbourhood on the firm's innovation activities

a) All innovation dynamics variables

	Innovation		Product Innovation		Market novelty		Product-line novelty		Process innovation		Continuous R&D	
	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.
<i>Innovation dynamics in t-1</i>												
in1	-0.014	0.020	-0.021	0.022	0.013	0.026	-0.008	0.024	0.008	0.022	-0.006	0.019
in0	0.025	0.020	0.008	0.017	-0.003	0.006	0.000	0.008	0.012	0.013	-0.027	0.013 *
ot1	0.013	0.019	0.001	0.020	0.038	0.016 *	0.012	0.018	-0.017	0.022	0.014	0.016
ot0	0.032	0.021	0.025	0.017	0.007	0.005	-0.004	0.008	0.019	0.013	0.019	0.013
ch1	0.010	0.005	0.012	0.005 *	0.009	0.005 *	-0.003	0.007	0.014	0.005 *	0.023	0.011 *
ch0	0.012	0.008	0.021	0.008 *	-0.001	0.005	0.006	0.006	0.007	0.006	0.021	0.008 *
cl1	-0.028	0.034	0.013	0.035	0.066	0.025 *	-0.031	0.026	-0.027	0.037	0.028	0.052
cl0	-0.024	0.035	-0.004	0.029	0.015	0.011	0.026	0.013 *	-0.019	0.025	0.000	0.022
nf1	0.133	0.143	-0.036	0.154	a)		a)		0.321	0.216	0.053	0.054
nf0	-0.032	0.040	0.014	0.034	0.004	0.013	0.005	0.016	-0.022	0.029	0.022	0.026
lna	-0.024	0.008 *	-0.023	0.008 *	0.009	0.002 *	0.014	0.002 *	-0.016	0.007 *	-0.024	0.007 *
lnb	0.067	0.005 *	0.053	0.004 *	0.148	0.082 *	0.056	0.031 *	0.059	0.004 *	0.044	0.004 *
LR Chi <sup>2</sup>	1,015.5		1,103.5		372.9		377.6		467.5		1,209.8	
Pseudo R <sup>2</sup>	0.112		0.131		0.168		0.143		0.062		0.179	
# observ.	6,736		6,742		4,915		4,905		6,754		6,311	

b) Aggregated innovation dynamics variables

	Innovation		Product Innovation		Market novelty		Product-line novelty		Process innovation		Continuous R&D	
	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.
<i>Innovation dynamics in t-1</i>												
into1	0.008	0.005	0.009	0.005	0.010	0.004 *	-0.003	0.007	0.013	0.005 *	0.019	0.009 *
into0	0.011	0.007	0.017	0.007 *	-0.001	0.003	0.004	0.004	0.007	0.005	0.007	0.006
outo1	0.003	0.017	0.003	0.017	0.044	0.014 *	0.002	0.017	-0.019	0.019	0.009	0.016
outo0	0.019	0.017	0.012	0.014	0.008	0.005	0.004	0.006	0.012	0.011	0.009	0.011
lna	-0.024	0.008 *	-0.023	0.008 *	-0.001	0.004	-0.002	0.004	-0.017	0.007 *	-0.025	0.007 *
lnb	0.067	0.005 *	0.053	0.004 *	0.009	0.002 *	0.014	0.002 *	0.058	0.004 *	0.044	0.004 *
LR Chi <sup>2</sup>	1,008.8		1,099.4		371.2		372.6		462.7		1,196.3	
Pseudo R <sup>2</sup>	0.111		0.131		0.168		0.141		0.061		0.177	
# observ.	6,736		6,742		4,915		4,905		6,754		6,311	

a) Variable omitted due to perfect correlation with dependent variable.

\* p<0.05

Table 6: Estimation results of probit models on the impact of innovation dynamics in a firm's neighbourhood on changes of the firm's innovation activities

a) Dependent variable: starting innovation

	Innovation		Product Innovation		Market novelty		Product-line novelty		Process innovation		Continuous R&D	
	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.
<i>Innovation dynamics in t-1</i>												
into1	-0.002	0.003	-0.004	0.003	0.001	0.001	0.001	0.003	0.002	0.003	0.001	0.003
into0	-0.005	0.004	-0.003	0.003	-0.001	0.001	-0.004	0.002	0.000	0.003	-0.002	0.003
outo1	0.015	0.008	0.014	0.008	0.003	0.003	-0.005	0.010	0.003	0.010	-0.005	0.007
outo0	0.013	0.008	0.008	0.007	0.002	0.001 *	0.000	0.003	0.014	0.006 *	0.001	0.004
lna	0.006	0.004	0.005	0.004	0.002	0.001	0.001	0.002	0.005	0.004	0.002	0.003
lnb	0.000	0.003	0.001	0.002	0.000	0.000	0.001	0.001	0.004	0.002	0.002	0.002
LR Chi <sup>2</sup>	27.0		32.5		53.2		42.6		44.3		48.6	
Pseudo R <sup>2</sup>	0.008		0.011		0.094		0.055		0.014		0.027	
# observ.	6,356		6,422		4,836		4,806		6,491		6,311	

b) Dependent variable: stopping innovation

	Innovation		Product Innovation		Market novelty		Product-line novelty		Process innovation		Continuous R&D	
	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.	m.E.	std.err.
<i>Innovation dynamics in t-1</i>												
into1	0.002	0.003	-0.002	0.003	0.002	0.002	-0.003	0.005	0.003	0.003	0.003	0.004
into0	-0.011	0.005 *	-0.006	0.004	-0.001	0.002	-0.003	0.003	-0.001	0.004	0.000	0.003
outo1	-0.004	0.011	-0.010	0.011	-0.006	0.010	-0.028	0.019	0.022	0.012	-0.002	0.007
outo0	0.012	0.010	0.016	0.009	0.003	0.002	0.000	0.004	-0.017	0.008 *	-0.004	0.005
lna	0.009	0.005	-0.001	0.005	0.001	0.002	-0.002	0.003	0.013	0.005 *	-0.009	0.003 *
lnb	-0.004	0.003	0.000	0.003	0.002	0.001	0.002	0.002	0.006	0.003 *	0.004	0.002 *
LR Chi <sup>2</sup>	37.6		31.5		56.4		55.1		63.3		55.1	
Pseudo R <sup>2</sup>	0.008		0.008		0.075		0.049		0.016		0.027	
# observ.	6,356		6,422		4,836		4,806		6,491		6,311	

\* p<0.05.

For the introduction of market novelties, we find that if other firms in the neighbourhood introduced a market novelty in the previous year, or if firms with market novelties moved away or ceased business, the probability to introduce a new-to-the-market innovation increases significantly. The marginal effects are quite large: 3.8 percentage points if firms with a market novelty moved out and 6.6 percentage points if firms with market novelties closed. The stimulating effect from neighbouring firms that introduced such innovations in the previous period is much smaller (0.9 percentage points, the average share of firms with market novelties is 5.9% in our sample). The results for the aggregated variables on innovation dynamics confirm the findings.

For the other dimension of novelty, product innovation that opens up a new product-line, we find fewer impacts of innovation dynamics in a firm's local environment. Only if neighbouring firms without such innovation have ceased business in the previous year we find a positive impact on the probability to introduce such innovations.

The probability to introduce a process innovation increases if neighbouring firms have introduced process innovations in the previous year. The marginal effect is 1.4 percentage points, while the average share of process innovators in the sample is 24.6%. This finding indicates a kind of local learning effect if firms can observe process innovation activities of their neighbours.

For continuous R&D, we find similar effects as for product innovation. If neighbouring firms started or stopped continuous R&D activities in the previous year, the probability to conduct continuous R&D increases. A high turbulence in R&D activities of others seems to stimulate own R&D. In addition, the probability decreases if firms without continuous R&D moved into the neighbourhood.

The results for the dependent variables measuring changes in firms' innovation activities (i.e. starting or stopping a certain type of innovation) show few impacts of dynamics in the innovative environment of a firm (Table 6). One reason for the largely insignificant results may relate to the fact that year-on-year changes in innovation are rather rare events, which is also caused by the definition of innovation (and R&D activities), which always relate to a 3-year reference period. What we do find are positive impacts on entering into market novelties and process innovations if non-innovating neighbouring firms moved out, closed down or changed from innovating to non-innovating in the previous year. For stopping innovation, we see a lower probability for firms with either product or process innovation if non-innovators moved in or neighbouring firms stopped innovation. For process innovation, the probability to start with this type of innovation decreases if neighbouring firms without process innovation moved out or ceased business.

## 6. Conclusion

This paper made an attempt to explore the role of microgeography for innovation activities in firms, using a comprehensive firm-level panel data set of Berlin-based firms. For each firm in the sample, we added information on the firm's neighbourhood, including proximity to science and education institutions, entertainment facilities, transport infrastructure, other firms (including start-ups and firms recently moving into the neighbourhood), population density and real estate price levels as well as dynamics in innovation activities of neighbouring firms. Based on exact address data, we applied a very detailed geographical scale, which allowed zooming into a firm's neighbourhood at the level of individual buildings. This database facilitates analyses of how neighbourhood configurations may affect knowledge exchange and firm innovation.

Using information on different types of innovation activities, the paper investigated location patterns of innovative and non-innovative firms and the impact of local innovation dynamics on innovation in firms. We find that innovative firms locate much closer to research institutes and universities than non-innovative firms. For innovating firms very close proximity (50m and less) to research institutes seems to be particularly attractive for. Product innovators and firms with continuous R&D seem to have a specifically high preference for locating close to science and education institutions. For firms with new-to-market innovations or with continuous in-house R&D, close proximity to other firms from the same sector is a distinctive feature. These firms also prefer locations with a high firm dynamic, measured by the number of firms moving into the neighbourhood and the number of start-ups.

We find virtually no relation between the innovativeness of firms and the proximity to entertainment facilities such as bars, restaurants, galleries, theatres, cinemas and leisure facilities. This finding is in some contrast to the literature on creative urban environments and their likely role to stimulate also creativity and innovation in firms. Innovative firms tend to prefer closer proximity to urban transport and are located in urban neighbourhoods with a slightly lower population density compared to non-innovative firms. We also see some difference in the price level for commercial units as long as firms have introduced product innovations with a high degree of novelty, with 4-5% higher rents.

The innovation dynamics in a firm's neighbourhood (defined as changes in innovation activities in other firms located within a 250m radius) do have some impact on innovation. The probability to introduce a product innovation or to conduct continuous in-house R&D activities increases if neighbouring firms have changed their innovation behaviour (for the respective type of innovation activity) in the previous year. As both 'positive' and 'negative' changes do play a role, it seems that local turbulence in innovation motivates firms to consider innovation for themselves. For market novelties, we find a positive effect if neighbouring firms have introduced such innovation in the year before or if firms with market novelties moved out of the neighbourhood or closed their business. While the first finding

may indicate some type of local learning, the latter may show that firms try to fill in innovative gaps left by leaving firms.

This research was a first attempt to zoom into the role of local conditions for innovation. While we think that some of our results widen our understanding of how the very local environment may affect innovation, many relevant research questions remain open. First, we can only observe correlations between innovation activities of firms and their local environment. We do not know to what extent a firm is connected to its neighbourhood, e.g. whether it actively observes the things going on in its neighbourhood or whether it communicates or cooperates with other firms and organisations located nearby. Though we do have panel data, the limited length of our time series (5 years) prevents more in-depth analysis of how changes in the local environment lead to changes in innovation strategies. There is hence an urgent demand for more detailed studies on the microgeography of innovation.

## 7. References

- Agrawal, A., A. Galasso, A. Oettl (2016), Roads and innovation, *Review of Economics and Statistics* (doi:10.1162/REST\_a\_00619).
- Allen, T.J. (1977), *Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information Within the R&D Organization*, Cambridge, Ma.: MIT Press.
- Arsanjani, J.J., P. Mooney, A. Zipf, A. Schauss (2015), Quality assessment of the contributed land use information from OpenStreetMap versus authoritative datasets, in: J.J. Arsanjani, A. Zipf, P. Mooney, M. Helbich (eds.), *OpenStreetMap in GIScience: Experiences, Research, and Applications*, Heidelberg et al.: Springer, 37–58.
- Arsanjani, J.J., E. Vaz (2015), An assessment of a collaborative mapping approach for exploring land use patterns for several European metropolises, *International Journal of Applied Earth Observation and Geoinformation* 35, 329–337.
- Audretsch, D.B. (2003), Innovation and spatial externalities, *International Regional Science Review* 26(2), 167–174.
- Audretsch, D.B., M.P. Feldman (1996), R&D Spillovers and the Geography of Innovation and Production, *American Economic Review* 86(3), 630–640.
- Bersch, J., S. Gottschalk, B. Müller, M. Niefert (2014), *The Mannheim Enterprise Panel (MUP) and Firm Statistics for Germany*, ZEW Discussion Paper No. 14-104, Mannheim.
- Boschma, R. (2005), Proximity and innovation. A critical assessment, *Regional Studies* 39, 61–74.
- Breschi, S., F. Lissoni (2001), Knowledge spillovers and local innovation systems: a critical survey, *Industrial and Corporate Change* 10(4), 975–1005.
- Breschi, S., F. Lissoni (2003), *Mobility and Social Networks: Localised Knowledge Spillovers Revisited*, CESPRI Working Paper 142, Milan.
- Cassiman, B., R. Veugelers (2006), In search of complementarity in innovation strategy: internal R&D and external knowledge acquisition, *Management Science* 52(1), 68–82.

- Catalini, C. (2016), *Microgeography and the Direction of Inventive Activity*, Rotman School of Management Working Paper No. 2126890, Cambridge, Ma.
- Cohendet, P., D. Grandadam, L. Simon (2010), The anatomy of the creative city, *Industry and Innovation* 17, 91–111.
- Cooke, P. (2001), Regional innovation systems, clusters, and the knowledge economy, *Industrial and Corporate Change* 10(4), 945–974.
- Cooke, P. (2002), Regional innovation systems, *The Journal of Technology Transfer* 27(1), 133–145.
- Dorn, H., T. Törnros, A. Zipf (2015), Geo-Information comparison with land use data in Southern Germany, *International Journal of Geo-Information* 4, 1657–1671.
- European Environment Agency (2012), *Population Density Disaggregated with Corine Land Cover 2000*, Copenhagen: EEA.
- Feldman, M.P. (2014), The character of innovative places: entrepreneurial strategy, economic development, and prosperity, *Small Business Economics* 43, 9–20.
- Feldman, M.P., D.B. Audretsch (1999), Innovation in cities: science-based diversity, specialization and localized competition, *European Economic Review* 43, 409–429.
- Feldman, M.P., R. Florida (1994), The geographic sources of innovation: technological infrastructure and product innovation in the United States, *Annals of the Association American Geographers* 84(2), 210–229.
- Flanagin, A.J., M.J. Metzger (2008), The credibility of volunteered geographic information, *GeoJournal* 72(3-4), 137–148.
- Florida, R. (1995), Toward the learning region, *Futures* 27(5), 527–536.
- Florida, R., K. King (2016), *Rise of the Urban Startup Neighborhood: Micro-clusters of Venture Capital and Startup Activity at the Neighborhood Level*, Martin Prosperity Research Working Paper Series 2016-MPIWP-003, University of Toronto.
- Forman, C., A. Goldfarb, S. Greenstein (2016), Agglomeration of invention in the Bay Area: not just ICT, *American Economic Review* 106(5), 146–151.
- Gallego, F.J. (2010), A population density grid of the European Union, *Population and Environment* 31(6), 460–473.
- Gertler, M.S. (2003), Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there), *Journal of Economic Geography* 3, 75–99.
- Girres, J.F., G. Touya (2010), Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS* 14(4), 435–459.
- Glaeser, E.L. (1999), Learning in cities, *Journal of Urban Economics* 46(2), 254–277.
- Glaeser, E.L. (2000), The new economics of urban and regional growth, in: G. Clark, M. Gertler, M. Feldman (eds.), *The Oxford Handbook of Economic Geography*, Oxford: Oxford University Press, 83–98.
- Haklay, M. (2010), How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets, *Environment and Planning B: Planning and Design* 37(4), 682–703.
- Hecht, R., C. Kunze, S. Hahmann (2013), Measuring completeness of building footprints in OpenStreetMap over space and time, *ISPRS International Journal of Geo-Information* 2, 1066–1091.

- Heckman, J., H. Ichimura, P. Todd (1998), Matching as an econometric evaluation estimator, *Review of Economic Studies* 65(2), 261–294.
- Howells, J.R.L. (2002), Tacit knowledge, innovation and economic geography, *Urban Studies* 39(5-6), 871–884.
- Jacobs, J. (1961), *The Death and Life of Great American Cities*, New York: Random House.
- Jacobs, J. (1970), *The Economy of Cities*, New York: Vintage.
- Jaffe, A.B., M. Trajtenberg, R. Henderson (1993), Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics* 108, 578–598.
- Kabo, F.W., N. Cotton-Nessler, Y. Hwang, M.C. Levenstein, J. Owen-Smith (2014), Proximity effects on the dynamics and outcomes of scientific collaborations, *Research Policy* 43, 1469–1485.
- Kloosterman, R.C., A. Brandellero (2016), "All these places have their moments": Exploring the Micro-Geography of Music Scenes: The Indica Gallery and the Chelsea Hotel, *M/C Journal* 19(3).
- Löfsten, H., P. Lindelöf (2002), Science Parks and the growth of new technology-based firms—academic-industry links, innovation and markets, *Research Policy* 31(6), 859–876.
- Marshall, A. (1891), *Principles of Economics*, London: Macmillan.
- Neis, P., D. Zielstra, A. Zipf (2011), The street network evolution of crowdsourced maps: OpenStreetMap in Germany 2007–2011, *Future Internet* 4(1), 1–21.
- Olson, J.S., S. Teasley, L. Covi, G. Olson (2002), The (currently) unique advantages of collocated work, in: P.J. Hinds, S. Kiesler (eds.), *Distributed Work*, Cambridge, Ma.: Academic Press, 113–136.
- Peters, B., C. Rammer (2013), Innovation panel surveys in Germany, in: F. Gault (ed.), *Handbook of Innovation Indicators and Measurement*, Cheltenham and Northampton, Ma.: Edward Elgar, 135–177.
- Phan, P.H., D.S. Siegel, M. Wright (2005), Science parks and incubators: observations, synthesis and future research, *Journal of Business Venturing* 20(2), 165–182.
- Porter, M. (1996), Competitive advantage, agglomeration economies, and regional policy, *International Regional Science Review* 19(1), 85–94.
- Rammer, C., D. Czarnitzki, A. Spielkamp (2009), Innovation success of non-R&D-performers: substituting technology by management in SMEs, *Small Business Economics* 33, 35–58.
- Saxenian, A. (1994), *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*, Cambridge, Ma: Harvard University Press.
- Singh, J., M. Marx (2013), Geographic constraints on knowledge diffusion: political borders vs. spatial proximity, *Management Science* 59, 2056–2078.
- Sternberg, R. (1990), The impact of innovation centres on small technology-based firms: the example of the federal republic of Germany, *Small Business Economics* 2(2), 105–118.
- Thompson, P. (2006), Patent citations and the geography of knowledge spillovers: evidence from inventor- and examiner-added citations, *Review of Economics and Statistics* 88, 383–389.

West, J., M. Bogers (2013), Leveraging external sources of innovation: a review of research on open innovation, *Journal of Product Innovation Management* 31(4), 814–831.

## 8. Appendix

### Measurement of location variables

#### *OSM data quality*

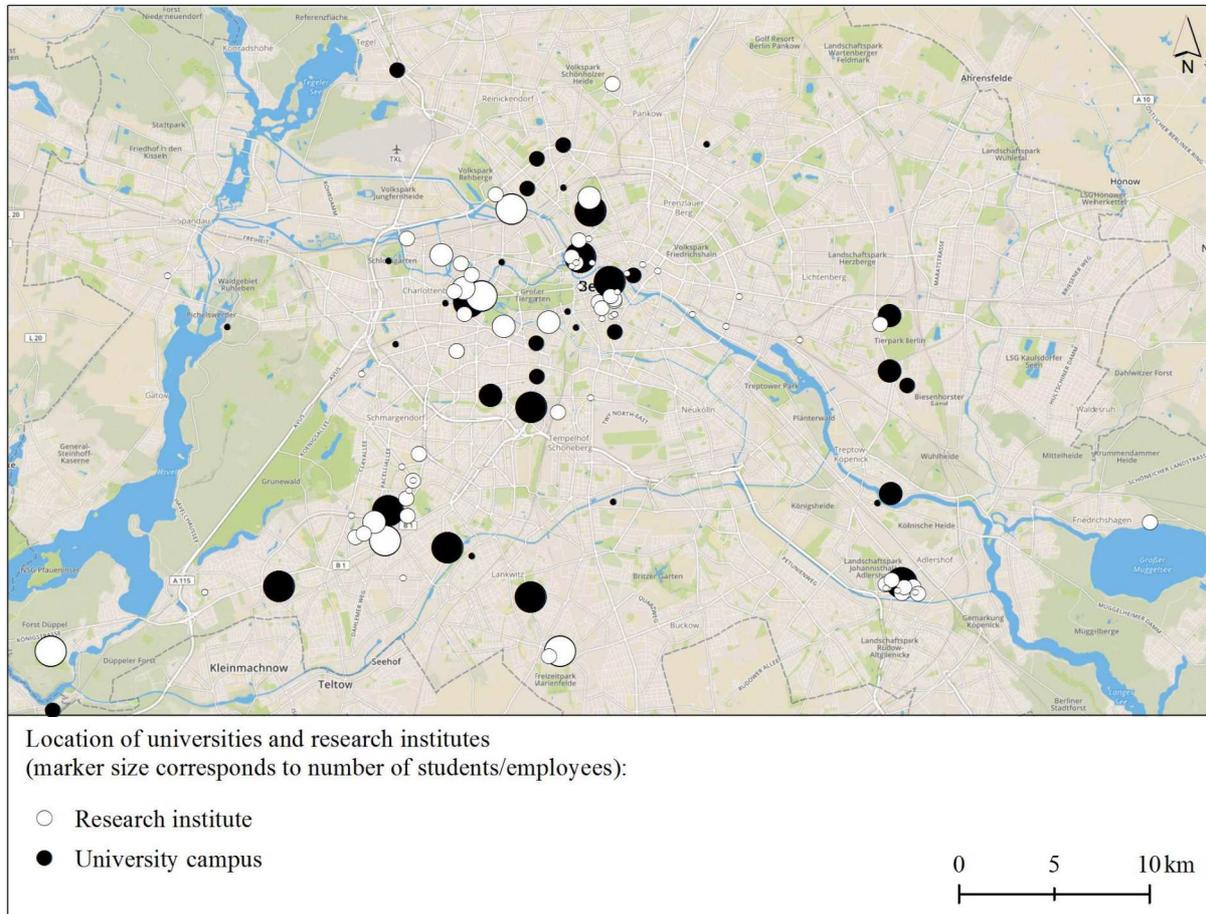
OpenStreetMap (OSM) is a collaborative mapping project, which allows users to create freely accessible geographic data. This kind of crowdsourced geographic information is also known as Volunteered Geographic Information (VGI). In addition to roads, OSM includes map features such as retail shops, public transport facilities, and a variety of natural features. Other than conventional geographic data, as provided by the authorities and commercial vendor, the OSM contributors are usually laymen. Users can add and modify map features after registration without any prior review of their skills or expertise. Concerns about the quality of this kind of user-generated geographic information seem natural and emerged shortly after the launch of the project in 2004 (Flanagin & Metzger 2008). Since then an array of studies investigated OSM data and assessed the geometric, attributive and temporal accuracy and completeness of the mapped features. Besides intrinsic approaches, most of these studies compare OSM data to established commercial or official geographic data of road networks (Haklay 2010; Girres & Touya 2010; Neis et al. 2011), buildings (Hecht et al. 2013), and land use data (Arsanjani et al. 2015; Arsanjani & Vaz 2015; Dorn et al. 2015). Their results show, first, that OSM data is only slightly inferior to official/commercial data in terms of accuracy. Second, OSM data completeness increases at a rapid rate and is assumed to have reached or exceeded the level of completeness of commercial data in the meantime. Third, OSM's completeness is positively correlated to population density and can be considered to be particularly suitable for spatial analysis in cities. Apart from these studies, the appreciation of user-generated geographic information is reflected by the implementation of crowdsourcing mechanisms in commercial web maps (e.g. Google Maps), too. Hence, the user-generated geographic data of OSM can be regarded as non-inferior to commercial and official products - especially in urban environments.

#### *Science & education*

The role of universities and research institutes as location factors for firms is determined by the count of students (for universities) and researchers (for research institutes) within multiple distance thresholds around each firm (50, 100, 250, 500, 1000, and 2500 meters). Both the geographic data of Berlin-based universities and research institutes are based on our own inquiry. The university dataset lists 63 locations (campuses) and includes the number of students there. If no exact number for different university campuses of a certain university was available, we allocated a proportional number of students to each location. This concerned only a small number of campuses of small universities. The other dataset includes 83 locations of public research institutes, scientific libraries, as well as agencies running

science and research funding programmes. In the case of multi-site institutes, proportional numbers of researchers were allocated to each branch office. This concerns only a few organisations.

Figure 4: Location of universities and research institutes in Berlin



Basemap: MapBox/OSM

### *Entertainment*

The presence of local leisure opportunities is determined by counting the number of respective facilities within multiple distance thresholds around each firm (50, 100, 250, 500, 1000, and 2500 meters). The location of 538 Berlin-based cultural facilities (e.g. theatres, arts centres), 9,224 leisure facilities (e.g. parks, playgrounds), and 9,077 gastronomy facilities (e.g. bars, pubs, restaurants) has been extracted from OSM (OpenStreetMap Foundation 2016) using relevant tags. Both point and areal features were considered. The OSM data were obtained in September 2016 as an unmodified full copy (*Planet.osm*).

### *Transport*

The connectivity to local, national and international passenger services is determined by measuring the distance between a considered firm and the nearest transportation hub. We distinguish between local public transports (e.g. bus, metro), main-line railway stations with

intercity connectivity, and international airports. The above mentioned OSM data set was used to extract the locations of 7,351 public transport stops and stations in Berlin. The locations of the two Berlin-based international airports (*Schönefeld* and *Tegel*) and ten main-line railway stations in the greater Berlin area are based on our own inquiry.

#### *Business activities*

We use the *Mannheim Enterprise Panel* and the *Mannheim Innovation Panel* to derive information on business activities in a firm's neighbourhood. Both the spatial dimension of business activities (relocations, start-ups, closures) and their innovation behaviour are considered. First, we determine the degree of sector localisation at each firm site by counting the number of firms of the same sector (2-digit level) within 50, 100, 250, 500, 1000, and 2500 meters. Second, we measure local spatial firm dynamics by counting the number of incoming firm relocations and the number of firm start-ups within the aforementioned distance thresholds. For incoming firms and start-ups, all sectors covered by the Berlin Innovation Panel (i.e. manufacturing and knowledge-intensive services) are considered.

#### *Population density*

We obtain the local population density for each firm location from a population density grid with 100m resolution (European Environment Agency 2012). This raster covers 27 EU member states and is based on a combination of population data per commune and CORINE Land Cover data. The population numbers per 100 by 100 meters raster cell were obtained via downscaling, as summarised by Gallego (2010).

#### *Real estate market*

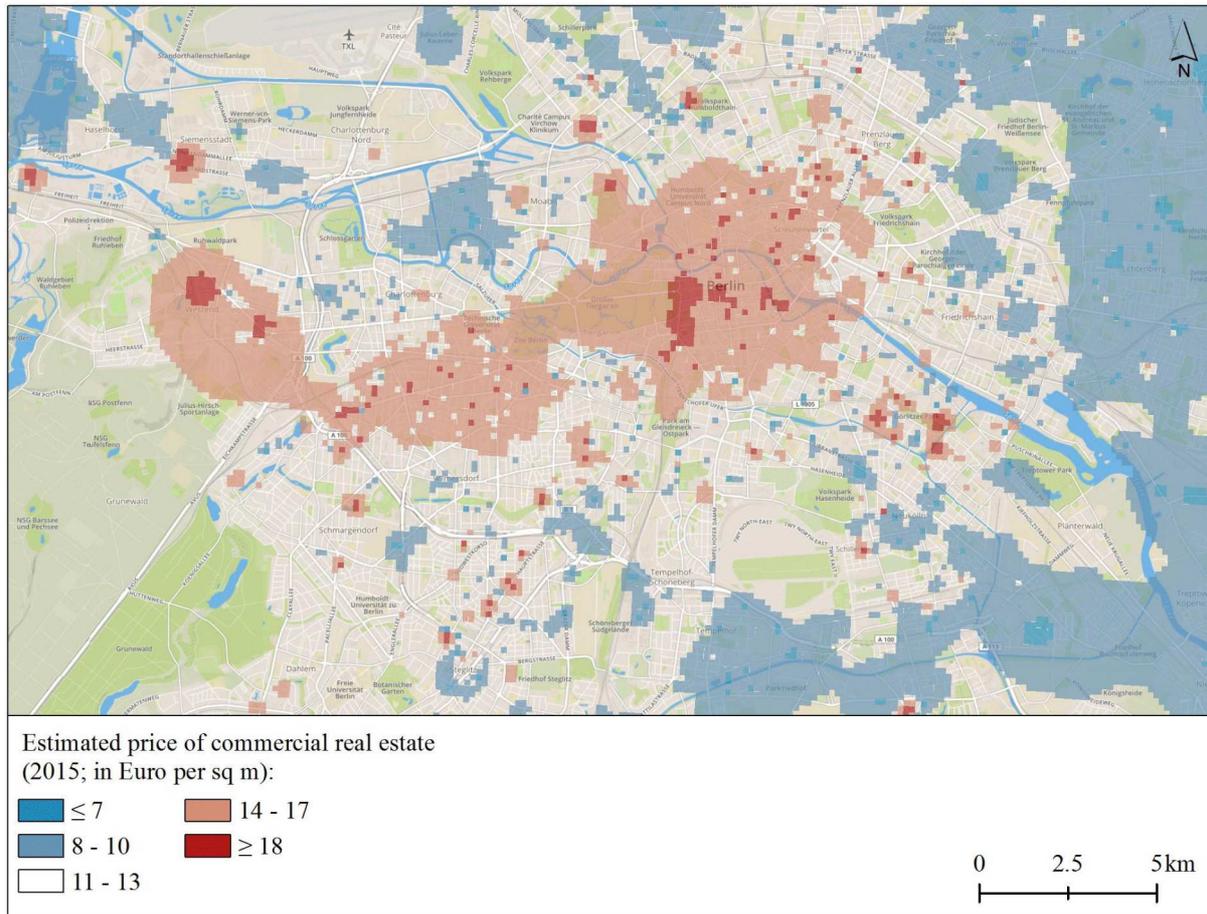
The local rent levels of both commercial and residential real estate are obtained from the *empirica-systeme Marktdatenbank* database. This database contains information on the German real estate market, aggregated from public real estate advertisements. Due to our detailed spatial level of analysis, we processed only those data entries that include location information on the street number level. The resulting annual datasets include 7,000 to 8,500 (commercial estate) and 24,000 to 36,000 (residential estate) point observations. We interpolated (inverse distance weighted) each annual dataset to derive a continuous estimate of local rent levels. These continuous surfaces allow us to calculate the annual change in local rent levels, too.

#### *Innovation dynamics*

We use information from the Berlin Innovation Panel to determine the innovation dynamics in the neighbourhood of each firm in each year. For the main model estimations, we only look at a 250m threshold. For this threshold, we determine the number of firms that have changed their innovation status in the previous year (considering different types of innovation activities), the number of incoming and outgoing firms in the previous year (by innovation

status in that year), and the number firms that have been newly founded or closed in the previous year (by innovation status in that year).

Figure 5: Price level of commercial real estate in the central parts of Berlin



Basemap: MapBox/OSM; Data: empirica-systeme.

Table 7: Descriptive statistics of model variables

*a) Innovation indicators for matching*

	# obs.	Mean	Std.dev.	Min	Max
Innovator (inn)	13,405	0.390	0.488	0	1
Product innovator (ipd)	13,286	0.313	0.464	0	1
Process innovator (ipc)	13,297	0.246	0.430	0	1
Market novelty (imn)	10,503	0.054	0.226	0	1
Product-line novelty (ipn)	10,471	0.073	0.261	0	1
Cost-reducing proc. inn. (ico)	10,930	0.043	0.202	0	1
Quality-improving proc. inn. (iqu)	10,895	0.067	0.249	0	1
Continuous in-house R&D (rdc)	12,363	0.212	0.409	0	1
Occasional in-house R&D (rdo)	12,363	0.138	0.345	0	1
Contracting-out R&D (rde)	8,962	0.054	0.225	0	1

*b) Innovation indicators for probit models*

Variable	# obs.	Mean	Std.dev.	Min	Max	Variable	# obs.	Mean	Std.dev.	Min	Max
inn	6,736	0.399	0.4990	0	1	ipd_out	6,422	0.095	0.294	0	1
ipd	6,742	0.316	0.465	0	1	imn_into	4,836	0.011	0.102	0	1
imn	4,915	0.059	0.236	0	1	imn_out	4,836	0.015	0.121	0	1
ipn	4,905	0.076	0.265	0	1	ipn_into	4,806	0.016	0.125	0	1
ipc	6,745	0.246	0.431	0	1	ipn_out	4,806	0.025	0.155	0	1
rdc	6,311	0.226	0.418	0	1	ipc_into	6,491	0.069	0.255	0	1
inn_into	6,356	0.074	0.261	0	1	ipc_out	6,491	0.095	0.293	0	1
inn_out	6,356	0.112	0.316	0	1	rdc_into	6,311	0.033	0.179	0	1
ipd_into	6,422	0.062	0.241	0	1	rdc_out	6,311	0.038	0.192	0	1

*c) Location variables for matching*

Variable	# obs.	Mean	Std.dev.	Min	Max	Variable	# obs.	Mean	Std.dev.	Min	Max
uni_50	12,851	8.8	236	0	10,917	leis_1000	12,851	13	11	0	49
uni_100	12,851	22	368	0	11,180	leis_2500	12,851	75	46	0	179
uni_250	12,851	123	976	0	11,180	pubtr	12,851	158	159	1	3,515
uni_500	12,851	782	2,621	0	12,133	air	12,851	9,160	3,385	333	23,990
uni_1000	12,851	2,390	4,909	0	30,395	train	12,851	3,881	2,542	124	16,209
uni_2500	12,851	11,750	13,394	0	49,342	fst_50m	12,851	0.4	1.3	0	14
ins_50	12,851	1.9	22	0	400	fst_100m	12,851	0.6	1.5	0	14
ins_100	12,851	5.6	38	0	445	fst_250m	12,851	1.4	2.7	0	21
ins_250	12,851	33	143	0	1,460	fst_500m	12,851	3.4	5.6	0	43
ins_500	12,851	88	251	0	2,286	fst_1000m	12,851	10	15	0	85
ins_1000	12,851	239	496	0	3,036	fst_2500m	12,851	39	53	0	236
ins_2500	12,851	1,130	1,388	0	6,007	fin_50m	12,851	2.1	7.2	0	96
bars_50	12,851	0.7	1.3	0	11	fin_100m	12,851	3.8	9.1	0	102
bars_100	12,851	2.1	3.3	0	27	fin_250m	12,851	13	24	0	274
bars_250	12,851	11	15	0	109	fin_500m	12,851	40	64	0	578
bars_500	12,851	40	48	0	245	fin_1000m	12,851	130	175	0	1,100
bars_1000	12,851	139	148	0	540	fin_2500m	12,851	595	602	0	2,184
bars_2500	12,851	692	576	0	1,690	st_50m	12,851	0.2	0.7	0	20
cult_50	12,851	0.0	0.2	0	3.0	st_100m	12,851	0.4	1.0	0	22
cult_100	12,851	0.1	0.5	0	8.0	st_250m	12,851	1.5	2.7	0	34
cult_250	12,851	0.7	1.5	0	11	st_500m	12,851	4.8	7.5	0	56
cult_500	12,851	2.5	4.0	0	27	st_1000m	12,851	16	22	0	149
cult_1000	12,851	9.0	12	0	66	st_2500m	12,851	74	78	0	375
cult_2500	12,851	45	48	0	168	popden	12,849	6,634	4,696	0	12,795
leis_50	12,851	0.0	0.2	0	3.0	comm	12,845	10.4	2.8	3	50
leis_100	12,851	0.1	0.4	0	4.0	comm_ch	10,139	0.3	1.8	-34	32
leis_250	12,851	0.8	1.3	0	9.0	resdt	12,849	8.4	1.5	4	17
leis_500	12,851	3.3	3.5	0	18	resdt_ch	10,140	0.4	0.8	-7	6.0

d) Innovation dynamics variables for probit models

Variable	# obs.	Mean	Std.dev.	Min	Max	Variable	# obs.	Mean	Std.dev.	Min	Max
in_inn1	9,748	0.081	0.342	0	4	in_ipn1	9,748	0.014	0.118	0	1
in_inn0	9,748	0.085	0.325	0	4	in_ipn0	9,748	0.118	0.399	0	5
ot_inn1	9,748	0.085	0.348	0	5	ot_ipn1	9,748	0.021	0.152	0	2
ot_inn0	9,748	0.080	0.325	0	3	ot_ipn0	9,748	0.123	0.425	0	6
ch_inn1	9,748	0.766	1.211	0	10	ch_ipn1	9,748	0.168	0.443	0	4
ch_inn0	9,748	0.478	0.837	0	6	ch_ipn0	9,748	0.194	0.514	0	4
cl_inn1	9,748	0.033	0.189	0	2	cl_ipn1	9,748	0.006	0.076	0	1
cl_inn0	9,748	0.034	0.181	0	1	cl_ipn0	9,748	0.044	0.211	0	2
nf_inn1	9,748	0.001	0.036	0	1	nf_ipn1	9,748	0.000	0.014	0	1
nf_inn0	9,748	0.023	0.151	0	2	nf_ipn0	9,748	0.029	0.171	0	2
into_inn1	9,748	0.848	1.285	0	10	into_ipn1	9,748	0.182	0.467	0	4
into_inn0	9,748	0.586	0.957	0	6	into_ipn0	9,748	0.341	0.720	0	6
outo_inn1	9,748	0.118	0.395	0	5	outo_ipn1	9,748	0.027	0.169	0	2
outo_inn0	9,748	0.114	0.388	0	4	outo_ipn0	9,748	0.167	0.491	0	6
in_ipd1	9,748	0.064	0.293	0	3	in_ipc1	9,748	0.049	0.247	0	3
in_ipd0	9,748	0.103	0.363	0	4	in_ipc0	9,748	0.119	0.402	0	4
ot_ipd1	9,748	0.074	0.310	0	4	ot_ipc1	9,748	0.042	0.241	0	4
ot_ipd0	9,748	0.092	0.356	0	4	ot_ipc0	9,748	0.127	0.425	0	4
ch_ipd1	9,748	0.670	1.105	0	8	ch_ipc1	9,748	0.613	1.083	0	10
ch_ipd0	9,748	0.428	0.780	0	5	ch_ipc0	9,748	0.490	0.901	0	7
cl_ipd1	9,748	0.026	0.166	0	2	cl_ipc1	9,748	0.019	0.136	0	1
cl_ipd0	9,748	0.040	0.202	0	2	cl_ipc0	9,748	0.048	0.215	0	2
nf_ipd1	9,748	0.001	0.030	0	1	nf_ipc1	9,748	0.001	0.025	0	1
nf_ipd0	9,748	0.024	0.156	0	2	nf_ipc0	9,748	0.030	0.179	0	3
into_ipd1	9,748	0.736	1.176	0	9	into_ipc1	9,748	0.663	1.117	0	10
into_ipd0	9,748	0.554	0.940	0	6	into_ipc0	9,748	0.639	1.079	0	7
outo_ipd1	9,748	0.100	0.348	0	4	outo_ipc1	9,748	0.061	0.280	0	4
outo_ipd0	9,748	0.132	0.426	0	4	outo_ipc0	9,748	0.175	0.493	0	4
in_imn1	9,748	0.013	0.115	0	1	in_rdc1	9,748	0.051	0.271	0	3
in_imn0	9,748	0.118	0.401	0	6	in_rdc0	9,748	0.124	0.432	0	4
ot_imn1	9,748	0.013	0.116	0	2	ot_rdc1	9,748	0.062	0.328	0	6
ot_imn0	9,748	0.129	0.444	0	6	ot_rdc0	9,748	0.119	0.408	0	3
ch_imn1	9,748	0.151	0.477	0	6	ch_rdc1	9,748	0.188	0.474	0	3
ch_imn0	9,748	0.171	0.492	0	4	ch_rdc0	9,748	0.269	0.606	0	6
cl_imn1	9,748	0.006	0.080	0	2	cl_rdc1	9,748	0.013	0.113	0	1
cl_imn0	9,748	0.045	0.213	0	2	cl_rdc0	9,748	0.054	0.240	0	2
nf_imn1	9,748	0.000	0.014	0	1	nf_rdc1	9,748	0.010	0.101	0	1
nf_imn0	9,748	0.026	0.163	0	2	nf_rdc0	9,748	0.028	0.172	0	3
into_imn1	9,748	0.164	0.514	0	7	into_rdc1	9,748	0.249	0.593	0	4
into_imn0	9,748	0.315	0.700	0	6	into_rdc0	9,748	0.420	0.816	0	6
outo_imn1	9,748	0.019	0.140	0	2	outo_rdc1	9,748	0.075	0.344	0	6
outo_imn0	9,748	0.174	0.508	0	6	outo_rdc0	9,748	0.173	0.483	0	4

in\_ : firms moving into neighbourhood  $n$  in  $t-1$

out\_ : firms moving out of neighbourhood  $n$  in  $t-1$

ch\_ : firms from neighbourhood  $n$  changing innovation status in  $t-1$

cl\_ : firms from neighbourhood  $n$  stopping business in  $t-1$

nf\_ : firms starting business in neighbourhood  $n$  in  $t-1$

Table 8: Matching results for type of product/process innovation

	Market novelty			Product-line novelty			Cost-reducing process			Quality-improving proc.		
	mean	$\delta$	t value	mean	$\delta$	t value	mean	$\delta$	t value	mean	$\delta$	t value
uni_50	6	0	0.00	5	11	0.08	15	100	1.42	5	100	1.00
uni_100	7	-9	-0.07	6	-7	-0.06	17	86	1.36	7	2	0.02
uni_250	157	32	0.69	210	74	2.46 *	150	41	0.73	82	-51	-0.68
uni_500	1517	48	3.28 *	1247	43	2.87 *	1256	43	2.24 *	1036	35	2.00 *
uni_1000	3657	43	4.47 *	3222	27	2.69 *	2760	15	1.05	2601	9	0.74
uni_2500	12875	11	1.54	12921	9	1.46	10877	2	0.17	11432	-3	-0.42
ins_50	6.9	66	2.24 *	4.6	63	1.92	1.8	69	1.19	2.6	40	0.82
ins_100	13	51	2.15 *	10	42	1.63	6.9	70	2.22 *	8.1	-8	-0.22
ins_250	76	51	2.85 *	61	48	2.76 *	57	45	1.92	63	47	2.69 *
ins_500	194	51	4.21 *	144	42	3.30 *	142	43	2.51 *	140	40	2.98 *
ins_1000	398	36	3.52 *	335	28	2.75 *	306	27	1.89	293	17	1.48
ins_2500	1275	9	1.16	1255	11	1.52	1138	5	0.48	1111	-10	-1.24
bars_50	0.6	-15	-1.03	0.6	-2	-0.19	0.5	-1	-0.03	0.6	-3	-0.24
bars_100	1.9	-12	-0.97	1.9	-10	-0.83	1.5	-25	-1.56	1.8	-14	-1.24
bars_250	10	-1	-0.08	10	-3	-0.33	9.0	-17	-1.28	10	-7	-0.78
bars_500	39	3	0.40	40	5	0.67	34	-15	-1.30	37	-9	-1.05
bars_1000	136	2	0.31	139	4	0.63	127	-8	-0.83	132	-8	-1.17
bars_2500	687	4	0.70	700	7	1.23	655	-2	-0.24	670	-5	-0.86
cult_50	0.0	-27	-0.52	0.0	19	0.56	0.0	-150	-1.48	0.0	-38	-0.76
cult_100	0.1	-58	-1.43	0.1	-20	-0.70	0.0	-88	-1.50	0.1	-24	-0.86
cult_250	0.6	0	0.00	0.7	-4	-0.29	0.5	-15	-0.69	0.6	5	0.37
cult_500	2.4	11	1.01	2.5	9	0.89	2.2	-13	-0.92	2.2	-6	-0.51
cult_1000	9.0	11	1.19	9.1	8	0.96	8.2	-5	-0.46	8.5	-2	-0.23
cult_2500	46	6	0.84	46	7	1.10	43	3	0.36	44	1	0.11
leis_50	0.0	-67	-1.04	0.0	-44	-0.97	0.0	0	0.00	0.0	-5	-0.12
leis_100	0.1	33	1.51	0.1	16	0.84	0.1	-8	-0.24	0.1	5	0.21
leis_250	0.8	-12	-0.93	0.8	-11	-0.98	0.8	-22	-1.41	0.8	0	0.02
leis_500	3.3	3	0.37	3.3	3	0.40	3.3	-4	-0.46	3.5	1	0.12
leis_1000	13	2	0.34	13	3	0.66	13	-5	-0.73	13	-3	-0.51
leis_2500	72	1	0.15	73	3	0.67	72	-2	-0.39	74	0	-0.09
pubtr	168	9	1.63	176	4	0.66	156	2	0.28	152	-2	-0.33
air	8436	-9	-2.83 *	8790	-6	-2.20 *	8850	-5	-1.44	8986	-2	-0.75
train	4342	9	2.15 *	4193	3	0.87	4098	5	0.94	4027	7	1.77
fst_50	0.8	39	2.92 *	0.6	27	1.92	0.5	24	1.30	0.5	4	0.23
fst_100	0.9	30	2.40 *	0.8	26	2.11 *	0.6	21	1.27	0.6	3	0.23
fst_250	2.5	34	3.71 *	2.1	25	2.66 *	1.8	35	2.94 *	1.8	21	2.03 *
fst_500	5.2	26	3.07 *	4.6	18	2.09 *	3.8	27	2.47 *	3.9	12	1.25
fst_1000	13	19	2.19 *	12	9	1.12	10	21	2.02 *	10	4	0.49
fst_2500	43	-2	-0.24	43	-3	-0.40	37	7	0.66	39	-9	-0.98
fin_50	4	49	2.69 *	3.2	29	1.57	1.8	-41	-1.03	2.5	-14	-0.66
fin_100	6	36	2.51 *	5.2	27	2.05 *	3.3	-26	-1.09	4.4	-13	-0.78
fin_250	18	29	2.84 *	18	23	2.43 *	14	0	-0.02	15	-4	-0.34
fin_500	47	14	1.36	49	17	1.98 *	39	-4	-0.29	41	-12	-1.08
fin_1000	142	6	0.60	151	8	1.05	130	-2	-0.17	133	-11	-1.18
fin_2500	650	5	0.71	675	6	1.04	608	1	0.11	623	-7	-1.07
st_50	0.3	51	3.14 *	0.3	34	2.14 *	0.2	-12	-0.42	0.2	22	1.14
st_100	0.5	34	2.34 *	0.5	33	2.61 *	0.3	-7	-0.33	0.5	22	1.59
st_250	1.9	24	2.02 *	1.8	19	1.80	1.4	-2	-0.11	1.7	18	1.66
st_500	5.5	13	1.20	5.4	13	1.34	4.5	-2	-0.13	5.1	7	0.68
st_1000	17	10	1.05	17	7	0.80	16	4	0.37	17	5	0.56
st_2500	79	8	1.02	80	8	1.16	74	3	0.30	78	2	0.23
popden	5807	-8	-1.28	6041	-3	-0.66	5787	-9	-1.34	6033	-7	-1.35
comm	11	5	2.76 *	11	4	2.87 *	10	1	0.71	10	2	1.27
comm_ch	0.4	74	1.82	0.3	21	0.60	0.3	11	0.25	0.3	15	0.45
resdt	8.5	2	1.56	8.6	2	1.46	8.4	0	-0.20	8.5	0	0.35
resdt_ch	0.4	-7	-0.54	0.4	-19	-1.35	0.3	-18	-1.02	0.3	-24	-1.61
# obs. <sup>a</sup>	446 / 7,326			592 / 7,158			336 / 7,813			539 / 7,572		

\* p<0.05.  $\delta$ : difference to control group in %. a: number of innovators / no. of non-innovators with common support.

Table 9: Matching results for R&D activities

Variable	Continuous in-house R&D			Occasional in-house R&D			Contracting-out of R&D		
	mean	$\delta$	t value	mean	$\delta$	t value	mean	$\delta$	t value
uni_50	11	41	0.61	1	-2523	-2.35	28	97	0.97
uni_100	26	69	1.77	23	-35	-0.49	32	90	1.02
uni_250	188	56	2.73 *	153	44	1.79	117	25	0.40
uni_500	1356	52	6.85 *	691	-22	-1.38	1480	53	3.24 *
uni_1000	3249	35	6.03 *	2301	-11	-1.26	3761	41	3.83 *
uni_2500	12949	11	2.70 *	11963	-1	-0.31	13264	12	1.45
ins_50	5	69	3.82 *	2	-11	-0.19	5	56	1.37
ins_100	12	51	3.81 *	5	17	0.56	11	-11	-0.27
ins_250	76	63	7.58 *	33	-39	-1.89	83	49	2.66 *
ins_500	161	56	9.10 *	84	-28	-2.16	200	56	4.23 *
ins_1000	342	34	6.18 *	229	-12	-1.27	382	27	2.30 *
ins_2500	1244	7	1.62	1183	4	0.77	1307	9	1.10
bars_50	1	-7	-0.95	1	12	1.59 *	1	-6	-0.38
bars_100	2	-2	-0.32	2	5	0.83	2	-2	-0.13
bars_250	11	-6	-1.08	12	6	1.06	11	4	0.35
bars_500	40	-4	-0.94	42	1	0.27	43	15	1.67
bars_1000	144	-3	-0.84	148	2	0.46	147	16	2.03 *
bars_2500	714	-2	-0.71	720	1	0.23	715	7	1.01
cult_50	0	-15	-0.70	0	30	1.50	0	-11	-0.19
cult_100	0	3	0.24	0	21	1.56	0	26	0.88
cult_250	1	3	0.40	1	-2	-0.21	1	-8	-0.47
cult_500	3	0	0.04	3	6	0.92	3	18	1.56
cult_1000	10	3	0.73	9	-1	-0.26	10	15	1.50
cult_2500	48	0	0.12	48	2	0.44	48	4	0.53
leis_50	0	7	0.27	0	-33	-1.08	0	0	0.00
leis_100	0	-17	-1.17	0	11	0.77 *	0	47	2.11 *
leis_250	1	-7	-1.04	1	6	1.02	1	30	2.87 *
leis_500	3	-4	-1.03	4	2	0.43	3	6	0.67
leis_1000	13	-5	-1.67	14	3	0.83 *	13	8	1.33
leis_2500	77	-2	-1.00	78	0	0.04	75	2	0.49
pubtr	160	-2	-0.55	155	0	0.03	164	-15	-1.53
air	8975	-2	-1.73	9419	2	1.54 *	8895	-7	-2.19 *
train	4034	8	3.55 *	3753	-2	-0.61	4175	0	-0.06
fst_50	1	26	3.14 *	0	-39	-2.57	1	32	1.90
fst_100	1	26	3.68 *	1	-32	-2.59	1	27	1.82
fst_250	2	23	4.34 *	1	-15	-1.87	2	31	2.75 *
fst_500	5	13	2.71 *	4	-10	-1.44	5	23	2.29 *
fst_1000	12	7	1.44	11	-8	-1.31	12	15	1.50
fst_2500	44	-7	-1.43	46	-1	-0.14 *	42	1	0.07
fin_50	3	25	2.75 *	3	-2	-0.16	4	34	1.56
fin_100	5	18	2.41 *	4	-5	-0.56	5	21	1.19
fin_250	16	6	1.01	14	-9	-1.19	16	7	0.57
fin_500	46	0	0.05	43	-6	-0.99	47	11	0.97
fin_1000	146	1	0.16	143	-4	-0.79	150	14	1.41
fin_2500	674	2	0.58	671	-1	-0.18	671	7	0.88
st_50	0	14	1.24	0	-13	-0.86	0	33	1.82
st_100	0	11	1.26	0	-6	-0.61	1	43	2.96 *
st_250	2	4	0.59	2	1	0.16	2	18	1.32
st_500	5	2	0.38	5	3	0.48	6	12	0.97
st_1000	18	4	0.72	17	1	0.23	19	20	1.98 *
st_2500	80	2	0.57	79	2	0.43	83	8	1.04
popden	6269	-6	-2.08 *	6559	-2	-0.78	6214	1	0.09
comm	11	1	1.42	11	0	0.32	11	5	2.38 *
comm_ch	0	-37	-1.18	0	-9	-0.35	0	12	0.26
resdt	8	-1	-1.00	9	0	0.49	9	2	1.71
resdt_ch	0	-7	-0.76	0	7	0.77	0	-1	-0.04
# obs. <sup>a</sup>	2,318 / 7,510			1,437 / 8,391			384 / 6,168		

\* p<0.05.  $\delta$ : difference to control group in %. a: number of innovators / no. of non-innovators with common support.