Job-switching patterns in knowledge-intensive industries within the Helsinki-Uusimaa Region

Tamás Lahdelma

Working papers
2019:1

Helsinki

City of Helsinki, Executive Office,
Urban Research and Statistics

Job-switching patterns in knowledge-intensive industries within the Helsinki-Uusimaa Region

Tamás Lahdelma

Working papers
2019:1

Helsinki

The study is part of the European ROBUST (Rural-Urban Outlooks: Unlocking Synergies) research project. The project has received funding from the European Union's Horizon 2020 research and innovation programme (RUR-01-2016) under grant agreement No 727988.

Inquiries

Tamás Lahdelma
Katja Vilkama, tel. +358 9 31078396
firstname.lastname(at)hel.fi

Publisher

City of Helsinki, Executive Office,
Urban Research and Statistics

Address

P.O.Box 550, 00099 City of Helsinki
(Siltasaarenkatu 18–20 A)

Telephone

09 310 36377

Internet

www.hel.fi/urbanfacts

Orders, distribution

tel. 09 310 36293
kaupunkitieto.tilaukset@hel.fi

Map

Tuomas Tavi

Graphs and general layout

Lotta Haglund

ISSN 2489-477X (online)

Introduction

In regional and urban economics, the agglomeration of economic activities has been regarded as closely related to the creation and spread of knowledge (Marshall, 1920). Accordingly, economic activities intensive in the use of knowledge inputs in production show particularly strong spatial clustering tendencies (Arzaghi and Henderson, 2008; Rosenthal and Strange, 2008; Andersson et al., 2009; Larsson, 2017). However, when assessing the influence of geographical proximity following from local concentration, it must be taken into account that spatial closeness in itself is not a sufficient condition for the flow of information and knowledge, but it requires active participation in networks of knowledge sharing. Hiring employees from competing firms, business partners or other firms can be an important way of accessing such networks (Rao and Drazin, 2002; Song et al., 2003; Breschi and Lissoni, 2009). Ties to previous co-workers have also been shown to serve as a channel for knowledge exchange between organizations (Agrawal et al., 2006; Somaya et al., 2008; Corredoira and Rosenkopf, 2010). Hence, in empirical literature, labour mobility is considered an important factor explaining firm productivity (Maliranta et al., 2008; Eriksson and Lindgren, 2009) and regional growth (Boschma et al., 2014; Lengyel and Eriksson, 2017).

The importance of networks generated by employee mobility raises the question about the mechanisms underlying matching processes between employers and employees. Previous research on labour markets has demonstrated the importance of social network ties in the mediation of information about job openings and potential candidates for both recruiters and job seekers (Granovetter, 1995; Fernandez et al., 2000; Ioannides and Loury, 2004). Social processes conditioning matching between employers and employees can have a spatial dimension, as short distances are found to promote direct contacts and thus the exchange of information and knowledge are more easily transferred by face-to-face communication (Storper and Venables, 2004). This may be specifically so in well-networked knowledge-intensive service sectors in which informal interaction plays a key role (see Arzaghi and Henderson, 2008). The spatial embeddedness of social interaction, in turn, points to the question about the influence of community structure on the dynamics of local networks. The aim of the present paper is to address this issue by studying the spatial dimension of inter-organizational employee mobility in the knowledge-intensive industries of the integrated labour market of 26 municipalities in the Helsinki-Uusimaa Region.

If short distances promote direct communication and more rapid transmission of information between individuals, it can be assumed that it is more likely that a link will form between organizations closer to each other, all other things being equal. The empirical section of this paper tests this proposition using network data in which organizations are nodes and ties are created by the movement of employees across organizations. The local social network of organizations is constructed on the basis of a unique longitudinal matched employer–employee database covering selected knowledge-intensive industries during the period 2001–2015. The database combines information on the location of organizations with data on the entire organizational population from administrative business registers and a range of demographic and socioeconomic information on the individuals employed within these organizations. The assembled database makes it possible to investigate empirically how physical distance affects the probability that a tie is created, given the effect of other non-spatial factors making organizations interact and how the role of geography varies across the regional landscape. The findings of the paper have also some implications for future research on the relations between community structure and labour market dynamics, as well as policy suggestions for regional planning.

Background

Economic activities, particularly when intensive in the use of knowledge as an input, tend to be geographically highly concentrated (Arzaghi and Henderson, 2008; Rosenthal and Strange, 2008; Andersson et al., 2009; Larsson, 2017). Following Marshall (1920), research literature in economic geography and related fields identifies generally three sources of agglomeration economies as reasons for the observed spatial concentration: input sharing, labour market pooling, and knowledge spillovers. The first two sources of agglomeration benefits are often referred to as pecuniary externalities realised through market interactions, while knowledge spillovers, on the contrary, are considered to materialize through non-market interactions accessible to all members of the local community (Breschi and Lissoni, 2001).

Inter-firm mobility of employees is considered one of the key mechanisms mediating knowledge across firms. Previous empirical studies have shown that such diffusion, in turn, can increase the productivity and competitiveness of firms, industries and regions (Maliranta et al., 2008; Eriksson and Lindgren, 2009; Boschma et al., 2014; Lengyel and Eriksson, 2017). However, the effect of labour mobility on individual firms is not straightforward. Some firms suffer from job-switches in the form of loss of experienced workers to competitors and higher wages to retain its other workers (Combes and Duranton, 2006). Employee mobility is nevertheless generally seen beneficial to firms, as it is considered to speed up knowledge dissemination and thus learning processes, and to create bonds between different organizations (Power and Lundmark, 2004). Mobile employees not only move from one professional network to another but also form links between these networks thus fostering the flow of new information and ideas (Granovetter, 1995).

Theories of knowledge externalities explaining the role of geographical distance in knowledge diffusion make a distinction between "codified" and "tacit" knowledge and information, the latter being assumed to be transferable only informally through direct and repeated face-to-face interaction requiring spatial proximity (Audretsch, 1998). A closely related assumption is that information and communication technologies enable the geographical transfer of codified knowledge over long distances at low cost, while in the case of tacit knowledge this is not possible (Morgan, 2004).

However, Breschi and Lissoni (2001) argue that in the case of labour mobility, the externality of face-to-face interaction is that it reduces the costs associated with search and screening procedures. From this point of view, the concentration of networks is more related to agglomeration benefits based on the functioning of labour markets than on knowledge spillovers facilitated by spatial proximity, which means that knowledge externalities would also be pecuniary externalities transmitted through market mechanisms. Again, matching between employers and employees does not only happen through the price mechanism of labour markets, but social factors can also be of significance. In his seminal study, Granovetter (1995) showed that a job search is not only an economic rational process, but it is embedded in social relations constraining and defining the progress and results of the search. A general agreement is that approximately half of all vacancies are filled with persons who know someone from the firm offering the job (Durlauf, 2004). If short distances promote direct communication between individuals, the transmission of information about job openings can be assumed to be more rapid in a neighbourhood with higher employment density. The economic analysis of labour market processes must therefore take into account the social and spatial embeddedness of actors.

However, geographic proximity is not, *per se*, a sufficient condition for the transfer of knowledge, since economic agents can be located within close physical proximity without forming direct linkages to each other. Boschma's (2005) extensive review of the literature on different forms of proximity shows that in addition to geographical proximity, there are also other, non-spatial dimensions of proximity making organizations interact, such as cognitive, social, organizational and institutional proximity. Common to different proximity types is that they all are seen to reduce uncertainty and offer solutions to problems related to the coordination of interaction.

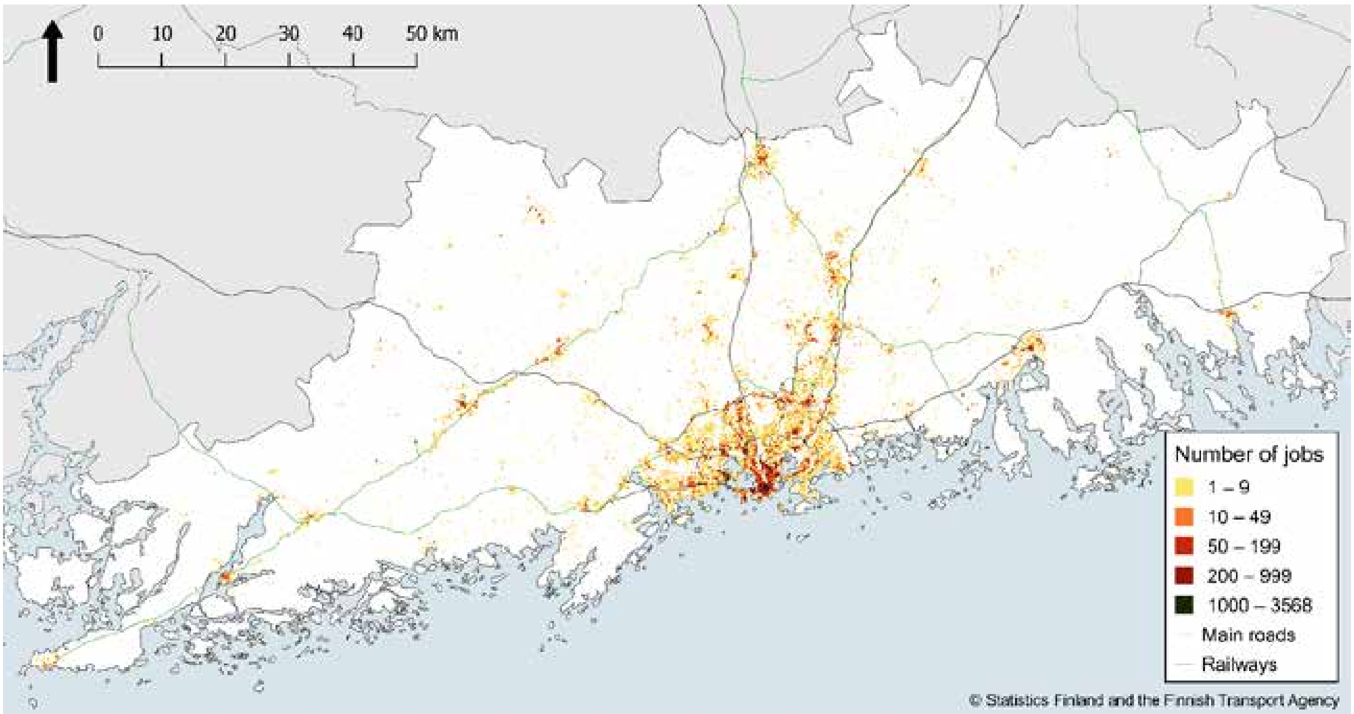
Cognitive proximity refers to the degree of similarity of knowledge bases of two actors. It is argued that actors are more likely to cooperate if they share similar knowledge bases, for it allows for more efficient communication, learning processes and knowledge sharing (Nooteboom et al., 2007). Organizational proximity is generally associated with a similarity of organizational positions among intra- and inter-organizational arrangements, as in the case, for example, for membership of the same organizational entity (Boschma, 2005). Further, the social dimension of proximity refers to the notion on the social embeddedness of economic relations at the micro-level. Relations between actors are considered to be socially embedded when they involve trust based on friendship, kinship and experience. The social embeddedness of organizations is due to the fact that organizational relations depend on the individual members of organizations and their relations (Granovetter, 1985). Finally, institutional proximity refers to the extent to which actors operate under the same set of formal rules, laws and norms as well as more informal cultural habits, routines, established practices and values (Boschma, 2005), e.g. when operating in the same social subsystem within industry (see Ponds et al., 2007).

The role of geographical proximity defined as the physical distance between organizations is considered in literature as a factor fostering interaction and cooperation. It is claimed that closely located firms have more face-to-face interaction and thus they are seen to be able to build trust more easily, which in turn is considered to lead to more personal relationships between firms. The growth of distance between actors weakens these positive externalities and makes communication more difficult (Boschma, 2005). In line with these arguments, empirical findings tend to confirm the sharp attenuation of knowledge spillovers with distance and the geographically bounded character of knowledge diffusion (e.g. Arzaghi and Henderson, 2008; Rosenthal and Strange, 2008; Andersson et al., 2009). This indicates that there can be assumed to be a substantial variation in the magnitude of spillovers across regions. However, it has also been presented that due to advanced information and communication technologies, networks through which learning takes place are no longer necessarily geographically limited, and face-to-face contacts required for the transfer of tacit knowledge can be arranged without permanent co-location (Rallet and Torre, 1999).

Data, model and variables

This study investigates an inter-organizational labour flow network in which links are generated by the mobility of employees from one organization to another in the knowledge-intensive sector[1] of the Helsinki-Uusimaa Region between year *t* and *t*+1. The empirical analysis focuses strictly on within-variation of job-switching patterns in the region consisting of 26 local municipalities integrated in terms of commuting flows.

Figure 1. Number of jobs, by neighbourhood (250 m × 250 m), in the Helsinki-Uusimaa Region (YKR). *Note*: the colour represents the absolute number of jobs in knowledge-intensive industries



The database used was assembled by combining individual level employment data with information on the entire organizational population of the region from administrative business registers and with a range of demographic and socioeconomic information on the individuals employed within the organizations. The empirical analysis focuses on all private work establishments of selected knowledge-intensive industries, subject to data availability (see Appendix 1 for the list of included NACE industries). Data sets were merged by unique ID numbers assigned to all individuals, establishments and firms of the registers. The database is a panel spanning from 2001 to 2015. All used registers are maintained by Statistics Finland and are generally of high quality. Each year, between 2 300 and 3 500 establishments were part of the network generated by the inter-firm mobility of employees.

The nature of the assembled database generates specific problems, which must be considered in the empirical setting. First, observations of network data are, by definition, non-independent, while conventional inferential formulas are based on the assumption of independence of observations. In a network approach, an individual or organization is, in

1 The number of establishments operating in the included industries accounted for about 29% of the total number of establishments in the region in 2015 (OSF).

contrast, positioned among the network of social or economic relations and the analysis focuses specifically upon the interdependencies between actors (see Abbott, 1997). Second, the rarity of job-switching events and the size of the data pose some challenges for the empirical strategy, as the vast number of potential dyads makes the analysis of the database at the whole-network level computationally infeasible, while drawing a random sample of the dyads of a sparse network would not fully utilise the available information.

For these reasons, a so-called matched case-control design is applied (see Sorenson and Stuart, 2008; Hosmer et al., 2013) using the potential dyads that make up the network. The strategy is similar to the approach used by Collet and Hedström (2013), who demonstrate on a data comparable to the database used in this study how the above-mentioned approach can be utilised to study network structure and tie-formation processes at the level of an entire labour market.

In the setting, each observation describes a pair of organizations, and the dichotomous outcome variable gets the value 1 if an employee switched from organization *i* to organization *j* between time *t* and *t*+1, and 0 if a link was not formed. In this paper, an organization is defined as a firm's work establishment with a unique geographic location. In order to test the relationship between different proximity dimensions and labour mobility, parameters of logistic regression models are estimated, specified as:

$$ln\left(\frac{p_{ijt}}{1-p_{ijt}}\right) = \alpha + \beta N_{ijt-1} + \gamma H_{ijt-1} + \delta I_{ijt-1} + \varepsilon X_{ijt-1} + \theta Y_{ijt-1}$$

where p_{ijt} is the probability that a link from organization *i* to organization *j* exists at time *t*; N_{ijt-1} is the variable measuring geographical distance between organization *i* and *j* at time *t*; H_{ijt-1} is a set of variables measuring how homophilous organization *i* and *j* are to one another at time *t*-1; I_{ijt-1} is a set of network-related variables measured at *t*-1, *t*-2 and *t*-3; X_{ijt-1} is a set of variables measuring the financial incentives for individuals to move from organization *i* to *j* at *t*-1; θY_{ijt-1} is a set of variables measuring other relevant properties of organization *i* and *j* at *t*-1 as well as regional interaction variables; and $\alpha,\beta,\gamma,\delta,\varepsilon,\theta$ are parameters to be estimated.

The database used for the estimation of the models described above is created so that all dyads directly linked to one another (with the value of 1 on the outcome variable) are included, forming the "cases" of the matched case-control design. Then a control group is defined for each of these cases from randomly selected five organizational pairs with a 0 on the outcome variable. Controls are matched with the cases so that the organizational dyads of the control group have the same industrial combination on the 2-digit NACE industry level as the case. This approach implies that all realized dyads are included and the controls are selected randomly, and thus there is no risk of biased estimates due to sampling strategy. In total, 59,318 unique cases and 296,590 controls are included in the analysis.[2]

The main variable of interest is the geographic distance between job-switchers' previous and current employer measured by calculating the number of kilometres between the two establishments. Practically, this measure is defined by assigning the latitude and longitude to the centre of each postal code area in which the establishments are located and by calculating the direct distance between the two points.[3] When both previous and

2 The "clogit" command was used in Stata 15 to estimate the parameters.

3 The postal code of an establishment's address is the most exact location information provided by Statistics Finland. The coordinates of the center of the postal code area were retrieved from Statistics Finland's Paavo - The open data by postal code area service.

current employer are located in the same postal code area, the distance between them is calculated as the mean distance of the establishments to the centre of the respective postal code area, weighted by the area's number of employees in knowledge-intensive industries.[4] In the models, geographic distance is logged to account for the fact that the frequency with which employees move from one workplace to the other does not change linearly over geographic space.

The above reviewed proximity literature shows that in addition to geographical proximity, there are also other types of proximities making two organizations interact. Social actors tend to form links with actors who are in some way similar with themselves, and the formation of knowledge ties requires a minimum level of cognitive proximity between two actors (Broekel, 2015). The homogeneity of social networks means that a link between organizations *i* to *j* is more likely to form if they are similar to one another in terms of aggregate statistics summarizing employees' sociodemographic characteristics. Homophilious tie-formation mechanisms are tested by including in the analysis variables measuring similarity of organizational pairs with regard to their gender, age and educational composition.

Similarity in organizational terms may also be of importance for the dynamics of the network. It is more likely that a link will form between two organizations if they belong to the same industry or multi-organizational firm. This is due to the fact that in these cases, the type of work carried out in the organizations is more similar, thus facilitating the mobility of employees between them. Two organizations belonging to the same industry may also be subject to a more similar institutional framework at the macro-level. For these reasons, it can be expected that a link from organization *i* to *j* is more likely to form if the two organizations belong to the same industry or multi-organizational firm or are similar in terms of the capital intensity of their production.

Social proximity is the most complex proximity dimension to operationalize, for it refers to some extent to the overlap of two firms' personnel's personal networks. In practice, this kind of information is rarely available. However, social embeddedness can also be defined through the repetition of network ties. In both social and economic networks, ties that exist at one point in time are likely to be repeated in the future (Rivera et al., 2010). Therefore, in addition to being a measure of the strength of a relationship, repetition can also be viewed as an indicator of trust (Gulati and Gargiulo, 1999) and the social embeddedness of economic action (Uzzi, 1996).

The movements of employees across organizations create channels through which both employers and employees can obtain information about job openings and potential candidates, thus influencing future mobility patterns (Granovetter, 1995; Fernandez et al., 2000; Ioannides and Loury, 2004). As a result, organizational pairs that are directly linked to one another are more likely to form a link at the next point in time. In addition to direct connections, longer chains are also found to matter for tie creation (Granovetter, 1995). In their study on inter-organizational employee mobility, Collet and Hedström (2013) found that movements of employees occur most frequently at geodesic distances of 2 and 3, suggesting that even though at greater distances the number of contacts expands considerably, the circulation of relevant information is very limited. To account for such tie-formation processes endogenous to the network (that is depending on existing network patterns), lagged network proximity variables indicating short path distances between organization *i* and *j* at *t*-1, *t*-2 and *t*-3 are introduced as controls.

Another hypothesized network-related effect is that employees in organization *i* would be more interested in moving to organization *j* if they observed that individuals from oth-

4 Employment-weighted coordinates were calculated based on data obtained from the Register of Enterprises and Establishments maintained by Statistics Finland.

er organizations moved to *j*. As in Collet's and Hedsröm's (2013) study, this hypothesis is tested in the analysis by including a variable indicating the indegree of organization *j* measured as number of individuals who moved to organization *j* from organizations other than *i*.

To control for unobserved regional fixed effects and to study spatial variation in the relationship between geography and tie-formation, dummies for the labour market area and interaction variables between area dummies and the geographical distance variable are introduced. Dummies for labour the market area reflect whether the job-switch was from the Capital Region to other urban municipalities; from the Capital Region to semi-urban or rural areas; from urban municipalities to the Capital Region; within urban municipalities; from urban to semi-urban or rural areas; from semi-urban or rural areas to the Capital Region; from semi-urban or rural to urban areas, or within semi-urban or rural areas, the base category being switches within the Capital Region.

The categories have been classified according to the statistical grouping of municipalities between urban, semi-urban and rural areas developed by Statistics Finland. The grouping divides municipalities into these three categories based on the proportion of people living in urban settlements and the population of the largest urban settlement. For analytical purposes, this classification has been applied in this study so that urban municipalities[5] have been further divided into the municipalities of the Capital Region and other urban municipalities, while semi-urban and rural municipalities have been merged into one group. The first adjustment has been made for the reason that in the municipalities of the Capital Region, which include the major cities of Helsinki, Espoo and Vantaa, employment densities are substantially higher compared to the other urban areas, as well. As Figure 1 illustrates, most of the densest neighbourhoods in terms of number of knowledge-intensive jobs are within the Capital Region.[6] Respectively, the merging of semi-urban and rural municipalities into one group has been done because only a small part of matches take place in the municipalities of either of these groups, as the significance of knowledge-intensive production is lower in these areas.

Financial incentives are also likely to be important for the network analysed, since job mobility decisions can be assumed to be influenced by prospective gains in wage earnings. The probability of a link being formed from organization *i* to organization *j* is therefore expected to be positively related with the wage in *j* and in a negative relationship with the wage in *i*. Finally, the probability that an employee will move between two organizations increases with their sheer workplace sizes. This effect is controlled for by including an estimate of establishment size in terms of number of employees.

5 *Urban municipalities* include those municipalities in which at least 90 per cent of the population lives in urban settlements or in which the population of the largest urban settlement is at least 15,000.

6 This is due to the fact that although compared to the rest of the country the business structure of the region is service oriented, knowledge-intensive production is concentrated mainly in the Capital Region. The municipalities surrounding the Capital Region are specialized in logistics and construction, while the areas on the fringe of the region are more industrial.

Results

A series of conditional logit models were estimated to study the association between geographical proximity and network tie-formation while other effects taken into account. The coefficients from the outcome equation for the models are found in Table 2, and a description of the variables is displayed in Table 1. The dependent variable in the logistic regression models indicates whether there was a direct link between workplaces *i* and *j*. The main variable of interest is the natural logarithm of the geographic distance in kilometres between the two workplaces. The estimation process advances in five steps, so the change in the distance variable can be studied as more controls are included. The matched case-control setting described above means that the matching variable, that is the industrial sector at the 2-digit NACE industry level, is controlled for in all model specifications.

Table 1. Descriptive statistics for variables included in logit models

Variable	Mean	St. dev.
Distance between workplace *i* and *j* [ln (km)]	2.142	1.244
Distance, capital → capital	1.660	1.007
Distance, capital → urban	3.548	0.456
Distance, capital → semi-urban/rural	3.737	0.530
Distance, urban → capital	3.538	0.452
Distance, urban → urban	2.780	1.433
Distance, urban → semi-urban/rural	3.720	0.678
Distance, semi-urban/rural → capital	3.742	0.551
Distance, semi-urban/rural → urban	3.760	0.690
Distance, semi-urban/rural → semi-urban/rural	3.060	1.469
Workplace employee size of *i* (ln)	2.836	1.588
Workplace employee size of *j* (ln)	2.886	1.556
Absolute difference in average age of employees in workplace *i* and *j*	7.101	5.640
Absolute difference in percentage of women in workplace *i* and *j*	0.281	0.228
Absolute difference in average years of schooling in workplace *i* and *j*	1.934	1.562
Average wage in workplace *i* (1 000 EUR)	3.452	5.919
Average wage in workplace *j* (1 000 EUR)	3.479	2.491
Workplace i and j are part of the same multi-organizational firm	0.00891	0.0940
Absolute difference in capital/employee (ln) in firm *i* and *j*	1.362	1.254
Sociometric distance of one between workplace *i* and *j* at *t*-1	0.0289	0.168
Sociometric distance of two between workplace *i* and *j* at *t*-1	0.0532	0.224
Sociometric distance of three between workplace *i* and *j* at *t*-1	0.0965	0.295
Sociometric distance of one between workplace *i* and *j* at *t*-2	0.0204	0.141
Sociometric distance of two between workplace *i* and *j* at *t*-2	0.0417	0.200
Sociometric distance of three between workplace *i* and *j* at *t*-2	0.0748	0.263
Sociometric distance of one between workplace *i* and *j* at *t*-3	0.0161	0.126
Sociometric distance of two between workplace *i* and *j* at *t*-3	0.0336	0.180
Sociometric distance of three between workplace *i* and *j* at *t*-3	0.0594	0.236
Indegree of workplace *j*	8.180	33.514
Capital → urban	0.0800	0.271
Capital → semi-urban/rural	0.0361	0.186
Urban → capital	0.0740	0.262
Urban → urban	0.0203	0.141
Urban → semi-urban/rural	0.00834	0.0909
Semi-urban/rural → capital	0.0261	0.159
Semi-urban/rural → urban	0.00729	0.0851
Semi-urban/rural → semi-urban/rural	0.00486	0.0696
N = 355,908		

Table 2. Conditional logit models of tie-formation

Variable/step	(1)	(2)	(3)	(4)	(5)
Distance between workplace *i* and *j* [ln (km)]	-0.482*** (0.00394)	-0.392*** (0.00500)	-0.335*** (0.00539)	-0.315*** (0.00632)	-0.255*** (0.00671)
Workplace employee size of *i* (ln)		0.521*** (0.00418)	0.380*** (0.00475)	0.383*** (0.00477)	0.387*** (0.00480)
Workplace employee size of *j* (ln)		0.463*** (0.00422)	0.261*** (0.00518)	0.259*** (0.00520)	0.263*** (0.00522)
Absolute difference in average age of employees in workplace *i* and *j*		-0.039*** (0.00124)	-0.033*** (0.00129)	-0.033*** (0.00129)	-0.034*** (0.00130)
Absolute difference in percentage of women in workplace *i* and *j*		-1.434*** (0.0316)	-1.262*** (0.0329)	-1.258*** (0.0330)	-1.256*** (0.0331)
Absolute difference in average years of schooling in workplace *i* and *j*		-0.315*** (0.00492)	-0.277*** (0.00510)	-0.272*** (0.00512)	-0.272*** (0.00513)
Average wage in workplace *i* (1 000 EUR)		-0.054*** (0.00406)	-0.049*** (0.00423)	-0.042*** (0.00423)	-0.041*** (0.00423)
Average wage in workplace *j* (1 000 EUR)		0.025*** (0.00403)	0.023*** (0.00422)	0.025*** (0.0424)	0.026*** (0.00425)
Workplace *i* and *j* are part of the same multi-organizational firm		5.315*** (0.1079)	4.635*** (0.1149)	4.624*** (0.1154)	4.663*** (0.1158)
Absolute difference in capital/employee (ln) in firm *i* and *j*		-0.035*** (0.00555)	-0.049*** (0.00591)	-0.051*** (0.00593)	-0.050*** (0.00595)
Sociometric distance of one between workplace *i* and *j* at *t*-1			2.989*** (0.0501)	2.976*** (0.0503)	2.967*** (0.0504)
Sociometric distance of two between workplace *i* and *j* at *t*-1			1.096*** (0.0254)	1.095*** (0.0254)	1.087*** (0.0255)
Sociometric distance of three between workplace *i* and *j* at *t*-1			0.406*** (0.0189)	0.407*** (0.0189)	0.403*** (0.0190)
Sociometric distance of one between workplace *i* and *j* at *t*-2			1.896*** (0.0602)	1.878*** (0.0603)	1.863*** (0.0604)
Sociometric distance of two between workplace *i* and *j* at *t*-2			0.473*** (0.0292)	0.471*** (0.0292)	0.469*** (0.0293)
Sociometric distance of three between workplace *i* and *j* at *t*-2			0.107*** (0.0213)	0.111*** (0.0213)	0.112*** (0.0214)
Sociometric distance of one between workplace *i* and *j* at *t*-3			1.337*** (0.0661)	1.332*** (0.0663)	1.312*** (0.0668)
Sociometric distance of two between workplace *i* and *j* at *t*-3			0.279*** (0.0317)	0.281*** (0.0318)	0.281*** (0.0319)
Sociometric distance of three between workplace *i* and *j* at *t*-3			-0.051* (0.0232)	-0.049* (0.0233)	-0.047* (0.0233)

Table 2. Conditional logit models of tie-formation (*continues...*)

Variable/step	(1)	(2)	(3)	(4)	(5)
Indegree of workplace *j*			0.006*** (0.0002)	0.006*** (0.0002)	0.006*** (0.0002)
Capital → urban				-0.161*** (0.0310)	1.031*** (0.183)
Capital → semi-urban/rural				-0.576*** (0.0531)	1.089*** (0.297)
Urban → capital				-0.0169 (0.0311)	0.823*** (0.187)
Urban → urban				0.849*** (0.0430)	2.414*** (0.0937)
Urban → semi-urban/rural				0.217** (0.0773)	3.413*** (0.402)
Semi-urban/rural → capital				-0.0357 (0.0511)	1.125*** (0.293)
Semi-urban/rural → urban				0.424*** (0.0773)	3.407*** (0.400)
Semi-urban/rural → semi-urban/rural				1.097*** (0.0789)	3.438*** (0.207)
Interaction of distance variable and capital → urban variable					-0.375*** (0.0525)
Interaction of distance variable and capital → semi-urban/rural variable					-0.496*** (0.0819)
Interaction of distance variable and urban → capital variable					-0.274*** (0.0533)
Interaction of distance variable and urban → urban variable					-0.637*** (0.0331)
Interaction of distance variable and urban → semi-urban/rural variable					-0.951*** (0.116)
Interaction of distance variable and semi-urban/rural → capital variable					-0.349*** (0.0791)
Interaction of distance variable and semi-urban/rural → urban variable					-0.865*** (0.112)
Interaction of distance variable and semi-urban/rural → semi-urban/rural variable					-0.874*** (0.0683)
Number of observations	355,908	355,908	355,908	355,908	355,908
Log likelihood	-98 416	-63 597	-56 657	-56 264	-55 844
Pseudo R^2	0.0740	0.402	0.467	0.471	0.475

The results of the baseline estimation implicate that labour flow networks generated by job-switches are geographically highly concentrated. The unadjusted odds ratio, that is the exponentiated value of the logistic regression coefficient not adjusted for confounders, for the variable measuring geographic distance between workplace *i* and *j* is 0.62 (≈ $e^{-0.482}$). This suggests that a one per cent increase in distance between previous and current employer is associated with a 38% decrease in the odds of forming a link between the organizations. However, in addition to spatial clustering, the observed network dynamics are also the outcome of several other tie-formation processes, which must be taken into account.

The second step controls for basic observables, such as estimates of establishment size in terms of number of employees, an array of variables describing proximity in terms of differences in age, schooling and gender compositions of organizations, as well as covariates measuring financial incentives for individuals to move from organization *i* to *j*. This step also includes variables examining the role of proximity at the organizational level; that is a covariate describing if the establishments are part of the same multi-organizational firm, and a variable measuring similarity in terms of capital intensity indicated by the natural logarithm of capital per worker in the firm.

The variables added in the second step do indeed explain a substantial part of the tie-formation processes. All the coefficients of the covariates added are highly significant and in the expected directions. Adding these controls also improves significantly the fit of the model as measured by the pseudo R^2. The parameter estimates of the size of organization *i* and *j* are both positive. Furthermore, the results suggest proximity at the organizational level matter: if organizations are part of the same multi-organizational firm, it is more likely that a tie will be formed between them. Similarity in terms of capital intensity on the firm-level also has a positive relationship with tie-creation. The results suggest that financial incentives are important, too. The probability of a tie from *i* to *j* is in a negative relationship with average earnings in *i* and in a positive relationship with the pay level in *j*. However, after controlling for basic observables capturing processes that could possibly affect an employee's movement between organizations, the magnitude of the coefficient of the geographic distance variable, which is the main variable of interest, remains important from a substantive point of view and highly significant.

Step three controls for tie-formation processes endogenous to the network by including dummy variables for lagged sociometric distances between establishments that were directly linked at time *t*. Sociometric distances are calculated along the shortest path from *i* to *j* or from *j* to *i* at time points *t*-1, *t*-2 and *t*-3. The dummies indicate whether the establishments had a lagged sociometric distance of one, two or three, path lengths greater than three or infinite[7] acting as reference category. In the estimation, only the effects of shorter paths are accounted for, because previous research studying information flows in social networks has shown that connections are rarely formed at path lengths of four or greater (Granovetter, 1995; Singh, 2005; Sorenson et al., 2006; Collet and Hedström, 2013).

As expected, dummies indicating shorter lagged sociometric distances are positively related to the probability of a link being formed at time *t*. The positive coefficient of sociometric proximity is lower at *t*-2 than at *t*-1 and is still lower at *t*-3 than at *t*-2, but remains highly significant until the sociometric distance of two at *t*-3. The results suggest that network proximity is an important factor explaining tie-formation processes. For example, organizations that were directly linked at *t*-1 had about 20 times greater odds of forming a tie at time *t* than non-connected organizations or organizations that were at path length of four or greater. Endogenous tie-creation processes are also correlated with

7 A path between two nodes is infinite, if the nodes are disconnected, that is there is no path between them.

spatial clustering, for the adding of variables measuring sociometric distances decreases somewhat the coefficient of the geographic distance variable (from -0.392 to -0.335).

The third estimation investigating the relationship between network-related covariates and tie-formation includes further a variable indicating the indegree of establishment *j* measured as the number of individuals who moved there from establishments other than *i*. This variable also has the expected highly significant positive association with tie-formation, suggesting that employees tend to move to organizations attracting individuals also from other organizations. The adding of these network-related controls raises the pseudo R^2 to a reasonably good level (47%), and the variables included in this step clearly explain a substantial part of tie-formation processes. However, geographic proximity still seems to have a substantial and highly significant independent positive association with the probability of forming a tie by the movement of employees between organizations.

To control for unobserved regional fixed effects, the fourth and the final step studying the average statistical relationship of geographical proximity and tie formation probability adds dummy variables describing the within-region geographical locations of the establishments forming the organizational pairs.[8] The internal dynamics of firms can vary regionally, as, for example, costs based on the price of land affect the location decision of firms through the selection of less profitable firms into lower and more profitable firms into higher cost areas.

After step four the odds ratio for the variable measuring geographic distance between workplaces forming the organizational tie is 0.73 ($\approx e^{-0.315}$). This means that when all control variables are accounted for, a one per cent increase in distance between the previous and the new employer is associated with a 27% decrease in the odds of forming a link between the organizations. However, the results of the fully specified model suggest that in addition to geographical proximity, the observed network dynamics are also the outcome of several other proximity processes. As expected, the larger the differences in terms of the gender composition, average age and average years of schooling in two organizations, the less likely it is that a tie will be formed between them.

The comparison of the odds ratios of the proximity variables implies that although all proximity processes controlled for in the model have a substantial role in the formation of network ties, particularly educational differences and geographical distance seem to be important for the inter-organizational network. As can be seen in Table 1, a typical variation as measured by a standard deviation in the age difference variable is 5.64, in the educational difference variable 1.56, in the geographical distance variable 1.24, and 0.228 in the gender composition difference variable. This means that the odds ratio for a typical variation in age composition is about 0.83 (0.97^5.64), compared to the odds ratio of 0.65 for a typical variation in average level of education, 0.68 for a typical variation in geographical distance and 0.75 for a typical variation in gender composition. Although it is difficult to compare the relative importance of different kinds of variables, the comparison above implies that geographical proximity seems to be important from a substantive point of view in relation to other types of proximities influencing network dynamics.

To study regional variation in the relationship between geographical proximity and tie-formation, the fifth step adds interaction variables of regional dummies and the variable measuring distance between workplace *i* and *j*. The aim of this step is to test wheth-

8 The regression was also run with controls measuring more specific locational characteristics, such as distances of the workplaces to the central business district of the region and to the employment centers of local municipalities, employment densities of the postal code areas and local clustering measured by the number of neighborhoods employing workers in the same industry relative to the total number of neighborhoods with economic activity. However, these variables show a high correlation with each other and with the distance measure, raising concerns with multicollinearity, and are thus excluded from the model. See Larsson (2017) on relations between job-switching distances, local density and occupation clustering.

er strong clustering tendencies as evidenced by job-switching patterns are characteristic to dense urban regions, or if labour flows cluster also between and outside of urban centres, or across longer distances in areas classified as semi-urban or rural. In model 5, the coefficient of the distance variable gives the estimate of the role of geographical distance in the base category of the regional dummies; that is amongst organizational pairs in which both *i* and *j* are located in the Capital Region. In Table 2, the coefficients of the interaction variables give the estimates of the distance variable in different regional groups relative to the organizational pairs of the base category.

The results of model 5 suggest that the coefficients of the distance variable are highly significantly smaller in all regional groups than in the base category. The comparison of the odds ratios of the interaction variables suggest that geographical distance is of the greatest consequence for the formation of network ties when the movement of employees occurs in less dense areas, that is between urban and semi-urban or rural and within semi-urban or rural areas. In these groups the odds ratios for the geographical distance variable are about 0.3, compared to the odds ratio of 0.77 in the base category. This finding indicates that relative to the surrounding areas, employees find new jobs over more concentrated areas of land when changing jobs outside the Capital Region or between other areas and the Capital Region.

It should be noted that the areas of the classification differ significantly in terms of geographical scope, and therefore the average distances moved by employees as well as potential distances between establishments vary greatly depending on the type of area. The average switching distance is 6 kilometres within the Capital Region, 12 kilometres within semi-urban or rural areas and in the case of urban-to-urban switches, while in the rest of the categories it varies between 32 and 46 kilometres. In the more densely built area of the Capital Region, ties are formed also between establishments located relatively far from each other, while in more extensive areas the possible distances of potential dyads become quite long, and ties are very rarely realized between establishments far apart. This regional diversity makes it difficult to assess the within-region variation of the coefficient of the geographical proximity variable from a substantive point of view. However, the logarithmic transformation of the distance variable makes it more comparable across different kinds of areas, as it reduces the effect of outliers by making the distribution less skewed and gives information on relative instead of absolute changes in the covariate. The main conclusion derived from these results is that knowledge intensive production shows clustering tendencies across longer distances outside of major cities and in semi-urban and rural areas as well, indicating that organizations performing tasks intensive in the use of knowledge and in which interaction is essential benefit from relatively shorter distances also outside high-density urban areas.

Conclusions

This paper examines the micro-foundations of agglomeration economies by studying the role of geography in a network created by the within-region mobility of employees across local organizations. Furthermore, the paper investigates how the role of physical proximity varies across the regional landscape, thereby analysing the relations between different areas of the region. It has been pointed out in the research literature on agglomeration economies that geographical proximity is not in itself a sufficient condition for the flow of information and knowledge, as economic agents can be co-located without forming direct linkages to each other. Transfer of knowledge requires active participation in knowledge exchange networks, and hiring employees from rival firms, business partners or other firms can be a key means to gain access to such networks (Breschi and Lissoni, 2009). From this point of view, causes of the concentration of economic networks can be seen to be related more to agglomeration benefits stemming from the functioning of labour markets than to the flow of information facilitated by physical proximity. However, the mobility of labour force and shared labour markets, which are found to be central in the mediation of knowledge, seem to have a strong spatial dimension. A series of conditional logit models examining the importance of various tie-formation processes show that the forming of a link by inter-organizational employee mobility is at all regional levels more likely between organizations geographically closer to each other, all other things being equal.

This finding is consistent with arguments about how matching between employers and employees does not only occur through the price mechanisms of labour markets, but the search for job opportunities is in addition determined by social processes requiring close personal contacts (Granovetter, 1995). Labour mobility between organizations links networks, thus creating social cohesion amongst firms changing personnel. This process generates social ties, which can further increase mobility. Social processes conditioning a search for jobs can be spatially embedded, as short distances are found to promote direct contacts, thus facilitating the exchange of knowledge more easily transferred by face-to-face communication (Storper and Venables, 2004). This may be particularly important in well-networked knowledge-intensive service sectors, in which informal interaction plays a key role (see Arzaghi and Henderson, 2008). Although knowledge networks are no longer necessarily geographically limited, due to advanced information and communication technologies, physical proximity still seems to have significance for economic activity.

However, in addition to organizational relations, labour mobility depends also on individual-level factors of employees changing jobs. One of the reasons employees tend to move locally can be, for example, the sunk costs of residence location (Breschi and Lissoni, 2001). There is also evidence that social interactions among neighbours affect labour market outcomes, such as the propensity to work together (Bayer et al., 2008). A subject of further research could therefore be how housing choices are associated to within-region employee mobility, which could not be examined in the setting of the present research. Including distances from previous and new employer to home as covariates would not only allow controlling for the effect of housing choices on employee mobility and job-switching distances but would also enable to study how the effect varies depending on the community structure of the area. Further research could also look at how mixing residential and employment uses within urban concentrations is related to localized job-switching rates, as both uses can support neighbourhood level development activities by contributing to creating a critical mass of activity.

The finding that shorter distances within a region is, amongst other mechanisms, a powerful predictor of tie-formation events corroborates the notion on the sharp attenuation of knowledge spillovers established in the empirical literature on human capital externalities. Particularly, economic activities intensive in the use of knowledge as an input are shown to have higher returns to local density (Arzaghi and Henderson, 2008; Rosenthal and Strange, 2008; Andersson et al., 2009; Larsson, 2017). These studies have shown the magnitude of the effect of non-market interaction to vary regionally so that they are the strongest in dense urban environments and depreciate quickly with distance. However, the present study demonstrates that knowledge-intensive industries show strong clustering tendencies as evidenced by job-switching patterns as well in lower density areas and across longer distances between different parts of the region. This finding implicates that organizations carrying out production in which human capital externalities are essential benefit from relatively shorter distances also in semi-urban and rural areas where they are locating, on average, more remotely to each other.

The physical aspects of labour market dynamics reported in this study implicate that there is a direct link between the functioning of labour markets and the region-wide coordination of land use planning and zoning. Although firms make their location choices and form their ties to other firms according to their own business criteria, municipalities and other relevant public agencies can, in cooperation with builders and developers, create conditions for the co-location and transit accessibility of businesses. The results of the study suggest that the promotion of existing business districts towards denser and more compact concentrations is essential in terms of the operating conditions of knowledge intensive production. The further intensification of concentrations also includes the infill and redevelopment of old structures in order that the supply of offices can adapt to changing demands of businesses in attractive business areas. From the perspective of region-wide development, it is also important to take advantage of investments in rail connections and the potential of well accessible transport station environments as locations for offices as well as to enhance the inter-municipal coordination of land use and transit. These issues are related directly to the future growth prospects of regions as the significance of knowledge in production processes increases.

Funding

This study has received funding from the European Union's Horizon 2020 research and innovation programme (RUR-01-2016) under grant agreement No 727988.

References

Abbott A (1997) Of time and space: The contemporary relevance of the Chicago school. Social Forces 75(4): 1149–1182.

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

Andersson R, Quigley JM and Wilhelmsson M (2009) Urbanization, productivity and innovation: Evidence from investment in higher education. Journal of Urban Economics 66(1): 2–15.

Arzaghi M and Henderson JV (2008) Networking off Madison Avenue. The Review of Economic Studies 75(4): 1011–1038.

Audretsch D (1998) Agglomeration and the location of innovative activity: Oxford Review of Economic Policy 14(2): 18–29.

Bayer P, Ross SL and Topa G (2008) Place of work and place of residence: Informal hiring networks and labour market outcomes. Journal of Political Economy 116(6): 1150–1196.

Boschma RA (2005) Proximity and innovation: a critical assessment. Regional Studies 39(1): 61–74.

Boschma RA, Eriksson RH and Lindgren U (2014) Labour market externalities and regional growth in Sweden: the importance of labour mobility between skill-related industries. Regional Studies 48(10): 1669–1690.

Breschi S and Lissoni F (2001) Knowledge spillovers and local innovation systems: A critical survey. Industrial and Corporate Change 10(4): 975–1005.

Breschi S and Lissoni F (2009) Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. Journal of Economic Geography 9(4): 439–468.

Broekel T (2015) The co-evolution of proximities – A network level study. Regional Studies 49(6): 921–935.

Collet F and Hedström P (2013) Old friends and new acquaintances: Tie formation mechanisms in an interorganizational network generated by employee mobility. Social Networks 35(3): 288–299.

Combes P and Duranton G (2006) Labour pooling, labour poaching and spatial clustering. Regional Science and Urban Economics 36(1): 1–28.

Corredoira RA, Rosenkopf L (2010) Should auld acquaintance be forgot? The reverse transfer of knowledge through mobility ties. Strategic Management Journal 31(2): 159–181.

Durlauf SN (2004) *Neighborhood effects*. Handbook of Regional and Urban Economics 4: 2173–2242.

Eriksson RH and Lindgren U (2009) Localized mobility clusters: impacts of labour market externalities on firm performance. Journal of Economic Geography 9(1): 33–53.

Fernandez R, Castilla E and Moore P (2000) Social capital at work: Networks and employment at a phone center. American Journal of Sociology 105(5): 1288–1356.

Granovetter M (1985) Economic action and social structure: The problem of embeddedness. American Journal of Sociology 91(3): 481–510.

Granovetter M (1995) *Getting a job: A study of contacts and careers*. Chicago: University of Chicago Press.

Gulati R and Gargiulo M (1999) Where do interorganizational networks come from? American Journal of Sociology 104(5): 1439–1493.

Hosmer DW, Lemeshow S and Sturvidant RX (2013) *Applied Logistic Regression*, 3rd ed. New York: Wiley.

Howells JRL (2002) Tacit knowledge, innovation and economic geography. Urban Studies 39(5–6): 871–884.

Ioannides YM and Loury LD (2004) Job information networks, neighborhood effects, and inequality. Journal of Economic Literature 42(4): 1056–1093.

Larsson JP (2017) Non-routine activities and the within-city geography of jobs. Urban Studies 54(8): 1808–1833.

Lengyel B and Eriksson RH (2017) Co-worker networks, labour mobility and productivity growth in regions. Journal of Economic Geography 17(3): 635–660.

Maliranta M, Mohnen P and Rouvinen P (2008) Is inter-firm labor mobility a channel of knowledge spillovers? Evidence from a linked employer-employee panel. Industrial and Corporate Change 18(6): 1161–1191.

Marshall A (1920) *Principles of Economics – An Introductory Volume*. London: MacMillan.

Morgan K (2004) The exaggerated death of geography: learning, proximity and territorial innovation systems. Journal of Economic Geography (4)1: 3–21.

Nooteboom B, Van Haverbeke W, Duysters G, Gilsing V and van den Oord A (2007) Optimal cognitive distance and absorptive capacity. Research Policy 36(7): 1016–1034.

Official Statistics of Finland (OSF): Regional statistics on entrepreneurial activity [e-publication]. ISSN=2342-6268. Helsinki: Statistics Finland [referred: 13.6.2019]. Access method: http://www.stat.fi/til/alyr/tau_en.html

Ponds R, Van Oort F and Frenken K (2007) The geographical and institutional proximity of research collaboration. Papers in Regional Science 86(3): 423–443.

Power D and Lundmark M (2004) Working through knowledge pools: Labour market dynamics, the transference of knowledge and ideas, and industrial clusters. Urban Studies 41(5–6), 1025–1044.

Rallet A and Torre A (1999) Is geographical proximity necessary in the innovation networks in the era of global economy? GeoJournal 49(4): 373–380.

Rao H and Drazin R (2002) Overcoming resource constraints on product innovation by recruiting talent from rivals: A study of the mutual fund industry, 1986–94. Academy of Management Journal 45(3): 491–507.

Rivera M, Soderstrom S and Uzzi B (2010) Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms. Annual Review of Sociology 36(1): 91–115.

Rosenthal SS and Strange WC (2008) The attenuation of human capital spillovers. Journal of Urban Economics 64(2): 373–389.

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

Somaya D, Williamson IO and Lorinkova N (2008) Gone but not lost: The different performance impacts of employee mobility between cooperators versus competitors. Academy of Management Journal 51(5): 936–953.

Song J, Almeida P and Wu G (2003) Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? Management Science 49(4): 351–365.

Sorenson O, Rivkin JW and Fleming L (2006) Complexity, networks and knowledge flow. Research Policy 35(7): 994–1017.

Sorenson O and Stuart TE (2008) Bringing the context back in: Settings and the search for syndicate partners in venture capital investment networks. Administrative Science Quarterly 53(2): 266–294.

Storper M and Venables AJ (2004) Buzz: face-to-face contact and the urban economy. Journal of Economic Geography 4(4): 351–370.

Uzzi B (1996) The sources and consequences of embeddedness for the economic performance of organizations: the network effect. American Sociological Review 61(4): 674–698.

Yhdyskuntarakenteen seurannan aineistot (YKR) Finnish Environment Institute SYKE and Statistics Finland, 2014.

Appendix 1

Table A1.1. List of included industries according to the NACE Rev. 2 classification

582	Software publishing
62	Computer programming, consultancy and related activities
63	Information service activities
72	Scientific research and development
69	Legal and accounting activities
70	Activities of head offices; management consultancy activities
71	Architectural and engineering activities; technical testing and analysis
73	Advertising and market research
74	Other professional, scientific and technical activities
78	Employment activities
82	Office administrative, office support and other business support activities
8532	Technical and vocational secondary education
854	Higher education
8559	Other education n.e.c.
856	Educational support activities
24	Manufacture of basic metals
25	Manufacture of fabricated metal products, except machinery and equipment
26	Manufacture of computer, electronic and optical products
27	Manufacture of electrical equipment
28	Manufacture of machinery and equipment n.e.c.
29	Manufacture of motor vehicles, trailers and semi-trailers
30	Manufacture of other transport equipment
33	Repair and installation of machinery and equipment
43292	Installation of lifts and escalators
20	Manufacture of chemicals and chemical products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations