INSTITUTE OF ECONOMICS
Centre for Economic and Regional Studies

Közgazdaság- és Regionális Tudományi Kutatóközpont
KÖZGAZDASÁG-TUDOMÁNYI INTÉZET

CERS-IE WORKING PAPERS | KRTK-KTI MŰHELYTANULMÁNYOK

# Co-worker networks and firm performance

BALÁZS LENGYEL – GUILHERME KENJI CHIHAYA –

LÁSZLÓ LŐRINCZ – RIKARD ERIKSSON

**CERS-IE WP – 2021/18**

May 2021

https://www.kti.krtk.hu/wp-content/uploads/2021/05/CERSIEWP202118.pdf

ABSTRACT

Firms and employees can benefit from information diffusion through social connections at other firms. Therefore, co-worker networks observed in collaborative projects or assumed from job co-occurrence have been analyzed in a wide literature ranging from management to economics, and economic geography. Yet, beyond case-studies, the actual information flows are seldom identifiable in these networks and previous focus on firm- or employee benefits was mainly limited to dyadic relations across firms. To address this gap, we simulate co-worker networks within firms from large-scale administrative data, for which we use parameters fitted to information networks that we collected with a survey and from social media profiles. Then, following all individuals through job moves over their career, we establish the dynamic co-worker network across firms of the entire ICT industry in Sweden. Fixed-effect regression models suggest that growth of average income is significantly higher in those firms that have diverse connections but are central to the network as well. We find that large firms benefit more from triadic closure in the co-worker network, stressing the role cohesive relations in sharing complex knowledge. Our results highlight that firm growth is embedded into the eco-system of co-worker networks that facilitate information flows across firms.

JEL codes: D85, L25, J62, O47

Keywords: co-worker networks, survey, social media, link prediction, administrative data, network simulation, firm growth

Balázs Lengyel
Agglomeration and Social Networks Lendület Research Group, ELKH Centre for Economic- and Regional Studies, 1097 Budapest, Hungary
and
NETI Lab, Budapest Corvinus University, 1093 Budapest, Hungary
e-mail: lengyel.balazs@krtk.hu

Guilherme Kenji Chihaya
Department of Geography, Umea University
and
Institute for Analytical Sociology, Linköping University
e-mail: guilherme.chihaya@umu.se

László Lőrincz
NETI Lab, Budapest Corvinus University, 1093 Budapest, Hungary
and
Agglomeration and Social Networks Lendület Research Group, ELKH Centre for Economic- and Regional Studies, 1097 Budapest, Hungary
e-mail: laszlo.lorincz@uni-corvinus.hu

Rikard Eriksson
Department of Geography, Umea University
and
Center for Regional Science, Umea University
e-mail: rikard.eriksson@umu.se

# Munkatársi kapcsolathálózatok vállalati teljesítmény

LENGYEL BALÁZS– GUILHERME KENJI CHIHAYA –

LŐRINCZ LÁSZLÓ– RIKARD ERIKSSON

ÖSSZEFOGLALÓ

A cégek és alkalmazottak profitálhatnak a más cégekhez fűződő társas kapcsolataikból. Ezért a projektekben való együttműködésekben megfigyelt vagy a közös munkahely alapján feltételezett munkatársi kapcsolathálózatokat széles körűen elemezték a menedzsment, közgazdaságtan és gazdaságföldrajz irodalmakban. Az esettanulmányokon túl azonban a tényleges információáramlás ritkán azonosítható ezekben a hálózatokban, és a cég- vagy alkalmazottak előnyeit vizsgáló való korábbi kutatások főként a vállalatok közötti diadikus kapcsolatokra korlátozódtak. Ebben a cikkben nagy adminisztratív adatok alapján szimuláljuk a cégeken belüli munkatársi kapcsolat-hálózatokat, amelyekhez kérdőívvel gyűjtött információs hálózatok becslésének paramétereit használjuk. Ezután követjük a munkavállalókat a karrierjük során, amivel egy dinamikus munkatársi kapcsolathálózatot hozunk létre a teljes svédországi IKT iparág összes vállalata között. A fix-hatás regressziós modelljeink eredményei szerint az átlagos jövedelem növekedése lényegesen nagyobb azoknál a cégeknél, amelyek diverz kapcsolatokkal rendelkeznek, de központi szerepet töltenek be a hálózatban is. Azt találjuk, hogy a nagyvállalatok jobban profitálnak a munkatársak hálózatában bezáródó háromszögekből, ami a kohézív kapcsolatok szerepét hangsúlyozza a komplex tudás megosztásában. Eredményeink rávilágítanak arra, hogy a vállalati növekedés a munkatársak személyes hálózatainak ökoszisztémájába ágyazódik be, ami megkönnyíti a vállalatok közötti információáramlást.

JEL: D85, L25, J62, O47

Kulcsszavak: munkatársi kapcsolathálózatok, kérdőíves felmérés, közösségi média, kapcsolat predikció, adminisztratív adat, hálózat szimuláció, vállalati növekedés

# Co-worker networks and firm performance

**Balázs Lengyel[1,2,*] - Guilherme Kenji Chihaya[3,4] - László Lőrincz[2,1] - Rikard Eriksson[3,5]**

1 Agglomeration and Social Networks Lendület Research Group, ELKH Centre for Economic- and Regional Studies, 1097 Budapest, Hungary
2 NETI Lab, Budapest Corvinus University, 1093 Budapest, Hungary
3 Department of Geography, Umea University
4 Institute for Analytical Sociology, Linköping University
5 Center for Regional Science, Umea University

* Corresponding author: lengyel.balazs@krtk.hu

Version 2021-05-07

**Abstract**
Firms and employees can benefit from information diffusion through social connections at other firms. Therefore, co-worker networks observed in collaborative projects or assumed from job co-occurrence have been analyzed in a wide literature ranging from management to economics, and economic geography. Yet, beyond case-studies, the actual information flows are seldom identifiable in these networks and previous focus on firm- or employee benefits was mainly limited to dyadic relations across firms. To address this gap, we simulate co-worker networks within firms from large-scale administrative data, for which we use parameters fitted to information networks that we collected with a survey and from social media profiles. Then, following all individuals through job moves over their career, we establish the dynamic co-worker network across firms of the entire ICT industry in Sweden. Fixed-effect regression models suggest that growth of average income is significantly higher in those firms that have diverse connections but are central to the network as well. We find that large firms benefit more from triadic closure in the co-worker network, stressing the role cohesive relations in sharing complex knowledge. Our results highlight that firm growth is embedded into the eco-system of co-worker networks that facilitate information flows across firms.

**Keywords**
co-worker networks, survey, social media, link prediction, administrative data, network simulation, firm growth

**Acknowledgements**
This work was funded by the Swedish Research Council (VR 2016-01803). Balázs Lengyel and László Lőrincz also acknowledge financial support received from Hungarian Scientific Research Fund (OTKA K-129207).

## 1. Introduction

The notion that firms are not isolated but instead are embedded in the surrounding information ecosystem has been present in economic thinking from the early work of Marshall (1920). Based on later contributions, the idea of firm-embeddedness has become mainstream in social sciences (Granovetter, 1985; Powell et al., 1996; Uzzi, 1997). More recently, techniques of network analysis have provided opportunities for scholars to map knowledge flows across firms (Broekel and Boschma, 2012; Giuliani 2007; Giuliani and Bell, 2005; Balland et al., 2016; Juhász and Lengyel, 2018) and demonstrate with case studies that the position of the firm in these networks is vital for its performance (Boschma and Ter Wal, 2007). However, due to limitations in data-access, large-scale analyses addressing knowledge interaction across firms have mostly focused on firm-level interactions (Walker et al., 1997) such as R&D&I collaborations (see for example, Broekel et al., 2015) or mergers and acquisitions (see for example, Shipilov, 2009). In this literature, the complex nature of firms' embeddedness in the social networks of their employees has remained overlooked despite strong claims that social interaction is a key channel through which learning processes occur (Arrow 1962).

Coworker relations are important channels of social interaction and inter-firm learning. Besides direct transfer of knowledge transmitted by labor flows (Almeida and Kogut, 1999; Boschma et al., 2009; Maliranta et al., 2009; Song et al., 2003), connections to colleagues developed at workplaces are important conduits of subsequent information access (Casper, 2007, Corredoria and Rosenkopf, 2010). Professional networks between co-workers (henceforth co-worker networks) that span company boundaries are important for firms because they improve matching through the diffusion of job-related information (Calvo-Armengol and Jackson, 2004; Granovetter, 1995; Hensvik and Nordström Skans, 2016), and facilitate inter-firm learning (Fleming et al 2007; Lőrincz et al, 2020; Ter Wal et al., 2016; Tóth and Lengyel, 2021). Since interaction related to jobs creates social bond and common understanding among co-workers (Storper and Venables, 2004), these contacts are often maintained across time and distance and have been found vital for continued knowledge inputs long after the co-workership is terminated (Agrawal et al, 2006; Breschi & Lissoni, 2009, Dahl & Pedersen, 2004).

Previous research on individual-level co-worker networks and firm performance is limited. Analyses of co-inventor relations have informed us that individual networks matter for firm innovation (Fleming et al., 2007; Ter Wal et al., 2016; Tóth and Lengyel, 2021). In labor economics, information flows on co-worker networks have been identified by demonstrating their efficiency in matching employees and jobs (Beaman and Magruder, 2012; Boza and Ilyés, 2020; Hensvik and Nordström Skans, 2016). In economic geography, Lengyel and Eriksson (2017) proposed a homophily-biased random network approach to generate the system of co-worker networks that can predict economic growth (Eriksson and Lengyel, 2019). Yet, the actual information flows are not identifiable in these latter networks.

In this paper, we contribute to this growing literature on co-worker networks with a new empirical framework that combines the systemic approach of generated co-worker networks with observed information flows. We first map real co-worker networks in firms by a survey on information flows in a local industry, and data collection from social media. We estimate the probability of information flows to fit parameters for individual, dyadic, and firm determinants. Then, using the parameters from link prediction, we simulate a series of co-worker networks within all national firms in the same industry using large-scale administrative data. Following individuals through job moves over their career, we establish individual links across firms. This allows for a fixed-effect panel regression, in which the position in the aggregate and dynamically changing co-worker network can be used to estimate the relation between co-worker networks and firm performance.

Our co-worker network data comes from 214 employees of 16 ICT (Information and Communication Technologies) firms in Umea, Sweden (Lőrincz et al, 2020). We find that homophily in age, sex, and education are present in these networks. Also, firm size and length of previous co-working experience had a significant correlation with the probability of links. Then, we extrapolated these parameter values to register data and generated inter-firm networks that represent the full set of ICT firms in Sweden.

Following this strategy, we find that the growth of average income is significantly higher in those firms that have diverse connections but are also central to this network. Our results suggest that large firms benefit more from co-worker ties that close triangles across firms, which highlights the role of cohesive networks in the transfer of complex knowledge. These findings altogether illustrate that inter-firm learning does not only mean learning from partners but rather happens in the ecosystem of social relations, in which each individual employee can be the mediator of knowledge transfer.

## 2. Data and Methods

Our empirical approach combines data collection through a survey and big data sources that are used to estimate the determinants of co-worker ties in firms. Next, we simulate co-worker networks in a large administrative dataset using the parameters of link prediction and trace co-worker links over time. This yields a dynamic inter-firm co-worker network that can be used to quantify network position of firms to predict firm-level covariates obtained from register data.

### 2.1 Data sources

We have conducted a survey data collection from 214 employees of 16 ICT firms in the Swedish city of Umeå (see Lőrincz et al. 2020 for details on the survey). Network data was collected via name-generator questions asking respondents to pick from a list of those co-workers, from whom they obtained essential work-related information. The survey collected further individual data on education, sex, and age. We asked the respondents to friend our account on LinkedIn that enabled us to collect their work and educational histories with their permission. This also allowed us to calculate the number of years that each pair of subjects worked at the same workplace. This information was then used for co-worker link prediction at workplaces with an explicit intention to use the parameters in network simulation from administrative data.

Matched employee-employer data comes from several Swedish administrative registers assembled into the ASTRID database hosted at Umeå University. It includes longitudinal records for every person in Sweden and contains information on employer, firm location, firm industrial classification, income, educational degree, field of education, date of birth, and sex, among others. The data covers the period 2001-2016 and enables us to replicate the network data format of the survey information. Our analysis uses data from firms in the following industries of the ICT sector (NACE Rev. 2 in parentheses): Manufacture of computers and peripheral equipment (26200); Computer consultancy activities (62020); Publishing of computer games (58210); Other software publishing (58290); Computer programming activities (62010); Computer consultancy activities (62020); Other information technology and computer service activities (62090); Computer facilities management activities (62030); Data processing, hosting and related activities (63110); Web portals (63120), Research and experimental development on biotechnology (72110), Other research and experimental development on natural sciences and engineering (72190). Over the period we have 60,566 unique firms in the sample ending with 18,084 firms and 134,910 workers in 2016.

### *2.2 Co-worker network estimation and network generation*

Using data from the survey and LinkedIn, we estimated the probability that two employees are linked by directed information flow ($P_{ij}$), using the formula:

$$P_{ij} = \gamma_0 + \gamma_1(Z_{i\epsilon f}, Z_{j\epsilon f}) + C_{ij} + S_f + \varsigma_f + \varepsilon_{ijf} \quad (1)$$

where *i* and *j* are employees at firm *f*, *Z* refers to their characteristics regarding age, sex, and education, $C_{ij}$ is the number of years that employees *i* and *j* have worked together previously, $S_f$ is the number of employees at firm *f*, $\varsigma_f$ is the random intercept of firm *f* assumed to be normally distributed and $\varepsilon_{ijf}$ is the error term. The equation is estimated by a logistic regression. We have chosen the random intercept model instead of the fixed effect model because it allows for firm specificities in the variation but does not require firm-specific parameters, which are not possible to replicate in the network simulation.

**Table 1.** Estimations of co-worker links

| | Coefficient | S.E. |
|---|---|---|
| Male-Male | 0.580*** | (0.102) |
| Female-Female | 0.829*** | (0.265) |
| Female-Male | -0.054 | (0.165) |
| University-University | 0.515*** | (0.133) |
| High school-High school | 0.489* | (0.282) |
| High school-University | 0.327** | (0.165) |
| Same Generation | -0.166* | (0.100) |
| Years Co-worked | 0.045* | (0.028) |
| Firm Size | -0.055 | (0.045) |
| Constant | -0.416 | (0.677) |
| N. of observations | 3,056 | |
| Log Likelihood | -1,786.386 | |
| Akaike IC | 3,594.773 | |
| Bayesian IC | 3,661.046 | |

Notes: *** p<0.001, ** p<0.01, * p<0.05

When estimating the likelihood of co-worker links, we indeed find that homophily strongly influence tie creation (Table 1). Compared to male-female links, ties between men and women, respectively, are more likely. Workers with similar education levels are more likely to share information, and high-school workers ask more help from colleagues with a university degree than the other way around. While age similarity deters information flows, longer mutual time spent within the same workplace increases the probability of information flows. Firm size does not seem to have a significant role in predicting the ties.

Next, we extrapolate the predictions from the survey to all other ICT firms in Sweden using register data. For each firm, we simulate the internal information networks by randomizing individual links using the formula

$$L_{ij,t} = \begin{cases} 1 \; if \; U\,(0,1) < \hat{P}\left(i_{g,a,e}, j_{g,a,e}, ij_{co-work,firmsize}\right) \\ \quad\quad\quad\quad 0 \; otherwise \end{cases} \tag{2}$$

where $L_{ij}$ is the simulated link between individuals *i* and *j*, $\hat{P}$ is the estimated probability based on parameters from Equation 1 and calculated from age, sex, and education of workers *i* and *j*, the number of years they worked together, and the firm size. $U(0,1)$ refers to the uniform distribution between 0 and 1.

Intuitively, we draw a random number from the uniform distribution and if this number is smaller than the estimated probability, we establish the tie. For the sake of simplicity, we assume that ties do not dissolve once formed and trace them over the subsequent years until *i* or *j* reached the age of 65 (see Lengyel and Eriksson, 2016).

Ties are kept fixed even if one or both co-workers in the dyad left the company. Such inter-firm mobility enables us to generate a dynamic network of companies in which two firms are connected if there is a past network connection between any of their workers. The weight of a tie between two firms equals the number of such connections between their workers. Formally:

$$w_{ab,t} = \sum_{ij,t} L_{ij,t} \; ; \; i\epsilon a, j\epsilon b \tag{3}$$

This above process yields a growing network because we add but do not delete individual edges year-by-year unless the worker exits the data (due to death, moving abroad etc. when we no longer can trace individuals in administrative data). Such growing networks are desirable as have been proven useful in fixed-effect panel regression specifications (Tóth and Lengyel, 2021; Eriksson and Lengyel, 2019). This process creates dynamic inter-firm networks in which two firms may be connected even if they do not exchange workers directly given that any of their workers have shared a workplace in the past (Lengyel and Eriksson, 2016). To account for stochastic variation in our simulated networks, we replicate the procedure above 25 times and estimate our models each time (see below on analytical strategy).

*2.3 Firm wage estimation framework*

There are different ways of proxying firm performance. Either as innovativeness by measuring innovation-rates or patenting (Herstad and Sandven 2020), or as productivity (Boschma et al 2009). In this paper we rather employ or readily available data on incomes due to a number of reasons: Not all firms patent or register innovations which might bias the sample while data on incomes are available for all firms (c.f., Eriksson 2009). Moreover, wages tend to be considered the best available proxy for worker productivity since the most innovative and productive firms usually are able to reward their employees with higher salaries (c.f., Kemeny and Storper 2015). Further, firm level output like productivity or innovativeness is registered on firm level and not for each given workplace in the firm. Although a majority of firms in this sample are single-plant firms, we also include workplaces that belong to larger firms with many branch-plants. Hence, for these workplaces, income is the most readily available proxy for performance and therefore we also use the term firm although in some cases the unit of analysis is a workplace within a firm. To assess the relationship between network position of firms and wage level, we measure wage-level as the logarithm of mean monthly income of all employees, calculated by dividing the annual income from wage and employment-related transfers (e.g., parental leave, sick leave) by twelve. Hourly compensation from work is unfortunately not available in our data, therefore we resort to this procedure. Income values were adjusted for inflation as of 2016 prices in Swedish crowns (10 SEK is equivalent to about 1 EUR).

We consider the following measures of network position: strength, closeness centrality, and Burt’s constraint. Strength (weighted degree, that is the number of co-worker links that connect the firm to others) represents the connectivity of the firm. Closeness centrality measures whether the firm’s position in the network is central versus peripheral, based on the inverse of the average distance of the selected firm from all other ones:

$$C_a^{cl} = \frac{n-1}{\sum_{a \neq b} \ell(a,b)} \qquad (4)$$

where $n$ is the degree of the node $a$, and $l(a,b)$ is the length of the shortest path between firms $a$ and $b$ in the network.

To measure how redundant firms’ ties are, we use Burt’s constraint measure. This is, defined as

$$C_a = \sum_{b \in V_{a,a \neq b}} \left( \sum_{q \in V_{a,k \neq a,b}} p_{a,b} + p_{a,k} p_{k,b} \right)^2 \qquad (5)$$

where $p_{ab}$ are proportional tie strengths, defined as

$$p_{a,b} = \frac{w_{a,b} + w_{b,a}}{\sum_{k \in V_{a,k \neq b}} (w_{a,k} + w_{k,a})} \qquad (6)$$

where $w_{ab}$ is the edge weights between nodes $a$ and $b$, $w_{ak}$ is the edge weights between nodes $a$ and *k*. High constraint values indicate that a node have redundant ties, while low values indicate having important, bridging ties in the network.

Our networks are generated from workers’ mobility between firms, and mobility can be directly related to wage levels. First, high wage-levels usually attract productive workers, while low levels may generate higher turnover-rates (e.g. Christensen et al. 2005; Holtom et al 2008). Second, attracting workers with high human capital generates positive knowledge spillovers, which leads to higher productivity and wage levels (Boschma et al 2009; Poole 2013). To keep track of these mechanisms, we control for the mobility of the workers, more precisely, for the incoming and outgoing human capital. For measuring human capital, we follow the method of Abowd, Kramarz, & Margolis (1999), and decompose worker-wages into the effect of observed traits as well as unobserved individual- and firm-specific effects, by estimating the following wage-equations including individual- and firm-specific fixed-effects:

$$wage_{m,a,t} = \propto + \beta z_{m,t} + \theta_m + \varphi_i + \varepsilon_{m,a,t} \quad (7)$$

where $wage_{m,i,t}$ is the logarithm of the wage of worker $m$ at firm $g$ in year $t$. $z_{m,t}$ include observable characteristics of workers: level of education, field of education, age and age squared. $\theta_m$ and $\varphi_i$ represent worker-level and firm-level fixed-effects respectively, while $\varepsilon_{m,a,t}$ is the error term.

From the Equation 6, we calculate human capital of the individual workers as the sum of the predicted observable traits and the worker-level fixed-effects:

$$HC_{m,t} = \hat{\beta} z_{m,t} + \hat{\theta}_m \quad (8)$$

Then, we can calculate the incoming and outgoing human capital for each firm ($HC_{a,t}^{in}$ and $HC_{a,t}^{out}$), as the sum of the human capital of workers who arrived at the firm, and of those who left the firm every given year.

When estimating the wage levels of firms by the network characteristics and the above controls, we used firm-level fixed-effect panel regressions, to control for unobserved heterogeneity across firms. Thus, due to the within-estimation we are able to assess whether wages within the firms increase over time as they achieve more favorable network positions. Since future wage-levels at firms are greatly determined by past realisations of wages, we add the previous year's wage-level to the right-hand side of the equation. Accordingly, we estimate the following regressions:

$$y_{a,t+1} = \propto + y_{a,t} + controls_{a,t} + \xi_a + \varepsilon_{at} \quad (9)$$

where $y_{a,t}$ represent the average log wage-level of the firm, *controls* include incoming and outgoing human capital, firm size (log number of employees), the share of female employees, and year dummies to control for time-specific events that influence all observations but are specific to certain years (e.g., macroeconomic trends and economy-wide changes in policy). $\xi_a$ stands for firm-level fixed-effects, and $\varepsilon_{at}$ is the error term. The baseline model (9) is followed by the "network" model (10), and an "extended" model (11), in which we included the interactions of the network variables with the firm size[1]:

$$y_{a,t+1} = \propto + \beta_1 Strength_{at} + \beta_2 C_{at}^{cl} + \beta_3 C_{at} + w_{a,t} + controls_{a,t} + \xi_a + \varepsilon_{at} \quad (10)$$

$$y_{a,t+1} = \propto + \beta_1 Strength_{at} + \beta_2 C_{at}^{cl} + \beta_3 C_{at} + \beta_4 Strength_{at} x\ Size_{at} + \beta_5 C_{at}^{cl} x\ Size_{at} + \beta_6 C_{at} x\ Size_{at} + w_{a,t} + controls_{a,t} + \xi_a + \varepsilon_{at} \quad (11)$$

We estimated the panel regressions (9-11) for each of the simulated networks, meaning that 25 different regressions were estimated for each model. We obtained pooled coefficient and

[1] We also estimated alternative versions of equations (9-11) omitting the reference year's wage levels from the right-hand side. Results considering the parameters of interest became highly similar in these specifications from which we can conclude that adding the lagged dependent variable does not cause any severe issues of time-series correlation.

standard error estimates by applying Rubin's rules (Rubin, 2004) for combining estimates from multiple imputations. Regression coefficients are pooled according to:

$$\underline{\beta} = \frac{1}{r}\left(\sum_{g=1}^{r} \theta_g\right) \tag{12}$$

where $\theta_g$ is the estimated coefficient for a regression model estimated network *g* of *r=25* simulated networks. In a similar vein, standard errors from *r* regression models are pooled using:

$$SE_{Pooled} = \sqrt{\frac{1}{r}\sum_{g=1}^{r} SE_g^2 + \left(1+\frac{1}{r}\right)\frac{\sum_{g=1}^{r}(\theta-\underline{\theta})^2}{r-1}} \tag{13}$$

Note, that the idea behind pooling standard errors (13) is that the combined variance is the sum of the average variance, and the variance of the parameter estimations between estimations, thus the combined standard errors also include these two components.

## 2.4. Descriptive statistics

Average monthly salary in the examined sectors were 29.46 thousand Swedish Crowns in 2016 values, about 25% higher than the average full-time wages in the private sector. The average firm in these sectors had approximately 7 employees, indicating the dominance of SMEs in these sectors, however, the large standard deviation signifies the presence of larger enterprises as well. The incoming and outgoing human capital measures are roughly half the monthly salary, which is about 5% of the yearly one, indicating an approximately 5% annual turnover in the labor force. Outgoing human capital is higher than incoming. As the salary of workers increases by experience, this implies that they on average earn more, when they exit compared to when they entered the companies. The share of women shows that a substantial male dominance exists in the IT sector even in Sweden (Table 2).

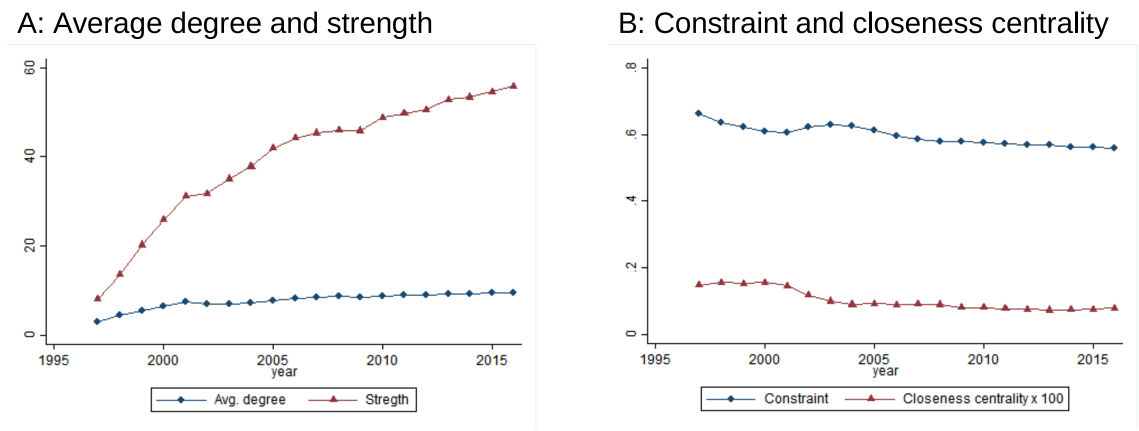
**Table 2.** Descriptive statistics

| Variable | Mean | Std. Dev. | Mean | Std. Dev. | N. obs. (firm x year) |
|---|---|---|---|---|---|
| | original | | log transformed | | |
| mean income monthly | 29.46 | 19.82 | 1.35 | 0.43 | 238,218 |
| HC in | 13.30 | 69.00 | 0.47 | 0.66 | 238,218 |
| HC out | 15.90 | 178.86 | 0.49 | 0.67 | 238,218 |
| size (employees) | 7.23 | 39.94 | 0.55 | 0.40 | 238,218 |
| % of women | 17.47 | 30.16 | 0.57 | 0.79 | 238,218 |
| | Mean | Std. Dev. | | | N. obs. (firm x year x simulation) |
| Network strength | 47.05816 | 118.4935 | | | 4,720,004 |
| Closeness centrality | 0.00087 | 0.00041 | | | 4,720,004 |
| Network constraint | 0.58 | 0.367 | | | 4,720,004 |

The bottom panel of Table 2 describes the network indicators of the firms over the examined period, derived from our 25 simulations. As we build the network from the mobility of employees between firms, this network is growing over the observed years. This feature is reflected by the increasing average degree and strength values in Figure 1A. However, the

Constraint, and Closeness centrality measures, which are our key interests, tend to decrease as the network grows (Figure 1B). Note that in our regressions (Eq. 9-11) we control these network trends by including year dummies.

**Figure 1.** Evolution of average degree and mean Strength (A) and Constraint and Closeness centrality (B) over time




## 3. Results

The results from the fixed-effect panel regressions on wage dynamics are presented stepwise according to equations 9-11 (Table 3). First, only the control variables are regressed (Model 1) followed by Model 2 in which the network variables are included. Finally, Model 3 presents interaction effects. Although one could argue that the first model will suffer from omitted variable bias, it is essential to get a sense of whether the control variables behave as expected. Shortly, this is the case. High income growth in the past is positively correlated with future income growth and losing highly skilled employees also deter income growth more than recruiting new high-skilled workers does. Growing firms are however experiencing a higher per-capita income growth. Despite having an expected negative sign due to the general wage penalty for women, the estimate on the share of women is not significant. Thus, despite that the ICT sectors often are being portrayed as male-dominated (James 2017) we do not observe a gender wage-gap within firms in the sense that if a male worker is replaced by a female one, the average wage levels do not change.

Moving on to the main results (Model 2), our results indicate that Closeness centrality and Constraint are related to income dynamics in the expected direction. Firms tend to increase wages after becoming more central in the social network. In addition, if their position includes more non-redundant linkages, such as bridging edges, they also tend to be more likely to experience faster income growth. On the other hand, the weighted number of connections in itself is not associated with wage increase or decrease.

**Table 3.** Estimations of wage dynamics

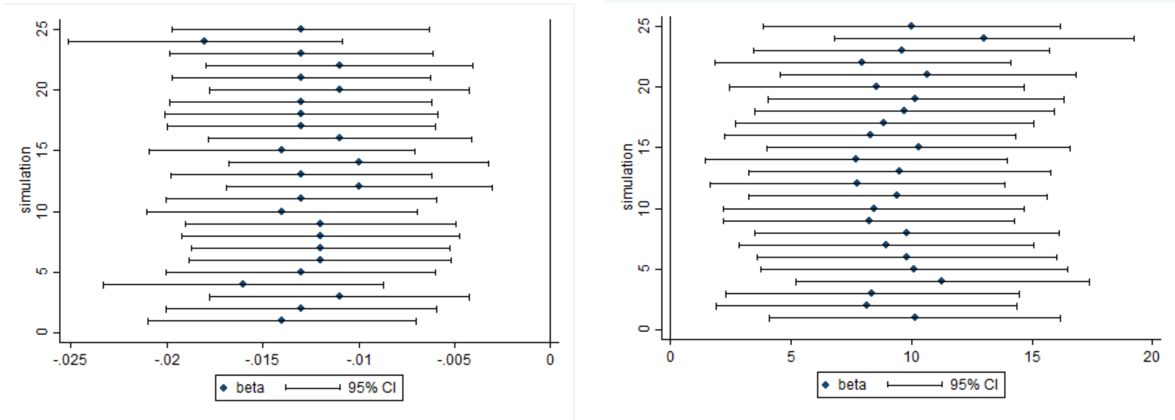
| | 1: baseline<br>mean income (log) | 2: network<br>mean income (log) | 3: extended<br>mean income (log) |
|---|---|---|---|
| *Firm characteristics (t-1)* | | | |
| mean income (log) | 0.297*** | 0.297*** | 0.297*** |
| | (0.0025) | (0.0025) | (0.0026) |
| HC in (log) | 0.0004 | 0.0004 | 0.0004 |
| | (0.0010) | (0.0010) | (0.0010) |
| HC out (log) | -0.0045*** | -0.0045*** | -0.0045*** |
| | (0.0013) | (0.0013) | (0.0013) |
| Share women (log) | -0.0012 | -0.0012 | -0.0012 |
| | (0.0012) | (0.0012) | (0.0012) |
| Size (log N employees) | 0.0492*** | 0.0494*** | 0.0465*** |
| | (0.0045) | (0.0045) | (0.0045) |
| *Network position (t-1) (information network)* | | | |
| Strength | | 0.0000 | 0.0000 |
| | | (0.0000) | (0.0000) |
| Closeness centrality | | 9.392** | 11.314* |
| | | (3.391) | (5.7415) |
| Constraint | | -0.0133** | -0.0258*** |
| | | (0.0041) | (0.0068) |
| *Interactions (t-1)* | | | |
| Strength x Size | | | 0.0001 |
| | | | (0.0000) |
| Closeness centr x Size | | | -3.2012 |
| | | | (6.8029) |
| Constraint x Size | | | 0.0200* |
| | | | (0.0082) |
| N (firm x year) | 176,586 | 176,586 | 176,586 |
| N (firms) | 39,489 | 39,489 | 39,489 |

Notes: Pooled coefficients (and standard errors in parentheses) of 25 regressions with firm fixed-effects. Additional controls: year dummies. *** p<0.001, ** p<0.01, * p<0.05

Figure 2 illustrates the estimated coefficients of our key network parameters over the 25 regressions. Both estimations are stable across the average values (with only one, number 24 being a bit out of the range). Still, both the constraint and the closeness centrality parameters are significant in each of the 25 regressions at p<0.05 which indicate that we indeed succeed to simulate this.

**Figure 2.** Coefficients and confidence intervals of the constraint (A) and closeness centrality (B) from the 25 simulations

A: Constraint B: Closeness centrality




Naturally, the size of a firm might influence both the structure of networks but also to what extent different links are more or less beneficial. For example, a small firm, compared to a large firm, is in greater need of belonging to a network in which new external knowledge can be retrieved while a larger firm potentially can internalize such information flows by enjoying economies of scale. On the other hand, a larger firm, as an effect of mere size, can potentially have much more extensive networks compared to a smaller firm with just a handful of employees.

When assessing the heterogeneity of these network covariates by firm size in Model 3, we find no significant interaction for Closeness centrality. This implies that being in the center of the network seems to be equally important for small and large firms. We find, however, a significant interaction effect for the Constraint measure. This suggests that having non-redundant networks is more beneficial for performance but that the role of non-redundancy decreases by firm size. Hence, smaller firms on average benefit more from having non-redundant ties than large firms do. For the latter having more redundant ties are less detrimental for performance proxied by income.

## 4. Discussion

The purpose of this paper was to contribute to our understanding of, and in such case how, co-worker networks provide firms with performance enhancing information that increase competitiveness. This was accomplished by a novel empirical framework that combined survey information on stated information channels in a group of regional ICT-sectors administrative data for the whole of Sweden for the period 2001-2016. By using the parameters from the link prediction, we then simulated a series of co-worker networks within all national firms. The detailed data allowed us to follow individuals through job moves over their career and establish and aggregate links across firms.

Our longitudinal data allowed us to estimate a fixed-effect panel regression, in which the position in the aggregate co-worker network can be used to estimate the relation between co-worker networks and firm performance.

Our findings suggest that homophily in age, sex and education are present in these networks. Apart from this, also firm size and length of previous co-working experience had a significant correlation with the probability of links. Our fixed-effect regressions that were employed to estimate the relation between co-worker networks and firm performance suggest that the growth of average income is significantly higher in those firms that have diverse connections but are also central to this network. These processes however seem to vary dependent on firm size. Our results suggest that large firms benefit more from co-worker ties that close triangles across firms compared to small firms in which non-redundant ties are more beneficial.

In conclusion, the paper has made two distinct contributions to the literature. First, methodologically this is, to our knowledge at least, the first study that has combined qualitative information on actual information networks with large-scale administrative data to generate co-worker networks by means of simulating tie creation and then applying Rubin's rules to obtain pooled coefficient and standard error estimates. Secondly, in so doing, we provide direct evidence on the assumed effect of co-worker networks on firm performance. Contrasting previous studies that either assumes that everyone knows everyone (Hensvik and Nordström-Skans, 2016), more homophily-biased approaches assuming a certain clustering of contacts within firms based on individual characteristics (Lengyel and Eriksson, 2017), or approaches conflating population density with network density and the potential for knowledge spillovers (Storper and Venables 2004), this paper has provided unprecedented insights to the micro-channels of knowledge spillovers.

More competitive firms in knowledge based and innovative activities like the ICT-sectors benefit from having diverse connections characterized by non-redundant information. This is particularly evident for smaller firms as the detrimental effect of having many closed triangles in information channels is less severe for larger firms with presumably higher internal capacities. This provides direct evidence on the social dimension of firm-learning as all learning processes are derived from social interaction (Arrow 1962). Firms that are well-connected in the network are more likely to get access to valuable information that can enhance performance. Since this is partly driven by labor mobility and because labor mobility predominantly is a local process, we can assume that this is a key-mechanism of successful industrial agglomerations as indicated at industry-level by Eriksson and Lengyel (2019).

In all, these findings open up for further analyses. For example, the evident homopihily of contacts in relation to sex could be scrutinized further to assess different possibilities for career progression depending on network position. We have neither explicitly accounted for the type of contacts each firm has. Thus, further studies could combine this firm-level information with information on the type of activities (e.g., sectors) that are more beneficial. In particular, assessing that in relation to the geography of these networks would provide unpreceded insights on the knowledge flows in regional clusters.

## References

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251-333.

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

Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, *45*(7), 905-917.

Arrow, K. J. (1962). The economic implications of learning by doing. Review of Economic Studies 29(3), 155–73.

Balland, P. A., Belso-Martínez, J. A., & Morrison, A. (2016). The dynamics of technical and business knowledge networks in industrial clusters: Embeddedness, status, or proximity?. *Economic Geography*, *92*(1), 35-60.

Beaman, L., & Magruder, J. (2012). Who gets the job referral? Evidence from a social networks experiment. *American Economic Review*, *102*(7), 3574-93.

Broekel, T., & Boschma, R. (2012). Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, *12*(2), 409-433.

Broekel, T., Fornahl, D., & Morrison, A. (2015). Another cluster premium: Innovation subsidies and R&D collaboration networks. *Research Policy*, *44*(8), 1431-1444.

Boschma, R., 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*(2), 169-190.

Boschma, R. A., & Ter Wal, A. L. (2007). Knowledge networks and innovative performance in an industrial district: the case of a footwear district in the South of Italy. *Industry and Innovation*, *14*(2), 177-199.

Boza, I., & Ilyés, V. (2020). Decomposition of co-worker wage gains. *IZA Journal of Labor Economics*, *9*(1).

Breschi, S., & 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.

Calvo-Armengol, A., & Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American Economic Review*, *94*(3), 426-454.

Casper, S. (2007). How do technology clusters emerge and become sustainable?: social network formation and inter-firm mobility within the San Diego biotechnology cluster. *Research Policy*, *36*(4), 438-455.

Christensen, B. J., Lentz, R., Mortensen, D. T., Neumann, G. R., & Werwatz, A. (2005). On-the-job search and the wage distribution. *Journal of Labor Economics*, *23*(1), 31-58.

Csáfordi, Z., Lőrincz, L., Lengyel, B., & Kiss, K. M. (2020). Productivity spillovers through labor flows: productivity gap, multinational experience and industry relatedness. *The Journal of Technology Transfer*, *45*(1), 86-121.

Corredoira, R. A., & Rosenkopf, L. (2010). Should auld acquaintance be forgot? The reverse transfer of knowledge through mobility ties. *Strategic Management Journal*, *31*(2), 159-181.

Dahl, M. S., & Pedersen, C. Ø. (2004). Knowledge flows through informal contacts in industrial clusters: myth or reality?. *Research Policy*, *33*(10), 1673-1686.

Eriksson, R. H., & Lengyel, B. (2019). Co-worker networks and agglomeration externalities. *Economic Geography*, *95*(1), 65-89.

Fleming, L., Mingo, S., & Chen, D. (2007). Collaborative brokerage, generative creativity, and creative success. *Administrative Science Quarterly*, *52*(3), 443-475.

Giuliani, E. (2007). The selective nature of knowledge networks in clusters: evidence from the wine industry. *Journal of Economic Geography*, *7*(2), 139-168.

Giuliani, E., & Bell, M. (2005). The micro-determinants of meso-level learning and innovation: evidence from a Chilean wine cluster. *Research Policy*, 34(1), 47-68.
Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360-1380.
Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481-510.
Hensvik, L., & Skans, O. N. (2016). Social networks, employee selection, and labor market outcomes. *Journal of Labor Economics*, *34*(4), 825-867.
Herstad, S. J., & Sandven, T. (2020). A closer look at the relationship between innovation and employment growth at the firm level. *Journal of Evolutionary Economics*, *30*(2), 375-399.
Holtom, B. C., Mitchell, T. R., Lee, T. W., & Eberly, M. B. (2008). Turnover and retention research: a glance at the past, a closer review of the present, and a venture into the future. *Academy of Management annals*, *2*(1), 231-274.
James, A. (2017). *Work-life advantage: Sustaining regional learning and innovation*. John Wiley & Sons.
Juhász, S., & Lengyel, B. (2018). Creation and persistence of ties in cluster knowledge networks. *Journal of Economic Geography*, *18*(6), 1203-1226.
Kemeny, T., & Storper, M. (2015). Is specialization good for regional economic development?. *Regional Studies*, *49*(6), 1003-1018.
Lengyel, B., & Eriksson, R. H. (2017). Co-worker networks, labour mobility and productivity growth in regions. *Journal of Economic Geography*, *17*(3), 635-660.
Lőrincz, L., Chihaya, G. K., Hannák, A., Takács, D., Lengyel, B., & Eriksson, R. (2020). Global connections and the structure of skills in local co-worker networks. *Applied Network Science*, *5*(1), 1-20.
Maliranta, M., Mohnen, P., & Rouvinen, P. (2009). 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.
Poole, J. P. (2013). Knowledge transfers from multinational to domestic firms: Evidence from worker mobility. *Review of Economics and Statistics*, *95*(2), 393-406.
Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 116-145.
Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81). John Wiley & Sons.
Song, J., Almeida, P., & Wu, G. (2003). Learning–by–hiring: When is mobility more likely to facilitate interfirm knowledge transfer?. *Management science*, *49*(4), 351-365
Storper, M., & Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, *4*(4), 351-370.
Ter Wal, A. L., Alexy, O., Block, J., & Sandner, P. G. (2016). The best of both worlds: The benefits of open-specialized and closed-diverse syndication networks for new ventures' success. *Administrative Science Quarterly*, *61*(3), 393-432.
Tóth, G., & Lengyel, B. (2021). Inter-firm inventor mobility and the role of co-inventor networks in producing high-impact innovation. *The Journal of Technology Transfer*, 117-137.
Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, 35-67.

Walker, G., Kogut, B., & Shan, W. (1997). Social capital, structural holes and the formation of an industry network. *Organization Science*, *8*(2), 109-125.