IBS WORKING PAPERS

Co-worker networks, productivity growth and cross-industry links in regions

(This version: January 12, 2015)

Working Paper 3/2015
http://web.ibs-b.hu/research

INTERNATIONAL BUSINESS SCHOOL

Záhony u. 7.

1031 Budapest, Hungary

Balázs Lengyel [a,b], Rikard Eriksson [c]

[a] Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Budaörsi út 45, 1112 Budapest, Hungary. Corresponding author: lengyel.balazs@krtk.mta.hu

[b] International Business School Budapest, Záhony utca 7, 1031 Budapest, Hungary

[c] Department of Geography and Economic History, Umeå University, SE-901 87 Umeå, Sweden

Abstract

This paper provides a new empirical perspective for analysing the roleof social networks in regional economic growth by constructing large-scale networks from employee-employee co-occurrences in plants in the entire Swedish economy 1990-2008. It is assumed that employees in different plants know each other if they have been engaged in a co-worker relationship previously. We argue that these personal acquaintances are important for local learning opportunities and consequently for regional productivity growth. Moreover, the pattern of cross-industrial network is likely to characterize the knowledge externalities of regions. The paper provides the first systematic evidence for a central claim in economic geography: social network density has positive effect on regional productivity growth. In a further step, we demonstrate that the co-worker network across industries differs from the skill-relatedness network now frequently used in economic geography. Social ties concentrate within same industries but this is increasingly true in small regions; while the share of edges across skill-related industries as well as across unrelated industries is higher in large metropolitan areas. A key finding suggests that social relatedness of industries have higher importance in middle-sized and small regions.

JEL codes: D85, J24, J61, R11, R23

Keywords: social network, probability and strength of tie, regional productivity growth, panel regression, relatedness

Acknowledgment: Help in geo-localization of the data was received from Magnus Strömgren. We thank Gábor Békés, Urban Lindgren and László Lőrincz for useful comments.

1. Introduction

Following Marshall (1920) there is a general agreement in economic geography and related fields that the agglomeration of economic activities is essential for understanding regional innovation and growth. In this respect, face-to-face interaction is increasingly emphasized as essential for why proximity still is crucial for sustaining learning and innovation (Storper and Venables, 2004), and that more dense environments enhance the probability of "learning by seeing" (Glaeser, 2000). Human interaction and the social networks created thereof are thus expected to be key drivers behind regional economic growth. This is basically because the effectiveness of learning and co-operation of individuals are enhanced by personal relations and this is expected to have both direct and indirect effects on productivity growth since firms gain extra benefits as well when accessing external knowledge through social ties. However, despite the above theoretical claims on the role of face-to-face contacts and social networks for learning and growth, very little empirical work has actually been devoted to analysing the role of social networks on regional productivity growth. Instead, scholars tend to proxy the socializing potential of regions by means of population density or industrial structure (Ciccone and Hall 1996, Glaeser 1999), and almost take the relation between density and social interaction for granted by assuming that the mere concentration of skilled workers automatically will increase the probability for social interaction and thus enhance learning and growth.

To address this potential shortcoming in the existing literature, the aim of this paper is to assess to what extent co-worker networks influence productivity growth in 72 Swedish labour market regions 1990-2008. This is made possible by a unique longitudinal matched employer-employee database from which we construct a social network of employees based on their co-occurrence at workplaces and analyse the effect of the network on regional dynamics. These type of networks are frequently called co-worker networks in labour economics and scholars assume that two employees know each other when they have worked in the same workplace simultaneously in a certain period of their career (for an overview see Beaman and Jeremy 2012). Evidence shows that information flow through these co-worker relations help people find better jobs and reduce unemployment time when dismissed (Calvo-Armengol and Jackson 2004, Glitz 2013, Granovetter 1995, Hensvik and Nordström Skans 2013). Given that the exchange of information and knowledge between workers and firms promotes the emergence and diffusion of innovation and subsequent productivity (Duranton and Puga, 2004), we claim that co-worker networks are important sources of regional economic dynamics. This is because valuable information flow more efficiently through co-worker relations and employees might learn more efficiently in dense co-worker networks as compared to the technological externalities assumed to be residing “in the air” of agglomerations (c.f. Breschi and Lissoni, 2009, Eriksson and Lindgren, 2009; Huber, 2012).

We claim to make two contributions to the existing literature. First, we develop a new probability measure of workplace-based acquaintance, building on the literature of homophily-biased random networks (Buhai and van der Lei 2006, Currarini et al. 2009). We calculate tie probability using the concept of baseline homophily and rank employee co-occurrence according to this probability. Then, we calculate the strength of individual ties for a selected number of most probable co-workers by counting years of co-working and assume that this tie strength decays over time after the termination of co-worker status. As result, we get a dynamically changing social network, with many weak and few strong ties. Despite co-worker networks and labour mobility networks are presumably interconnected because people establish new links in the co-worker network through mobility from one firm to another (Collet and Hedström 2012), we illustrate in details that our approach differs from previous labour mobility studies (e.g., Breschi and Lissoni, 2009: Eriksson and Lindgren, 2009) –and consequently from recent skill-relatedness approaches (cf Neffke and Henning, 2013) – in three basic aspects.

The second contribution is that this paper provides the first empirical evidence that the density of the social network has a positive effect on regional productivity growth as compared to population density per se. Furthermore, we find that majority of co-worker ties remain within same industries but this is increasingly true in small regions. The share of edges across skill-related as well as unrelated industries is higher in large metropolitan areas; a key finding suggests that social relatedness of industries have higher importance in middle-sized regions.

2. Literature, hypothesis and expectation

The spatial dimension of network-related learning is a core interest of economic geography (Bathelt and Glückler 2003). It is well understood now that transaction costs are diminished by physical proximity as well as personal connections, which enhance the efficiency of mutual learning (Borgatti et al. 2009, Maskell and Malmberg 1999). It is also claimed that most of the learning processes occur within certain spatial proximity despite distant, and presumably weak, ties might provide new knowledge not accessible in the region (Bathelt et al. 2004, Glückler 2007). We also understand that not the social network per se but its' interplay with industry structure is crucial for learning because cognitive, institutional, and organizational proximities are very important for mutual understanding (Boschma 2005, Sorensen et al. 2006). Despite the central interest, our knowledge about the network effect on regional productivity growth is still limited, which is partly due to data access difficulties. Our paper aims to contribute to the literature in this regard by constructing and analysing a large-scale co-worker network. The argument stresses two points: first, the network density is very important for regional productivity growth; and second, the industry-wise structure of the network varies according to the size of the region.

Regional productivity growth has been repeatedly found to depend on population density. This is because spatial agglomeration facilitates the sharing of common facilities, increase the chances of a productive job-worker matching, and enhances interactive learning through the concentration of firms and workers (Duranton and Puga, 2004), which has a direct effect on productivity growth differences (Ciccone and Hall 1996, Glaeser 1999). We argue that looking at not only the co-location of individuals but investigating also the density of social networks will improve our understanding because face-to-face relations and personal acquaintance are important for knowledge sharing (Storper and Venables 2004). As argued by Glaeser (2000) workers in dense environments are more likely to acquire human capital through learning by seeing which make dense regions more productive as well as more attractive for skilled workers with large potential returns for learning which will further increase productivity. Workplaces and consequently the co-worker networks that bind workplaces together are major fields of such knowledge sharing even after the termination of the co-worker relation because people maintain their professional contacts over time and might even follow the career of former colleagues in order to map out the knowledge-base they have potential access to (Dahl and Pedersen, 2003). Thus, co-worker network are important for local learning and consequently on regional productivity growth.

H1: Density of the local co-worker network enhances regional productivity growth.

The hypothesis is not only a further step in understanding spatial learning processes; it also refers to a central debate in the social networks literature. Network density has been considered as a major indicator of social capital for decades in sociology (Burt 1992, Coleman 1990, Walker et al. 1997, Wasserman and Faust 1994) because the closure of social relations enhances trust, authority and sanctions among local actors, all of which supports learning from contacts. However, density alone does not sufficiently describe the full horizon of information-flow tendencies in a network; the strength of social ties is a crucial factor and results in two fundamental processes (Granovetter 1973). On the one hand, people trace strong ties frequently, which offers the possibility of incremental innovation and increase in individual productivity because they learn effectively from each other (Borgatti and Cross 2003). On the other hand, weak ties and the presence structural holes among separated sub-networks offers access to new information and combination of non-redundant knowledge can lead to radical innovations (Ahuja 2000).

Similar ideas to the network-related learning have been present in the economic geography literature (for an overview see Ter Wal and Boschma 2009); one can associate the classical debate between MAR and Jacobsian externalities to the structure of social networks (e.g., Porter 2003; Glaeser et al. 1992, Henderson et al. 1995). For example, strong social ties within certain sectors in specialized industrial districts enhance incremental innovation and productivity growth (Amin, 2000, Asheim 1996, Malmberg 1997), whereas diverse networks across industries in urban areas are associated with potential new combinations of information, creation of new knowledge and radical innovation (Feldman 1999). However, the emerging literature of evolutionary economic geography suggests that spatial learning depends on a complex combination of various proximity dimensions between individual firms and that not only regional specialization or diversity per se but relatedness between co-located industries determine regional productivity growth (Boschma 2005, Frenken et al. 2007).

A growing number of papers look at spatial labour mobility links between firms and industries (Almeida and Kogut 1999, Eriksson and Lindgren 2009) and assess the effect of related labour flows on regional and firm dynamics (Boschma et al. 2009, Timmermans and Boschma 2014). In this literature, two industries are considered as skill-related when the observed labour flow is higher than expected based on industry characteristics because very similar employee skills are needed in these industries (Neffke and Henning 2013). Apart from improving the potential regional matching of skills, Boschma et al (2014) show that high concentrations of skill-related flows in a region strongly influence productivity growth in Sweden due to the production complementarities produced by such labour market externalities. Despite the methodological differences, our co-worker approach is closely connected to the labour mobility approach and we assume that former colleagues maintain their relations even after moving from one workplace to another, which is a proposition often made in labour economics and even in evolutionary economic geography as well (Boschma and Frenken 2011). Despite the lasting characteristics of co-inventors have been found important for later patenting collaborations (Agrawal et al. 2006, Breschi and Lissoni 2009), this paper is the first attempt to analyze co-worker networks in economic geography. Thus, we provide information on how the network structure varies according to regional size and industry structure in the region by testing two additional hypotheses.

> *H2: Co-worker networks dominated by skill-relatedness and across unrelated industries are typically present in highly populated regions.*
>
> *H3: Specialized co-worker networks are more prevalent in small regions; whereas social-relatedness solely based on co-worker networks prevail in middle-sized regions.*

3. Methodology

We propose that employee *i* and employee *j* working for in the same workplace at the same period of time know each other with probability P_{ij} $[0,1]$ and maintain a tie L_{ij} with strength W_{ij} even after the termination of the co-workership. For practical reasons, we select the most probable 50 co-workers of highest P_{ij} for each employee in each year and trace these co-occurrences over the full period and look at those L_{ij} when employee *i* and employee *j* work for two different firms. Then, W_{ij} is given by the length of time period they had worked together decayed by the length of time after the termination of their co-workership. Formulation is as follows.

Probability calculation starts from the assumption of random tie formation at workplaces, which means that a tie between every pair of employees is established with equal probability. Intuition suggests that the larger workplace the less likely that employees know each other. Thus, we first set tie probability proportional to the size of workplace. However, this tie probability creates a large fraction of isolated ties in random network simulations, which is not our intention. Therefore, we use the probability threshold where isolated nodes tend to disappear in a random network setting (Jackson, 2008) and formulate random probability (P_{ij}^{r}) by

$$P_{ij}^{r} = \frac{\ln N}{N}; \qquad (1)$$

where N is the number of employees in the workplace.

In a second step, we consider that individual similarity increases the probability of tie formation, which is called homophily in the large range of social sciences (for an overview see McPherson et al. 2001). It has been shown repeatedly that much more friendship ties are formed across those individuals who are similar in terms of age, gender, race, education, occupation etc. than expected by random tie establishment (Granovetter 1995, Lincoln and Miller 1979, McPherson and Smith-Lovin 1987, Sias and Cahill 1998). Two types of homophily are distinguished in the literature: baseline homophily and inbreeding homophily. Baseline homophily means that individual choice of selecting friends is generated by the structure of the group because the larger subgroup of similar individuals the larger possibility of choosing similar friends. Thus, baseline homophily (H_b) can be measured by the share of subgroup in the firm by

$$H_b = \frac{N_m}{N}; \qquad (2)$$

where N_m denotes the size of the subgroup characterized by feature *m*.

We will assume that H_b influences P_{ij} because relations are more likely between those employees who are of similar age and sex and have the similar educational background. However, Currarini et al. (2009) showed that friendship ties usually exhibit larger homophily than H_b due to additional inbreeding homophily and individuals' choice is even more biased towards akin. Thus, using H_b we will most likely underscore the real probability of the tie between co-workers. We define employee characteristics like age, gender, and education as those subgroup features that are expected to increase tie probability then we can calculate H_b in a repetitive manner as explained above.

In the third step, we have to realize that the size of the subgroups – defined by employee characteristics – has a similar effect on tie probability than the firm size itself. Thus, we have to diminish the probability by $ln(N_m \ / \ N)$ in each case when employee *i* and *j* are similar.

Finally, we simply sum the probabilities calculated from firm size, baseline homophilies and group size effects in order to get probability of co-worker ties (Buhai and van der Lei 2006). Probability is formulated as

$$P_{ij} = \frac{\ln N}{N} + \sum_{G=1}^{M} \left(\frac{\ln N_m}{N_m} / \frac{N_m}{N} \right) \times \delta_{ij}; \quad (3)$$

where G ∈ $\{1, 2, \dots M\}$ denotes those characteristics we use for similarity measurement, N denotes plant size, N_m denotes subgroup size according to feature *m* and δ_{ij} equals 1 if employee *i* and *j* are similar according to feature *m* and 0 otherwise.

We maximize co-worker tie probability at 1, rank co-workers for every employee and follow the 50 most probable co-workers of every employee over time. The time spent together is claimed to be the most important factor of tie strength (Granovetter 1973); the length of the co-worker relationship increases tie strength. Marsden and Campbell (1984, p. 488) postulated that *"returns in terms of tie strength to increased duration of a relationship decline with increasing length of acquaintance"* and suggest to use natural logarithm of years of co-worker relation to index duration. Since we look at the ties between employees working for two distinct firms, we first count the years spent in a co-worker relation for each co-worker pair, controlling for the fact that the logarithm of 1 equals zero. Tie strength directly after the termination of the co-workership is formulated as

$$W_{ij}^{1} = \ln(\mathrm{t_f} - \mathrm{t_l} + 1); \quad (4)$$

where t_f refers to the first and t_l to the last year of co-worker status of employee *i* and *j*.

Intuition suggests that the tie looses most from strength in the first years after the termination of co-workership then the slope of decay diminishes in later years. The simplest time decay effect on tie strength is introduced in this paper. This provides results in a dynamically changing tie strength for every co-worker pairs, and is formulated as

$$W_{ij}^{t} = \frac{W_{ij}^{1}}{t}; \quad (5)$$

where t refers to the length of time period after the termination of co-workership.

The above steps result in a weighted individual-level co-worker network for every year that we can aggregate on firm, industry or regional level by simply counting the links and calculating the sum of tie weights.

4. Data and network creation

We use matched employer-employee data obtained from official registers from Statistics Sweden that –among a wide variety of data– contains age, gender, and detailed education code of individual employees and enables us to identify employee-employee co-occurrence at plants for the 1990-2008 period. Data is generated on a yearly basis and if employees change workplace over the year, they are listed repeatedly with different plant codes in the same year. Geo-location of plants is defined by transforming the data from a 100m x 100m grid setting into latitudes and longitudes.

For practical reasons, and in order to keep the size of the sample at the limit the analysis can handle, we exclude those without tertiary education from the data. Including all employees would exponentially increase computation demand without contributing much to the analysis. This is motivated by the fact that skilled workers (bachelors) are assumed to benefit more from learning by seeing and interacting (Glaeser, 2000). We therefore propose that workers without bachelor degrees rely to a greater extent on tacit or embodied knowledge and therefore might learn less from an individual level social network with colleagues at other plants. If an employee who has already been in the data obtains graduation at a later point in time, she will be included in our sample afterwards. As a result, the data contains 366.336 individuals in 1990 and 785.578 individuals in 2008 and those plants are excluded where none of the employees had BA degree or above (Table 1).

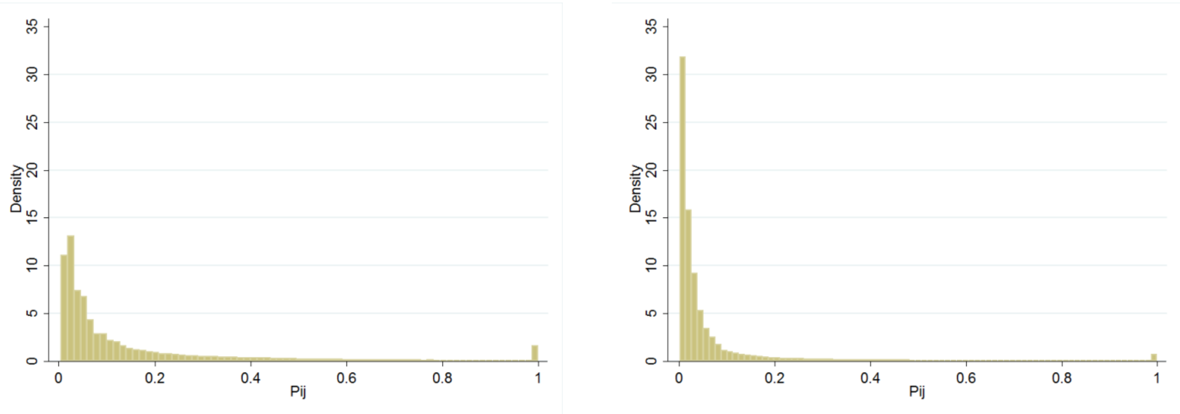
Table 1: Number of employees, plants, and co-occurrence in 1990 and 2008

		1990	2008
Total number of employees	Employees	2,628,306	3,824,182
	Plants	254,445	402,610
Employees with BA degree or above	Employees	366,336	785,578
	Plants	52,872	113,441

We first generated the list of employee pairs as co-occurrence at plants for every year, then calculated the probability of the co-worker relation for each employee pair using Equation 3. We used three characteristics of employees to generate subgroups: Direction of education (6 groups), gender (2 groups) and age (3 groups). For further information of group definitions and descriptive statistics, see Appendix 1.

Figure 1: Distribution of tie probability, 1990 and 2008

(A) (B)



Note: distribution for 1990 in the left and distribution for 2008 in the right.

Figure 1 illustrates that distribution of P_{ij} is left skewed towards zero and decreases monotonously in both 1990 and 2008. However, one can observe that the distribution is more left skewed in 2008 than in 1990 because plants are larger in 2008, which produces lower probabilities. Density of P_{ij} is relatively high at 1 because we set the upper limit there. Nevertheless, the probability that the tie is established is very low for the vast majority of employee co-occurrence.

Employee co-occurrence is exponentially higher in large plants than in small plants and our aim is to find a reasonable number of ties per person, which can be handled by the analysis. There is no clear suggestion in the literature in this regard. Management papers report on task-oriented ego-networks based on survey data and the number of personal ties in these networks are below ten on average (Brass 1985, McPherson et al 1992, Lincoln and Miller 1979, Morrison 2002). Recent papers in labour economics tend to construct much larger co-worker networks assuming that everyone knows each other in a firm not larger than 500 employees (Hensvik and Nordström Skans 2013), or 3000 employees (Saygin et al. 2014). Glitz (2013) only looked at firms with employees between 5 and 50.

Table 2: Tie and degree distribution and isolates at P_{min} thresholds, 1990

Sizecat	Number of Employees	Plants	Mean plant size	Ties above P>0	Ties above P≥0.1	Ties above P≥0.2	Ties above P≥0.3	Ties above P≥0.4	Avg. Degree, P ≥0	Avg. Degree, P ≥0.1	Avg. Degree, P ≥0.2	Avg. Degree, P ≥0.3	Avg. Degree, P ≥0.4	Isolates, P≥0.2	Isolates, P≥0.3	Isolates, P≥0.4
2-9	71,794	19,033	4.88	139,418	139,418	139,418	132,473	128,624	3.88	3.88	3.88	3.69	3.65	0	16	1,468
10-19	46,249	3,420	14.10	302,931	302,931	286,457	280,517	258,711	13.10	13.10	12.38	12.13	11.19	0	2	12
20-49	78,175	2,531	33.24	1,260,292	1,207,405	1,106,538	923,647	722,915	32.24	30.88	28.31	23.63	18.52	0	6	127
50-99	63,102	949	69.11	2,148,933	1,952,091	1,458,559	970,228	592,219	68.11	61.87	46.23	31.01	20.81	5	525	6,172
100-249	34,608	245	151.25	2,600,067	1,851,674	90,533	449,126	252,907	150.25	107.01	55.04	33.83	25.09	1,711	8,063	14,456
250-499	16,831	49	355.47	2,983,041	1,101,615	29,688	156,692	101,567	354.47	133.30	60.59	43.97	35.21	7,032	9,704	11,062
500-999	15,414	24	671.37	5,166,533	73,522	244,708	125,972	76,328	670.37	135.25	79.69	59.47	46.30	9,273	11,178	12,117
1000-	13,553	11	1243.98	8,423,092	659,004	175,234	86,453	84,235	1242.98	182.23	92.52	65.67	65.60	9,765	10,920	10,985
Sum	339,726	26,262		23,024,307	7,287,660	3,531,135	3125108	2,217,506						27,786	40,414	56,399

Evidently, co-occurrences are more likely to be real social ties in small plants and are less likely in large plants (see reports on P_{ij} distribution according to plant size categories in 1990 and 2008 in Appendix 2). Since P_{ij} distribution is similar at the first and last years of the sample, we identify the number of ties per person on base of 1990 network characteristics and apply that number consequently for upcoming years.

We reported in Table 2 on how the number of co-occurrence changes according to plant size categories when we exclude employee pairs under certain P_{ij} minimum threshold. The number of co-occurrences falls dramatically in large plant categories but remains quite stable in small firm categories. We calculated average degree in order to see how many ties an employee has according to plant size categories and also the number of isolates that the P_{ij} threshold generates. The average number of ties is stable until large P_{min} values in very small plants as well as the average degree, and number of isolated employees are very low until the P_{min}=0.4 threshold in plants smaller than 50. This is a large P_{min} threshold and suggests that we can use a 50 best friends approach because everyone might know everyone in small plants. We thus simplify our task and look only at the most likely 50 co-workers of every employee in large plants.

Accordingly, we rank employee pairs based on their P_{ij} values. In case employee pairs have the same probability, we rank those with same educational background and smaller age difference higher, respectively. P_{ij} values are calculated and relations are ranked on a yearly basis, which most likely make co-worker ties appear and disappear from the employees' portfolio in large plants from year to year. To handle this problem, we trace all those co-worker ties that were ranked among the top 50 at least in one year over the full period.

Table 3: Average degree of plants in the co-worker network, 1991-2008

Year	Nodes	Avg. Degr. Plants	Avg. Degr. Ind.	Avg. Degr. Weight
1991	31,391	8.15	71.20	49.35
1992	46,445	11.89	89.72	59.46
1993	53,599	14.46	100.37	59.38
1994	63,299	17.87	112.28	62.53
1995	71,513	22.03	126.23	67.95
1996	79,499	26.04	142.92	71.49
1997	87,072	29.96	152.50	75.03
1998	87,950	32.77	150.82	69.30
1999	95,080	36.89	162.19	71.31
2000	107,423	42.71	179.18	76.51
2001	115,948	47.69	191.51	79.47
2002	120,026	51.25	202.81	76.85
2003	127,355	52.86	208.32	73.92
2004	132,791	54.02	209.27	70.46
2005	140,042	55.77	216.89	72.58
2006	148,318	58.27	223.65	74.90
2007	159,529	64.12	243.35	82.13
2008	166,109	67.12	251.09	85.30

As a result of the above selection process, there are 49,630,691 employee pairs that we trace over 19 years creating a balanced panel of pairs. From the total number of 942,983,129 rows in the panel, we exclude those that have not been appeared in the data yet (481,973,234 pairs), those when at least one employee is already above 65 years of age (42,016,069 pairs). Then, we calculated the strength of ties for the remaining 418,993,826 pairs using Equations 4 and 5. Finally, we excluded those pairs, when either one or both individuals are not present in the labour market for unknown reasons (95,689,892 pairs) and those cases when the employees work in the same plant (167,632,360 pairs).

The remaining unbalanced panel of 155,671,574 employee pairs constitute a dynamic co-worker network over the 1990-2008 period we look at in the analysis. This network can be analysed on the individual level, and ties can be aggregated on the plant and industry levels. However, we must keep in mind, that this is a constantly growing network, because the number of employees in the sample increases monotonically, which is not balanced by labour market exits. For example, after aggregating the network on the plant level, we observe that the number of plants in the network increases over the full period (Table 3, Column 2). As a result, both the number of plants an average plant is connected to (Table 3, Column 3) and the number of individual links from an average plant to any other plants (Table 3, Column 4) increases monotonically. However, one finds that the average weighted degree of plants becomes relatively stable after year 1996 if we introduce the time-decay function. Thus, the dynamic co-worker network that our empirical perspective provides contains a large fraction of weak ties.

5. Geographies of the co-worker network

The analysis of the network is divided into a geographical description and comparison to skill-relatedness networks frequently used in evolutionary economic geography. The spatial level of the regional growth model will be selected on the basis the network geography and we shall provide information on how and why strong co-worker ties scatter across space.

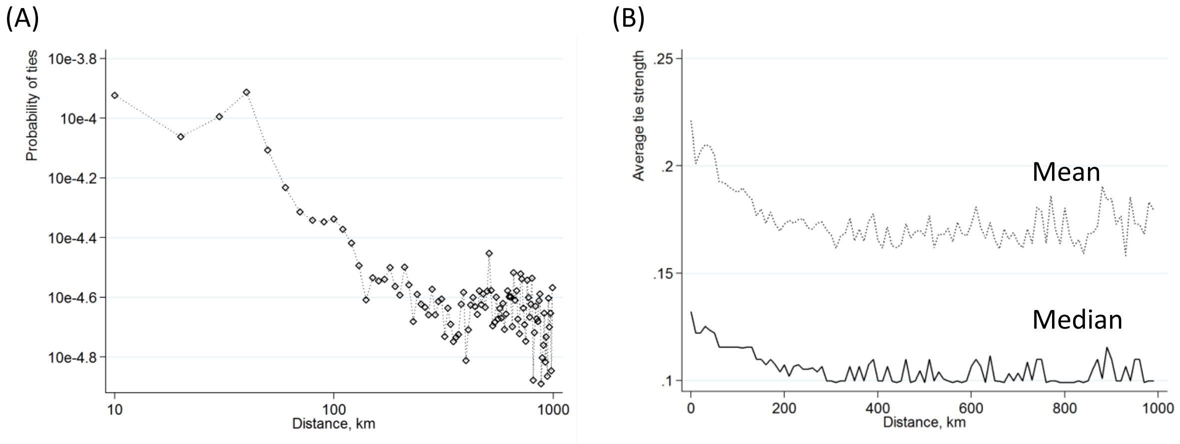
Not surprisingly, the network is spatially concentrated, more than 30% of all individual links were within municipality borders (the smallest administrative division in Sweden) in 2008 and this share is 60% when we look at functional regions (Table 4). The latter regions represent labour market areas and cover the whole territory of Sweden without overlapping each other. When we aggregate the network on the plant level we find a very similar pattern. Tie weights are even more concentrated: 42% of total weights are within municipality and 72% within functional region borders.

Table 4: The number and cumulative weight of ties within regional borders, 2008

	Number of links		Weight of links
	Individual level	Plant level	
Within municipality	7,826,977	1,470,603	3,040,410.8
Within functional region	14,066,872	3,170,695	5,147,408.5
SUM	20,855,160	5,574,879	7,084,738.5

The previous observation gets further support when we look at the probability of having a tie between two arbitrary employees as a function of distance. We define L_d as the number of observed ties between employees separated from each other by distance *d*; and N_d the number of possible ties at distance *d*. Then, we can calculate the probability that individuals have links to others given distance *d* by the formula $P_d=L_d/N_d$. A 10 km resolution was used for binning distance distribution. The probability of a co-worker tie is close to be constant until 40-50 kilometres, after which it falls sharply (Figure 2A). Since the average distance of commuting to another town in Sweden is 45 km, we find that labour market areas and thus functional regions are the proper ground for testing our hypothesis.

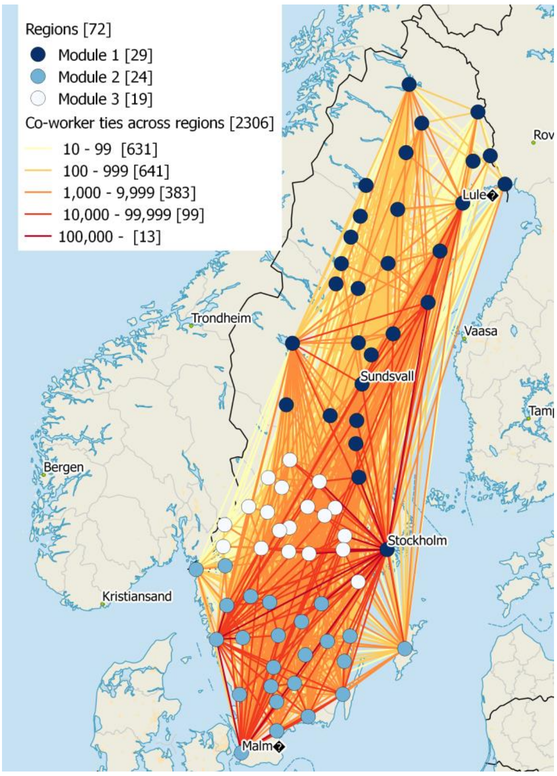
Figure 2: The effect of distance on the network, 2008



Even though the co-worker network concentrates in space, cross-regional labour flow adds a complex interregional dimension to it, in which the majority of strong ties are local but some strong ties can cover large distances. For an illustration consult Figure 2B where we calculated the average weight for each region-region pair from individual tie weights across these given regions and plotted the median and the mean of this average weight distribution using 10 km resolution for binning. One finds that both the median and the mean falls sharply until a certain distance threshold after which they become relatively stable. We also find that the mean exceeds median in all of the bins indicating a left-skewed distribution with relatively few outlier strong ties. In other words, majority of interregional co-worker ties are weak but there are also many exceptionally strong ties across regions, which is most probably due to recent labour flows.

The spatial-base of the co-worker network across functional regions is very plausible when the strength of tie between two regions is the number of individual co-worker links (Figure 3). Not surprisingly, Stockholm (the capital city region) is the centre of the interregional co-worker network meaning that the city has many individual-level ties to other regions. One can also find that Northern regions are very loosely connected with the exception of coastal towns like Umea or Lulea and the network is denser in the South than in the North. The Louvain community detection algorithm finds three modules that clearly represents a spatial divide in the co-worker network meaning that an employee in the South is more likely to know another employee in the South than in the Centre or in the North. Interestingly, Stockholm belongs to the Northern part in the network, which is probably due to a higher share of mobility from the North to the capital compared to mobility from Southern regions to the capital (Eriksson and Lindgren, 2009).

Figure 3: Number of co-worker ties across Swedish functional regions, 2008



Note: Same coloured nodes belong to same network module; modularity index is 0.074. Edges below the threshold number of links=10 are filtered out.

6. The effect of network density on productivity growth

The hypothesis is tested in this section by estimating the following regional growth equation:

$$y_{i,t+3}/y_{i,t} = \beta' X_{i,t-1} + \eta_{i,t} + \varepsilon_{i,t}; \qquad (6)$$

where *y* denotes productivity growth, *t* denotes one-year intervals from 1993 to 2005, *i* denotes the region, *X* stands for the set of explanatory variables, η denotes and unobserved effect of region-specific of time-invariant determinant of growth and ε is the error term.

We construct a panel dataset that contains all variables at the regional level. Regional productivity is measured by value added per capita in region *i* at time *t* and regional productivity growth (*ProdGro*) at time *t* means regional productivity at time *t* compared to regional productivity in region *i* at time *t+3*. Two density indicators are used as explanatory variables. Population density (*PopDens*) is calculated analogously to previous studies (e.g., Ciccone and Hall 1996, Glaeser 1999) as the number of total population over the size of the region measured in square-kilometres. Network density (*NetDens*) measures the share of existing links among all possible links in the co-worker network and can be formulated as

$$D = \frac{2 \times L}{N \times (N-1)}; \qquad (7)$$

where L is the number of existing links and N is the number of nodes.

Human capital (*HC*) is used as a control variable and is calculated by the share of employees among the total number of employees in the region. The models also control for the level of regional productivity (*RegProd*) at time *t* to control for potential catch-up effects. *PopDens, NetDens* and *HC* variables are log-transformed so that the distribution of all variables is close to the normal distribution, which is desired for a linear regression setting. Appendix 3 contains descriptive statistics of each variable.

Table 4: Descriptive statistics and correlation, pooled cross-section 1993-2005

Variable	Definition	Correlations			
ProdGro	Rate of value added per capita at *t* compared to *t+3*	1			
PopDens	Natural logarithm of population density in the region	-0.147	1		
NetDens	Natural logarithm of co-worker network density in the region	0.127	-0.808	1	
HC	Natural logarithm of the share of highly educated employees	-0.081	0.368	-0.692	1
RegProd	Value added per capita	-0.487	0.277	-0.318	0.222

Note: All co-efficients are significant at the 5% level.

Due to data-management considerations, we have to limit the time-span of the regression analysis in two respects. First, we set 1993 as the first year in the panel, because the network construction exercise started in year 1990 and there might be inefficient number of links in the first years. Second, the three years lag in the dependent variable let us to investigate 2005 as the last observed year.

Table 4 contains pooled cross-sectional pairwise Pearson correlation values of the variables. *NetDens* is weakly but positively correlated with *ProdGro* and all other variables are negatively correlated with *ProdGro*, which suggests that population density does not promises a positive impact on regional productivity growth from a cross-sectional perspective; whereas the higher density of the co-worker network the faster growth of regional productivity. *NetDens* is strongly and negatively correlated with *PopDens* and *HC* variables; these coefficients warn us about potential multicollinearity in the multiple regression models. Most importantly, the correlation suggests, contrary to previous studies (e.g. Storper and Venables, 2004; Glaeser, 1999) that population density does not reflect the density of social networks in a region. This implies that network density and population density capture somewhat different aspects of density that are not to be interchangeable.

We estimate the regional production growth model with a linear panel regression with year fixed-effects to control for unobserved time-specific heterogeneity. Two different model specifications are applied. A regional between effect (BE) model which emphasises the cross-sectional variation in the data, and a regional fixed effect (FE) model that emphasizes the within regional variation over time. Since the correlation analysis in Table 4 indicated some multicollinearity between the density variables and also between *NetDens* and *HC* variables we first introduce the explanatory variables into the models separately (Model 1-2 and Model 3-4), then introduce their interaction (Model 5-6) and at last also introduce the interaction between *NetDens* and *HC* (Model 7-8). Results are summarized in Table 5.

Table 5: Linear panel regression - regional productivity growth, 1993-2005

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
PopDens	-0.054	***	1.537	**					-0.119	***	-0.146		-0.134	***	0.283	
	(0.012)		(0.605)						(0.024)		(0.867)		(0.033)		(0.978)	
NetDens					0.065	***	0.552	***	0.059	*	1.038	***	0.184		1.488	***
					(0.019)		(0.141)		(0.030)		(0.215)		(0.123)		(0.377)	
HC	0.101		0.480	**	0.253	**	0.549	**					0.215		0.925	**
	(0.069)		(0.233)		(0.107)		(0.229)						(0.205)		(0.389)	
RegPop	-0.000	***	-0.002	***	-0.000	*	-0.002	***	-0.000	*	-0.002	***	-0.000		-0.002	***
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
PopDens× NetDens									-0.022	***	-0.268	***	-0.028	***	-0.309	***
									(0.007)		(0.091)		(0.010)		(0.095)	
NetDens× HC													0.048		0.158	
													(0.049)		(0.108)	
Constant	1.707	***	-0.270		2.138	***	5.432	***	1.604	***	3.593	*	2.119	***	4.225	**
	(0.169)		(1.578)		(0.272)		(0.717)		(0.077)		(1.862)		(0.499)		(1.944)	
Year FE	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Region	Between		Fixed		Between		Fixed		Between		Fixed		Between		Fixed	
R-sq	0.357		0.365		0.286		0.372		0.438		0.380		0.448		0.385	
Adj. R-sq	0.329		0.301		0.255		0.308		0.405		0.317		0.397		0.320	
N	936		936		936		936		936		936		936		936	

Notes: standard errors in parentheses; *, **, *** sign the level of significance at the 0.1, 0.05 and 0.001 levels, respectively.

The most important finding is that co-worker network density enhances regional productivity growth. The co-efficient of *NetDens* is always positive and loses significance only in Model 7, which reflects that *NetDens* has no significant impact on productivity growth when an interaction with *HC* is introduced to the BE regression. In general, the coefficients of the between effect models are weaker than the coefficients of the fixed effects models, which suggests that higher network density across regions imply higher productivity growth rates across regions but that increasing network density over time in the region enhances productivity growth much more.

Population density does not have a clear effect on productivity growth in our model. The between effect settings suggest that growth rates are lower in areas with high population density. Although Model 2 suggests that regions with increasing population density over time exhibit faster growth rates; the positive effect of *PopDens* disappears when *NetDens* is introduced to the regression.

The interaction term between *PopDens* and *NetDens* variables is negative and significant, which implies that population density weakens the positive effect of network density on regional productivity growth.

In sum, we find that co-worker network density enhances and speeds up regional productivity growth, which accords with Hypothesis 1. This result is novel contribution to the regional growth literature, since we find new systematic evidence that not population density per se, which was highlighted by previous studies (Ciccone and Hall 1996, Glaeser 1999) but the density of the social network is decisive for growth. Learning between co-located individuals and consequent productivity gains are more likely if these individuals have been in a co-worker relation before.

7. Co-worker network versus skill-relatedness network

Social networks might differ across regions in terms of cross-industrial patterns. Co-worker ties presumably concentrate within industry borders stronger in specialized regions and employees are more likely to know employees in other industries in urban areas due to the diversity of such regions. Since the important notion has a direct effect on the outcomes of local learning we analyze the co-worker network in this regard. We argue that the contribution will offer a new approach in understanding spatial knowledge externalities.

Inter-industry labour flows are considered a major source of cross-industrial learning in economic geography (Boschma et al. 2009) and the co-worker network is claimed to depend endogenously on labour flows (Collet and Hedström 2012). We compare the co-worker approach with skill-relatedness that is a recent concept coined in evolutionary economic geography and exploits inter-industry labour flows (Neffke and Henning 2013, Neffke et al. 2013). Two industries are defined skill-related when the sum of observed labour flows between them is larger than the expected value based on industry characteristics because similar employee skills are needed in those two industries. Recent studies suggest that regions endowed with many skill-related labour flows also have much higher productivity growth rates (Boschma et al, 2014), which motivates the comparison between the two types of networks. We introduce social relatedness in the chapter and two industries are considered socially related when the number of co-worker links between them is higher than expected.

The co-worker approach differs from skill-relatedness approach in three basic aspects. First, differences lie in the effect of labour mobility on links between plants. For example, if employees move from industry A to industry B and also to industry C there are links A to B and A to C in the labour mobility network but an additional link will be created between B and C in the co-worker network. Second, labour mobility is considered to be a single transaction of personally embodied knowledge in the skill-relatedness approach, while the co-worker approach can also consider the strength of ties. Third, and most importantly, our approach enables us to directly see the relatedness between local plants and industries and –in contrast to the skill-relatedness approach– we do not project national level relatedness networks to the region.

Skill-relatedness and social-relatedness industry networks are compared on a national level. In a further step, we provide new insights regarding the importance of skill-relatedness, social-relatedness and unrelated co-worker networks as a function of region size.

7.1 National level comparison

We compute skill-relatedness of industries with a method introduced by Neffke et al. (2013) and compare the observed labour flow to the expected labour flow between industry p and q over the 2000-2008 period by

$$R_{pq} = \frac{F_{pq}F_{..}}{F_{p.}F_{.q}}\ ; \qquad (8)$$

where F_{pq} is the observed number of flows between industry p and q, $F_{..}$ is the total number of flows, $F_{p.}$ is the number of workers leaving industry p and $F_{.q}$ is the number of workers joining industry q. Then, we transform R_{pq} onto the interval [-1; 1) as follows:

$$\bar{R}_{pq} = \frac{R_{pq}-1}{R_{pq}+1}. \qquad (9)$$

Two industries are skill-related when $\bar{R}_{pq}$>0 because the labour flow between p and q is higher than expected from the full labour flow matrix. Borrowing the above logic, we compute the same indicator for the co-worker network between industries. Let L_{pq} be the number of co-worker links between industry p and q, L_p and L_q the total number of links employees have in industry p and industry q, and L the total number of links. Then, relatedness based on co-worker links is given by

$$R^l_{pq} = \frac{L_{pq}L}{L_pL_q},\ and \qquad (10)$$

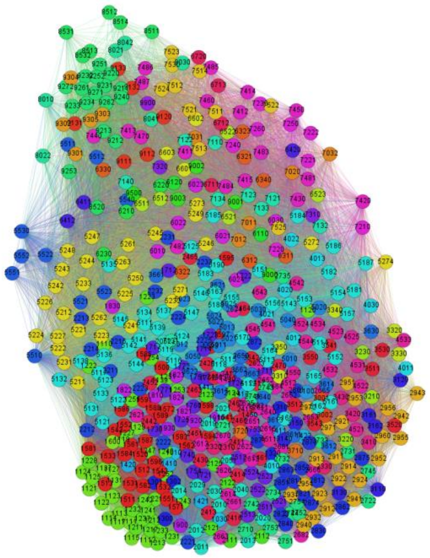
$$\bar{R}^l_{pq} = \frac{R^l_{pq}-1}{R^l_{pq}+1}. \qquad (11)$$

In a similar fashion, let W_{pq} be the accumulated strength of co-worker links between industry p and q, W_p and W_q the accumulated strength of links employees have in industry p and industry q, and W the sum of strength of all links. Then, relatedness based on the strength of co-worker links is given by

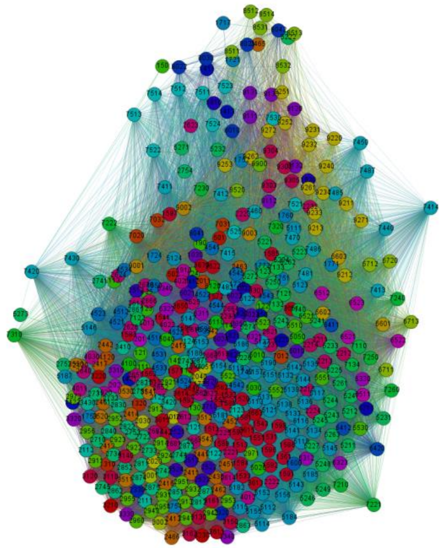
$$R^w_{pq} = \frac{W_{pq}W}{W_pW_q}\ , and \qquad (12)$$

$$\bar{R}^w_{pq} = \frac{R^w_{pq}-1}{R^w_{pq}+1}. \qquad (13)$$

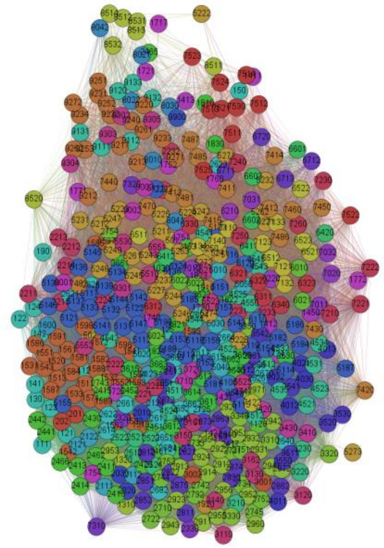
Figure 4: Skill relatedness network (2000-2008), un-weighted co-worker network (2008), and weighted co-worker network (2008) at the national level



(A)

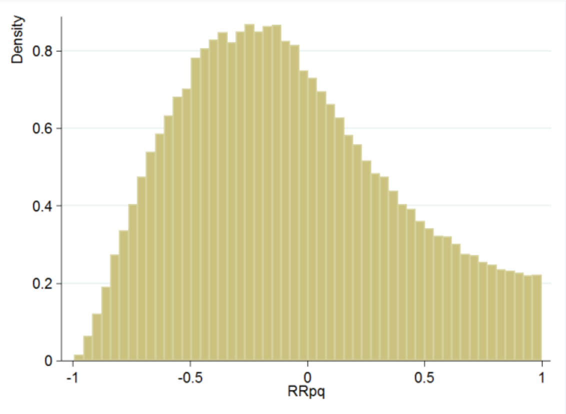


(B)

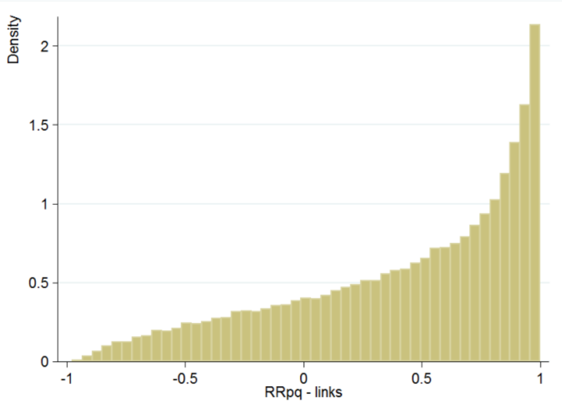


(C)

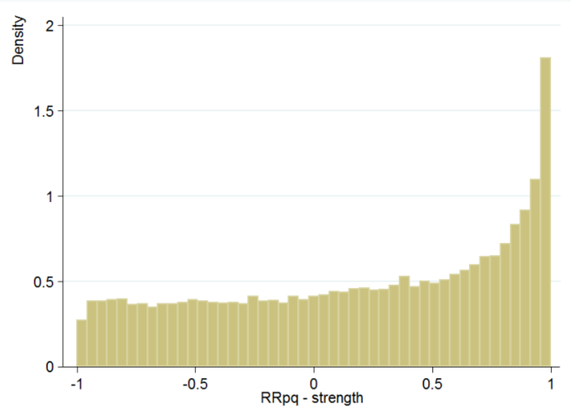
(D)



(E)



(F)



Note: Networks drawn from edges above the zero-threshold of the respected $\bar{R}$ value. Nodes in the networks represent NACE 4-digit industries. Same coloured nodes belong to the same NACE 2-digit sectors. Force Atlas algorithm was used in Gephi; parameters were optimized for visualization.

We find similar patterns of the industry network when using $\bar{R}_{pq}$, $\bar{R}^{l}_{pq}$, and $\bar{R}^{w}_{pq}$ values as edge weights in visualization (Figure 4ABC). Industries that one expects to be similar cluster together in all three approaches. However, the distribution of the above indicators shows significant differences (Figure 4DEF). While the distribution of skill-relatedness is close to be normal, relatedness based on co-worker links as well as on strength of co-worker links is a monotonically increasing.

This observation suggests that the co-worker network perspective is different from the labour flow perspective and offers a more complex picture because it captures those personal linkages between industries that a labour flow matrix overlooks. However, the notion that the distribution smoothes when strength of ties is in consideration suggests that that the co-worker network is more similar to the skill-relatedness network if we look at the accumulation of tie strength instead of aggregating the links only, which is due to the large effect of recent labour flows on edge weights between industries.

Table 6: Descriptive statistics of skill-relatedness and co-worker networks

		Edges RR>0	Density RR>0	Avg. Path length RR>0	Diameter RR>0	Modularity RR>0	Edge-weight correlation	
1	Skill-relatedness	51,778	0.319	1.682	3	0.22		
2	Co-worker – links (L)	43,615	0.378	1.633	4	0.118	0.597	
3	Co-worker – strength (W)	34,659	0.3	1.715	4	0.148	0.617	0.838

Note: Pairwise Pearson correlation values are significant at the 1% level.

The above finding is further supported by correlating basic network indicators across the three approaches at hand. For example, $\bar{R}^{l}_{pq}$ and $\bar{R}^{w}_{pq}$ are strongly correlated, but $\bar{R}^{w}_{pq}$ correlates stronger with $\bar{R}_{pq}$ than $\bar{R}^{l}_{pq}$ (Table 6). Similarly, when looking at industry centralities one finds that the co-worker network based on tie strength is more similar to the skill-relatedness network than the co-worker network based on accumulated links (Table 7).

Table 7: Correlation of industry centralities in skill-relatedness and co-worker networks

Degree centrality	Skill	Link	Closeness centrality	Skill	Link
Link	0.638	1.000	Link	0.629	1.000
Strength	0.700	0.971	Strength	0.678	0.970
Weighted degree centrality	Skill	Link	Betweenness centrality	Skill	Link
Link	0.627	1.000	Link	0.532	1.000
Strength	0.650	0.962	Strength	0.616	0.952

Note: Pairwise Pearson correlation values are significant at the 1% level.

Appendix 4 reveals the ego-network of industry 2442 (Aluminium production) and most important related industries using a spring algorithm. Only seven industries appear in all three networks out of 30 industries and the different layout of these networks indicates different industry spaces as well.

7.2 Co-worker and skill-relatedness networks in regions

The biggest advantage of the co-worker network perspective compared to the skill-relatedness network is that one does not need to construct the network on the national level and then project it on the regions. Consequently, we don't have to assume that the same relatedness matrix applies for metropolitan regions as well as rural areas because the co-worker network offers us a micro approach and reveals the actual social relations of plants or industries in every spatial unit. In the following, we show that this feature indeed provides us with novel information regarding spatial social network.

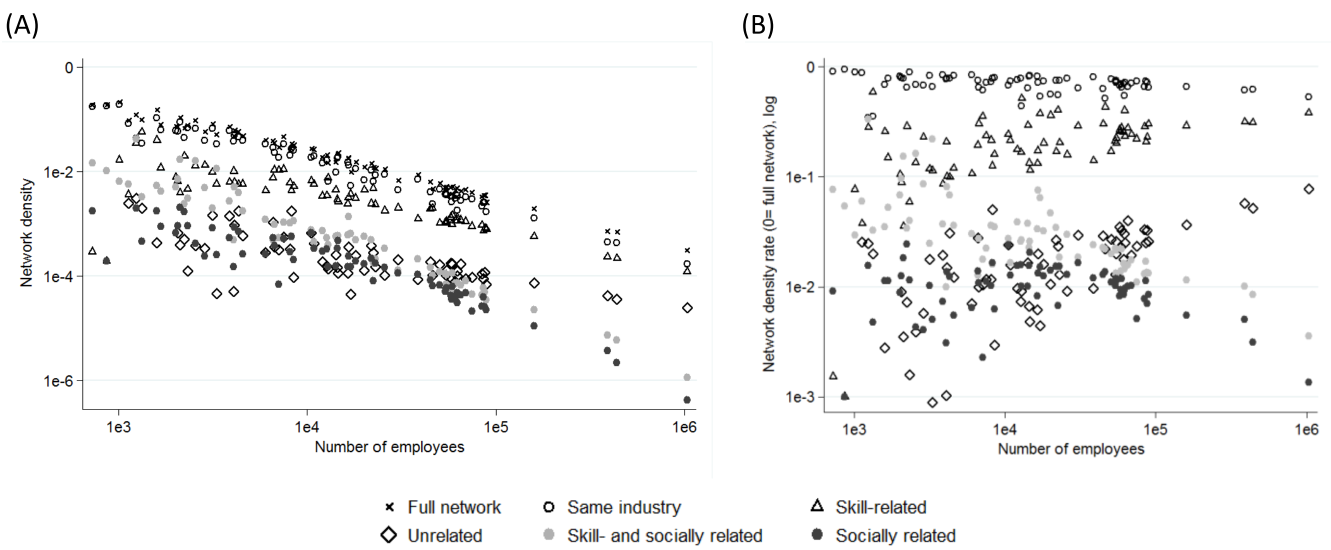
Industry-industry links in the co-worker network are characterized into five non-overlapping groups. *Same industry* category means that individuals belong to the same NACE4 industries but work in different plants. *Skill-related* category includes co-worker ties between industries among which labour mobility is higher than expected (Eq. 9). Co-worker ties are stronger than expected among *socially-related* industries (Eq. 13). One can also find a large share of industry pairs that are *skill- and socially-related* at the same time. Finally, there are industry pairs that are *unrelated*.

After labelling the inter-industry co-worker network, we can look at densities within industries, across skill-related industries, across socially-related industries, across skill- and socially related industries, and non-related industries. Thus, we can decompose the network density indicator we used in the previous section (Eq. 7) into density within industries and density across industries by relatedness types by a simple decomposition algorithm:

$$D = \sum_p \frac{2\times L_{pp}}{N_p\times(N_p-1)} \times \frac{N_p\times(N_p-1)}{N\times(N-1)} + \sum_{pq}^{r} \frac{2\times L_{pq}}{N_p\times N_q} \times \frac{N_p\times N_q}{N\times(N-1)} \times \delta_{pq}^{r}\ ; \quad (14)$$

where L_{pp} is the existing number of links within industry p, L_{pq} is the existing number of links between industries p and q; N_p and N_q are the number of employees in industries p and q; δ_{pq}^{r} equals 1 if industries p and q are related according to relatedness category r and 0 otherwise.

Figure 5: Decomposed network density and region size, 2008



Note: (A) depicts log-transformed density indicators. (B) depicts the log-transformed value of the rate of density indicators compared to the full network density.

Decomposed density indicators are depicted against the number of employees in the regions in Figure 5 on a log-log scale. In general, density indicators are smaller in bigger regions: the more employees the smaller rate of observed links among possible links. Interestingly, the same industry network is the densest in almost every region. Skill-relatedness is found to decrease on a smoother slope as the function of region size; and looking at the density rate we find that the skill-related network is relatively denser in large regions than in small regions. The overlap of skill-relatedness and social relatedness is prevalent in small regions. A very important finding suggests that the density of socially related network alone is relatively low in small regions; it becomes relatively higher in middle-size regions and becomes relatively low again in large regions. The importance of co-worker networks across unrelated industries is low in small industries and it becomes high in large regions.

Table 8: Share and strength of co-worker ties across industries by functional region types, 2000

Regions		Same industry	%	Skill-related	%	Socially-related	%	Skill- and socially related	%	Unrelated	%	SUM
1	W	7,220,826	81.76	1,342,861	15.20	13,739.25	0.16	40,993.98	0.46	213,488.8	2.42	8,831,909.03
	L	6,657,358	63.65	3,141,419	30.03	28,442	0.27	75,565	0.72	556,533	5.32	10,459,317
	Avg	1.08		0.42		0.48		0.54		0.38		0.84
2	W	3,470,098	88.91	359,694	9.22	16,011.97	0.41	32,311.17	0.83	25,840.02	0.66	3,902,955.16
	L	2,793,496	75.45	754,195	20.37	28,825	0.78	58,142	1.57	67,800	1.83	3,702,458
	Avg	1.24		0.48		0.55		0.55		0.38		1.05
3	W	592,723.4	87.41	69,646.84	10.27	3,136.66	0.46	8,170.92	1.20	4,420.15	0.65	678,097.97
	L	463,194	75.17	120,987	19.63	5,868	0.95	15,934	2.59	10,208	1.66	616,191
	Avg	1.28		0.57		0.53		0.51		0.43		1.1
4	W	87,405.9	88.01	7,802.27	7.86	664.91	0.67	3,207.96	3.23	234.54	0.24	99,315.58
	L	67,776	79.88	10,888	12.83	1,014	1.20	4,571	5.39	601	0.71	84,850
	Avg	1.28		0.71		0.65		0.7		0.39		1.17
5	W	76,386.93	92.37	1,487.78	1.80	4,575.56	5.53	124.20	0.15	124.19	0.15	82,698.66
	L	56,508	83.47	8,071	11.92	299	0.44	2,383	3.52	435	0.64	67,696
	Avg	1.35		0.18		15.3		0.05		0.28		1.22

Note: W denotes the accumulated strength of ties; L denotes the number of ties; Avg. denotes the average strength of ties (Avg=W/L). Functional region codes stand for metropolitan regions (1), bigger regional centres (2), smaller regional centres (3), small regions with private employment (4), small regions with public employment (5).

In a next step, we can also calculate the number and accumulated strength of co-worker ties according to the above network categories (Table 8). We find that co-worker ties in same industries represent the majority of the network in all region types and these ties are the strongest on average. However, the share of ties and their strength is relatively lower in large regions and relatively higher in small regions. Skill-related industries accumulate a large share of relatively weak co-worker links in metropolitan areas but this share is significantly lower in small regions where ties are relatively stronger. Thus, Hypothesis 2 is verified.

Social relatedness of industries is more important in middle-sized and small regions than in metropolitan regions since the share of links and accumulated weight in socially-related as well as in skill- and socially-related industries increases as regions size decreases. There are more co-worker ties across unrelated industries in metropolitan regions than in smaller regions. Thus, Hypothesis 3 is verified.

In sum, we find that the co-worker approach provides methodological improvement primarily in middle-sized and small regions, because skill-relatedness is found to describe the social network less than in large regions. For a visualization of socially related industries in selected regions, consult Appendix 8. An important finding implies that the importance of co-worker networks across unrelated industries increases in large regions, which is due to the increased complexity of social networks in metropolitan areas.

Based on the above findings, we also propose that the co-worker approach opens up new horizons for further research in the field of networks and economic geography. One might envisage that new empirical understanding can be reached in variety of theories including social capital, spatial inter-firm learning, and Jacobsian externalities. Therefore, further research is needed to unveil detailed network characteristics and dynamics in regions.

8. Conclusion

A new empirical perspective for social network analyses in economic geography was introduced in this paper; we constructed the co-worker network in Sweden for the period 1990-2008 and analyzed its spatial dimension. We believe that this approach can offer a wide variety of new answers for questions addressing the role of social networks in regional economic development. The current paper focused on two issues: (1) the effect of network density on productivity growth; (2) the difference in cross-industrial network structure according to regional size.

The paper provides the first systematic evidence that social networks are important for regional dynamics. People might learn more efficiently from those they have been in a co-worker relation with previously rather than from co-location per se. Thus, learning through the co-worker network is expected to enhance the productivity of the region. Indeed, in contrast to previous studies advocating the immense role of density (e.g., (Ciccone and Hall 1996, Glaeser 1999) our empirical analysis indicate that it is not population density per se but the density of the co-worker network that is important for regional productivity growth. This finding verifies our first hypothesis: network density triggers productivity growth.

Another contribution of the paper concerns the illustrated potential of the co-worker approach in mapping inter-industry links. We have demonstrated that links across skill-related and unrelated industries are more likely and are also stronger in large regions compared to in small regions. This is in line with our second hypothesis and indicates that the co-worker perspective is useful for further research addressing the role of social networks in generating Jacobsian externalities. It was also shown in the paper that the co-worker approach works better than skill-relatedness outside the largest regions, especially in middle-sized regions, since social-relatedness has the highest importance in those areas; thus, we verified our third hypothesis.

Since our methodology offers a micro perspective, one can analyse networks aggregated on various levels including individuals, plants, firms, industries or regions. Further research might devote attention to the effects of co-worker network's structure on other aspects of regional dynamics like firm entry, investment flows, entrepreneurship or employment growth introducing sector-specific characteristics into the analysis. For example, employees might learn more in those co-worker networks where the industry-specific knowledge is easier to transfer. Another potential in the co-worker approach is the tie strength and one might be interested how the strength of weak ties –as Granovetter put it– applies to the effect of co-worker networks on innovation performance.

References

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

Ahuja, G. (2000) Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, 45(3): 425-455.

Almeida, P., Kogut, B. (1999) Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, 45: 905-917.

Amin, A. (2000) Industril districts. In Sheppard, E. And Barnes T.J. (eds.) *A Companion to Economic Geography*, pp. 149-168. Oxford: Blackwell Publishing.

Asheim, B. (1996) Industrial districts as learning regions: a condition for prosperity. *European Planning Studies* 4, 4, pp. 379-400.

Bathelt, H., Glückler, J. (2003) Toward a relational economic geography. *Journal of Economic Geography*, 3(2): 117-144.

Bathelt, H., Malmberg, A., Maskell, P. (2004) Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography* 28, 1, pp. 31-56.

Beaman, L., Magruder, J. (2012) Who gets the job referral? Evidence from a social networks experiment. *American Economic Review*, 102(7): 3757-3593.

Blau, P. M., Schwartz, J. E. (1984). *Crosscutting social circles; testing a macrostructural theory of intergroup relations*. Orlando: Academic Press.

Borgatti, S.P., Cross, R. (2003) A Relational view of information seeking and learning in social networks. *Management Science*, 49(4): 432-445.

Borgatti, S.P., Mehra, A., Brass, D.J., Labiance, G. (2009) Network Analysis in the Social Sciences. *Science* 323, pp. 892-895.

Boschma, R.A. (2005) Proximity and innovation: a critical assessment. *Regional Studies*, 39(1): 61-74.

Boschma, R.A., Eriksson, R., Lindgren, U. (2009) How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity. *Journal of Economic Geography*, 9: 169-190.

Boschma, R., Eriksson, R.H. & Lindgren, U. (2014): Labour market externalities and regional growth in Sweden: the importance of labour mobility between skill-related industries. *Regional Studies*, 48 pp. 1669-1690

Boschma, R.A., Frenken, K. (2011) The emerging empirics of evolutionary economic geography. *Journal of Economic Geography*, 11(2): 295-307.

Brass, D.J. (1985) Men's and women's networks: a study of interaction patterns and influence in an organization. *Academy of Management Journal* 28, 2, pp. 327-343.

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

Buhai, S., van der Lei, M. (2006) A social network analysis of occupational segregation. Tinbergen Institute Discussion Paper, No. 06-016/1.

Burt, R. L. (1992) *Structural Holes*. Cambridge, MA: Harvard University Press.

Calvo-Armengol, A., Jackson, M.O. (2004) The Effects of Social Networks on Employment and Inequality. *American Economic Review*, 94(3): 426–54.

Ciccone, A., Hall, R.E. (1996) Productivity and the Density of Economic Activity. *American Economic Review* 86, 1, pp. 54-70.

Coleman, J. S. (1958): "Relational Analysis: The Study of Social Organizations With Survey Meth- ods," *Human Organization* 17: 28-36.

Coleman, J. S. (1990) *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Currarini, S., Jackson, M.O., Pin, P. (2009) An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica* 77(4): 1003-1045

Collet, F., Hedström, P. (2012) Endogenous tie formation mechanisms in a directed network generated by employee mobility. IFAU Working Paper, 2012: 25.

Dahl, M.S., Pedersen, C.O.R. (2003) Knowledge flows through informal contacts in industrial clusters. Myth or reality? *Research Policy*, 33: 1673–1686.

Duranton, G. & Puga, D. (2004) Micro-foundations of urban agglomeration economies. In: Henderson J.V. and Thisse J.V. (Eds) *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2063–2117. North-Holland, Amsterdam.

Eriksson, R., Lindgren, U. (2009) Localized mobility clusters: impacts of labour market externalities on firm performance. *Journal of Economic Geography*, 9: 33-53.

Feldman, M.P. (1999) The new economics of innovation, spillovers and agglomeration: a review of empirical studies. *Economics of Innovation and New Technology* 8, 1-2, pp. 5-25.

Frenken, K., Van Oort, F.G., Verburg, T. (2007) Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5): 685–697.

Glaeser, E.L., Kallal, H. D., Scheinkman, J. D., Shleifer, A. (1992): Growth in Cities. *Journal of Political Economy*, 100: 1126- 1152.

Glaeser, E. L. (1999) Learning in cities. *Journal of Urban Economics* 46, pp. 254-277.

Glaeser, E.L. (2000) The New Economics of Urban and Regional Growth. In: Clark, G.L., Gertler, M.S. and Feldman, M. (eds.) *The Oxford Handbook of Economic Geography*, pp: 83-98. Oxford University Press, Oxford.

Glitz, A. (2013) Co-worker networks in the labour market. CESifo Working Paper Series No. 4250. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2274613

Glückler, J. (2007) Economic geography and the evolution of networks. *Journal of Eonomic Geography* 7, 5, pp. 619-634.

Granovetter, M. (1973) The strength of weak ties. *American Journal of Sociology*, 78(6): 1360-1380.

Granovetter, M. (1995) *Getting a Job: a Study of Contacts and Careers*. Chicago: The University of Chicago Press.

Henderson, J. V. – Kuncoro, A. – Turner, M. (1995): Industrial Development in Cities. *Journal of Political Economy*, 103: 1067-1085.

Hensvik, L., Nordström Skans, O. (2013) Social networks, employee selection and labor market outcomes. IFAU Working Paper, 2013: 15.

Huber, F. (2012): Do clusters really matter for innovation practices in Information Technology? Questioning the significance of technological knowledge spillovers. *Journal of Economic Geography* 12: 107-126

Jackson, M.O. (2008) *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

Lincoln, J.R., Miller, J. (1979) Work and friendship ties in organizations: a comparative analysis of relation networks. *Administrative Science Quarterly* 24, 2, pp. 181-199.

Malmberg, A. (1997) Industrial geography: location and learning. *Progress in Human Geography* 21, 4, pp. 573-582.

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

Marsden, P.V., Campbell, K.E. (1984) Measuring tie strength. *Social Forces* 63, 2, pp. 482-501.

Maskell, P., Malmberg, A. (1999) Localised Learning and Industrial Competitiveness. *Cambridge Journal of Economics* 23, pp. 167-185.

McPherson, M., Popielarz, P.A., Drobnic, S. (1992) Social networks and organizational dynamics. *American Sociological Review* 57, 2, pp. 153-170.

McPherson, M., Smith-Lovin, L. (1987) Homophily in voluntary organizations: status distance and the composition of face-to-face groups. *American Sociological Review* 52, 3, pp. 370-379.

McPherson, M., Smith-Lovin, L., Cook, J.M. (2001) Birds of a feather: homophily in social networks. *Annual Review of Sociology* 27, pp. 415-444.

Morrison, E.W. (2002) Newcomers relationship: the role of social network ties during socialization. *The Academy of Management Journal* 45, 6, pp. 1149-1160.

Neffke, F., Henning, M. (2013) Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3): 297-316.

Neffke, F., Henning, M., Boschma, R.A. (2011) How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography*, 87(3): 237-265.

Neffke, F., Hidalgo, C., Otto, A., Weyh, A. (2013) Inter-industry labor flows. Mimeo.

Saygin, P.O., Weber, A., Weynandt, M.A. (2014) Coworkers, networks, and job search outcomes. IZA Discussion Paper No. 8174.

Sias, P.M., Cahill, D.J. (1998) From co-workers to friends: the development of peer friendship in the workplace. *Western Journal of Communication* 62, 3, pp. 273-299.

Sorensen, O. (2003) Social networks and industrial geography. *Journal of Evolutionary Economics*, 13: 513-527.

Sorensen, O., Rivkin, J. W., Fleming, L. (2006) Complexity, networks and knowledge flow. *Research Policy* 35, 7, pp. 994-1017.

Storper, M., Venables, A.J. (2004) Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography* 4, 4, pp. 351-370.

Ter Wal, A., Boschma, R.A. (2009) Applying social network analysis in economic geography: framing some key analytic issues. *Annals of Regional Science*, 43: 739-756.

Timmermans, B., Boschma, R.A. (2013) The effect of intra- and inter-regional labour mobility on plant performance in Denmark: the significance of related labour inflows. *Journal of Economic Geography*, 14(2): 289-311.

Walker, G., Kogut, B., Shan, W. (1997) Social Capital, Structural Holes and the Formation of an Industry Network. *Organization Science*, 8(2): 109-125.

Wasserman, S., Faust, K. (1994) *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Appendix 1a: Categories of employee education by direction of studies

			1990		2008		1990	2008
		code	N	%	N	%	%	%
1	Pedagogy and teaching	14	107,853	29,441	168,497	21,44879	29,44	21,45
2	Arts and media	21	5.100	1.392165	12.018	1.529829	6.91	5.84
	Journalism and media	32	3.491	0.95295	11.053	1.40699		
	Humanities	22	16.725	4.565481	22.825	2.905504		
3	Social sciences	31	27.273	7.444805	47.950	6.103786	22.43	21.40
	Business. trade and administration	34	40.262	10.99046	92.489	11.77337		
	Law	38	14.640	3.996331	27.662	3.521229		
4	Biology and environment	42	1.821	0.497085	9.571	1.218339	4.54	6.08
	Physics and chemistry	44	3.191	0.871058	10.265	1.306681		
	Mathematics	46	9.381	2.560764	10.637	1.354035		
	Data	48	2.256	0.615828	17.288	2.200673		
5	Engineering	52	36.910	10.07545	105.734	13.45939	14.68	18.09
	Manufacturing	54	1.476	0.402909	4.072	0.518344		
	Construction	58	10.915	2.979505	23.481	2.989009		
	Agriculture and forestry	62	2.835	0.77388	5.767	0.734109		
	Environmental protection	85	467	0.127479	1.828	0.232695		
	Transport services	84	1.175	0.320744	1.265	0.161028		
6	Animal care	64	807	0.22029	1.865	0.237405	21.00	24.37
	Health care	72	58.451	15.95557	151.420	19.27498		
	Social work	76	17.647	4.817162	36.679	4.669046		
	Personal services	81	42	0.011465	1.472	0.187378		
0	Security and military	86	52	0.014195	3.634	0.462589	0.99	2.77
	Unknown	99	3.566	0.973423	18.106	2.3048		
	SUM		366.336	100	785.578	100	100.00	100.00

Note: Employees with educational background code 0 are excluded from the analysis.

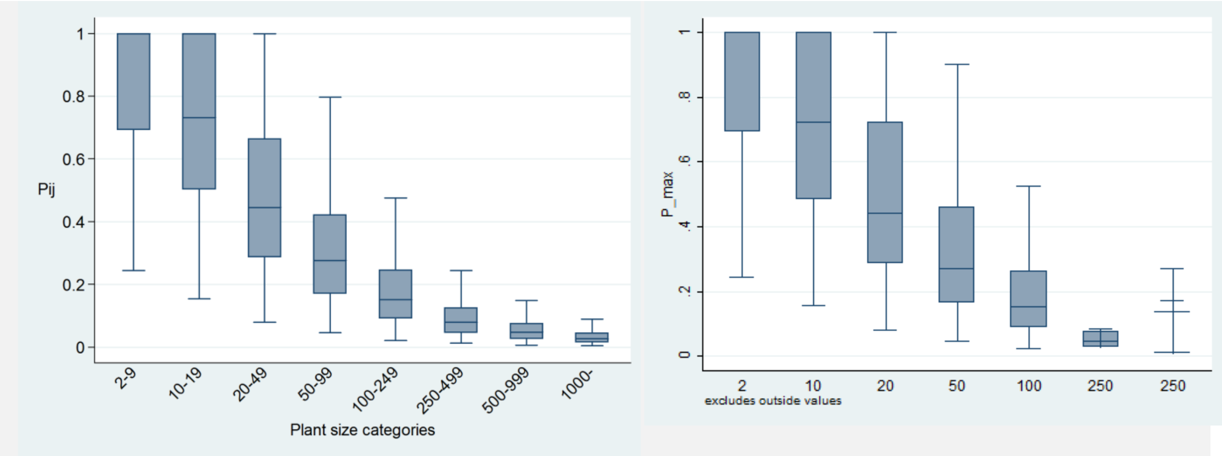
Appendix 1b: Number of employees by gender categories

Gender	1990	2008
0	182874	451303
1	183462	334275
SUM	366336	785578

Appendix 1c: Number of employees by age categories

Age	1990	2008
-34	79437	217813
35-49	201334	317635
50-	85565	250130
SUM	366336	785578

Appendix 2: Tie probability distribution and firm size categories. 1990 and 2008

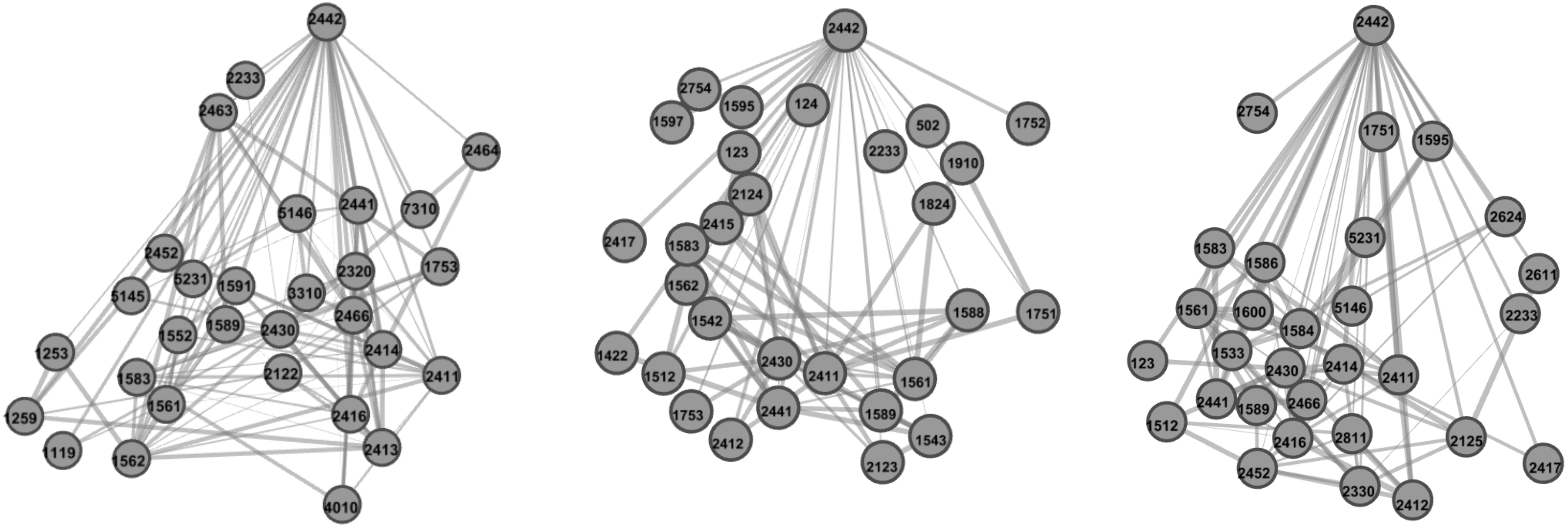


Note: Distributions for 1990 in the left and for 2008 in the right. We excluded those large number of outlier observations that are below or above the whiskers.

Appendix 3: Descriptive statistics of the variables in the regional growth model, 1993-2005

Variable		Mean	Std. Dev.	Min	Max	Obs	
ProdGro	overall	1.233	0.539	0.372	7.368	N =	936
	between		0.154	1.040846	2.033	n =	72
	within		0.517	0.085	6.568	T =	13
PopDens	overall	2.299	1.474	-1.399	4.950	N =	936
	between		1.483	-1.316	4.904	n =	72
	within		0.034	2.188	2.408	T =	13
NetDens	overall	-3.662	1.445	-7.946	-0.721	N =	936
	between		1.446	-7.861	-1.158	n =	72
	within		0.151	-4.24	-3.18	T =	13
RegProd	overall	326.843	121.078	30.641	1288.007	N =	936
	between		84.886	114.656	634.421	n =	72
	within		86.872	45.549	1218.82	T =	13
HC	overall	-2.133	0.279	-2.965	-1.233	N =	936
	between		0.243	-2.643	-1.439	n =	72
	within		0.14	-2.455	-1.719	T =	13

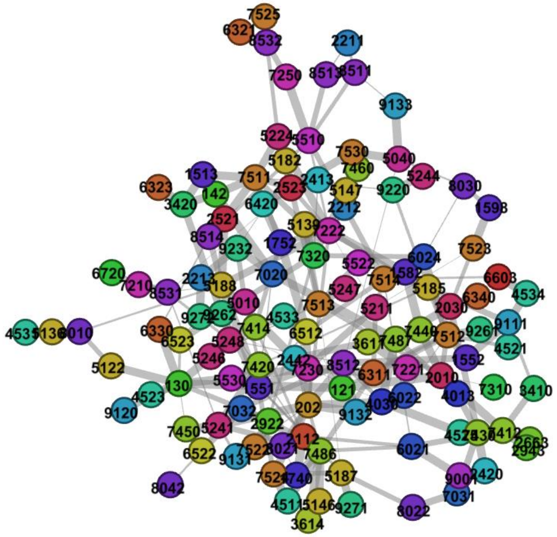
Appendix 4: Ego networks of industry 2442 according to skill-relatedness, co-worker links, and co-worker link strength settings

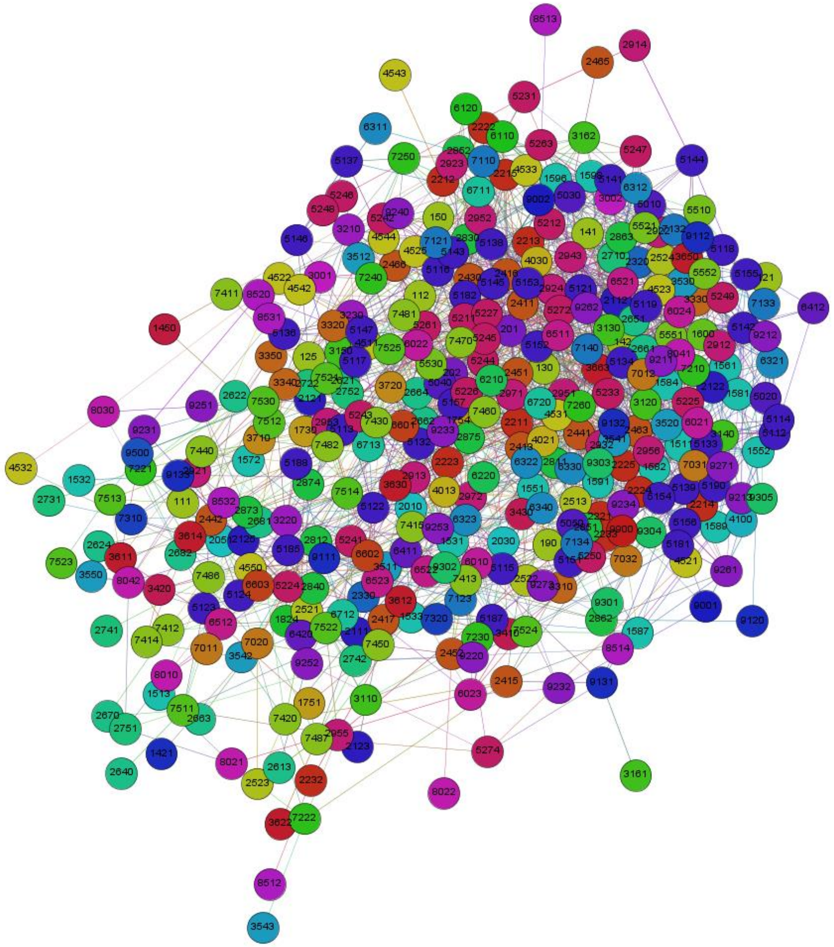


Note: Networks drawn with Fruchterman-Reingold algorithm in Gephi after Node 2442 was settled above its ego-network. Nodes have been systematically sorted out by setting the minimum threshold of edges weight until the top 30 neighbours remained in the sample. The minimum threshold (A) 0.852, (B) 0.9465 (C) 0.921.

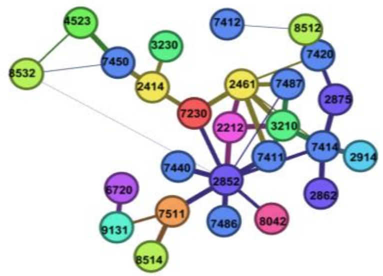
Figure 5: The socially related industries in Stockholm, Umea, and Karlskoga, 2000

Umea, D = 3.26





Stockholm, D = 4.04



Karlskoga, D = 6.26