*István Boza[1] and Virág Ilyés[2],**

Decomposition of co-worker wage gains

Abstract

We address the presence, magnitude, and composition of wage gains related to former co-workers and discuss the mechanisms that could explain their existence. Using Hungarian linked employer–employee administrative data and proxying actual co-workership with overlapping work histories, we show that the overall wage gain attributable to former co-workers consists of multiple elements: a contact-specific, an individual-specific, a firm-specific and a match-specific component. Former co-workers, besides the direct effect of their presence, may funnel individuals into high-paying firms, enhance the sorting of good quality workers into firms, and may contribute to the creation of better employer–employee matches. By introducing and applying a wage-decomposition technique, we demonstrate that there are non-negligible differences between linked and market hires in all empirically separable wage elements. By focusing on specific scenarios, we provide additional empirical evidence in favor of employee referral and information transmission as the main drivers of co-worker gains.

1   Department of Economics and Business, Central European University, Budapest, Hungary.
     Email: bozaistvan@gmail.com
2   Doctoral School of Sociology and Communication Science, Corvinus University of Budapest, Budapest, Hungary.
     Email: ilyesvirag@gmail.com

1 Introduction

The group of former co-workers forms an essential part of our social networks. As shown by early survey-based evidence, one's co-worker acquaintances can be essential sources of job-related information and they may also play an important role in the job-acquiring process (Corcoran et al., 1980; Granovetter, 1995; Holzer, 1988). Besides this type of studies, which typically exploited self-reported information about the individuals' job search process, in recent years, several studies used administrative registers to address the labor market effects of co-worker networks. Although having their limitations, such as the lack of direct information on social links or hiring methods, these datasets contain precise and reliable information about employment and wages that can be utilized to bypass these shortcomings. Using various techniques, recent studies showed that former co-workers can positively affect different individual labor market outcomes such as hiring probabilities (Cingano and Rosolia, 2012; Glitz, 2017; Saygin et al., 2019), tenure length and turnover (Glitz and Vejlin, 2019), and quite notably, wages (Glitz and Vejlin, 2019; Hensvik and Skans, 2016). The explanations for the existence of these beneficial effects mostly highlighted the role of two mechanisms: information transmission and employee referral.

In this article, we address the presence and magnitude of wage gains related to former co-workers and discuss the mechanisms that could potentially drive them. In our empirical estimations, we rely on administrative data from Hungary and use former co-workership as a proxy for actual social connections. Using a wage-decomposition technique, we document not only an overall wage gain of those job-switchers who have a former co-worker present in the receiving firm upon entry but also show that there are non-negligible differences in all empirically separable wage elements, namely, in the individual-specific, firm-specific, and match-specific components as well.

Studies that utilized a similar approach to assess the wage effects of former co-workers documented that gains can be mainly attributed to referral activity (Dustmann et al., 2016; Glitz and Vejlin, 2019). However, a few other papers revealed additional channels through which gains are generated. Hensvik and Skans (2016) showed that homophily in co-worker networks can lead to the selection of better individuals into firms. Schmutte (2015), on the other hand, established that selection to high-wage firms is also prevalent. Furthermore, Eliason et al. (2019) found that referral is more likely to happen when the applicants are of better quality and their social contacts' firm has higher wages.

We contribute to the literature of co-workers, employee referral, and wage differences in three ways. First, by being the first to document the presence of wage gains commonly attributed to the referral activity of former co-workers through the estimation of a two-way fixed effects wage equation on starting wages. We also claim that the gain estimated this way consists of two distinct factors: the *presence* effect of referral—which assumes the continuous presence of a referrer—and the selection of individuals into better matches. Although these mechanisms are empirically indistinguishable with our proposed methodology and data, the distinction is important for theoretical clarity. Second, to assess the presence and relative importance of selection channels in overall wage gains in detail, we augment and apply the decomposition method proposed by Woodcock (2008). To interpret our findings, we link differences in wage components to the established theories in the referral and co-worker literature. Finally, to reinforce our arguments, we provide additional empirical evidence by focusing on scenarios where referral activity is expected to be more prevalent, or conversely, where it is considered less probable.

To identify the effects of co-workers, ideally, we would compare hiring events to counterfactual observations of the same worker entering the same firm, but without/with a connection at the firm. As such variation is not present in the data, we control for observed and unobserved firm and individual heterogeneity by using a two-way fixed effects approach. We find a 2.1% wage gain for male workers, which could either reflect productivity sorting or other aspects of referral. This gain is accompanied by a 1.7% and 0.9% wage advantage attributable to better worker and average firm quality, respectively, that is high-quality employees are sorted into firms where co-workers are present and workers with former co-worker links are sorted into high-wage firms. These better firms, however, tend to hire high-quality workforce even without the co-worker links. The superior skills of new hires will be responsible only for a 1.3% wage advantage relative to market hires. The remaining 0.4% difference in worker effects is coming from an already established assortativeness among the involved firms and high-quality workers. Selection into better firms is more substantial when it is compared to the individuals' own work history, which typically consists of a somewhat inferior firm pool. The latter difference dampens the 1.2% within-individual gain by 0.3%. Considering female workers, most of the gains are attributable only to the selection of high-quality workers both in absolute and relative terms. Regarding occupational heterogeneity, we observe that two-way fixed effects parameters are generally stronger and individual selection is weaker in higher occupations. Moreover, the presence of firm selection is stronger in skilled occupations with stronger educational requirements. When relying on mass layoffs as exogenous sources of variation, we found similar results. Based on the implications of the theoretical literature and some reasonable assumptions, we interpret these figures as a result of referral and information transmission.

We supplement these arguments by showing that referral-related wage gains are stronger when the contact is of relatively higher occupation, had a longer tenure at the receiving firm, or if the length of the previous co-working spell with the job entrant was longer. We try to identify the referrer-dependent (presence) effects from separations of referrers and the prevalence of various occupation-specific skills. We find only small and insignificant differences, which may suggest that match-specific selection accounts for a substantial portion of referral-related gains.

The rest of the article is structured as follows. Section 2 summarizes previous empirical and theoretical literature and based on those studies it systematically presents the channels through which wage differences could be generated. Section 3 establishes our model and proposes a decomposition strategy. Section 4 presents the utilized dataset, the definition of co-worker links and discusses identification issues. Section 5.1 contains the main results of the decompositions, Section 5.2 provides additional evidence by utilizing exogenous job losses, while Section 5.3 presents alternative estimation strategies for capturing referral and information transmission effects. Finally, Section 6 concludes.

2 Background

2.1 Mechanisms and possible explanations of wage gains

The literature identifies two mechanisms through which former co-workers (and in some cases, other social contacts) might shape the individuals' labor market outcomes: information transmission and employee referral. The former refers to the phenomenon that former co-workers

might have access to relevant work-related information, which they can pass on to job-seekers. Employee referral, on the other hand, covers those cases when employees of certain firms (referrers) bring together their acquaintances (applicants) and the vacancies at their companies. The main difference rests in the direction of information flows. In the former case, only job-seekers receive information about the quality of some potential employers. However, in the latter case, information about worker type based on the shared co-working experience is also revealed to the employer in the form of recommendation.[1] To this distinction, we would add an additional layer of cases, when, upon hiring a new applicant, the referrer continues to act as a provider of information, either about the applicant's behavior to the employer or about firm-specific knowledge to the new co-worker. While keeping the above distinction in mind, we collect and systematically review various potential components of wage gains generated by former co-workers and aim to map the theories that might explain their existence.

The first component of co-worker wage gains consists of those elements, which essentially depend on the presence of a referrer. The related theories typically utilize the relationship between referrers and applicants. One group of such explanations is related to the mitigation of the employers' monitoring costs (Bartus, 2001; Kugler, 2003). Referrers can affect the performance of the newly hired workers both directly—by voluntarily monitoring their effort (Ekinci, 2016; Saloner, 1985; Smith, 2005)—and indirectly, if the applicants increase their productivity to compensate the referrers' favor (Smith, 2005). Also, referrers might have an important role in the integration of the workforce, as their presence might support smooth knowledge sharing and better cooperation at work (Castilla, 2005; Fernandez et al., 2000). The enhanced productivity of workers and lower monitoring costs could increase the firm's profits, but it is not trivial whether the firm shares the emerging rent with the applicant. If the firm does so, we will observe a wage advantage of referred workers. For the sake of brevity, we refer to everything that is dependent on the active presence of a referrer and is perceived, valued, and compensated by the firm as *presence* effects.[2]

Besides the monetary benefits attributable to the above mechanisms, wage gains might originate from three types of selections as well: those based on match-specific productivity, worker-specific general skills, and firm-specific wage levels. Gains attributable to these selections, which capture previously existent productivity differences, are essentially different from referrer-dependent effects, as those actually increase the worker's productivity. In understanding the detailed role of co-workers in the labor market, we believe that the description of these selections is equally important as focusing only on causal channels.

First, referral activity might facilitate the sorting of workers into better employer–employee matches. The presence of such synergy implies a higher wage relative to both the firm's wage level and the individual's outside options.[3] Dustmann et al. (2016) showed that the wage prospects of nonreferred workers are more uncertain as their match-specific productivity is not revealed in the hiring process. Therefore, they will potentially turn down job offers that would be good matches, leading to a higher expected match element for referral hires.

---

1   Referral without informing the applicant may happen, but is rather unlikely.

2   Favoritism can be also considered a source of these gains as the applicants only acquire wage gains if a particular referrer resides at their new company (Bian et al., 2015).

3   Employers could, however, withhold the gains from these productivity improvements. A firm mitigating a moral hazard problem with efficient wages may prefer to hire workers through referral, as social factors already incentivize them to work hard. Thus, the wages of such applicants could be lowered. (Dhillon et al., 2015).

However, the emergence of better matches could happen even without the active participation of a referrer if employees pass information to only those who would be a good fit for a given vacancy at their firms.

The use of employee referrals might also promote the selection of those workers who generally have better skills and would earn more at any firm compared to someone with similar observable characteristics.[4] As referrers can decrease screening costs either by providing information about their former co-workers or by signaling worker quality with their own productivity based on the assumption of network homophily in productivity (Hensvik and Skans, 2016; Montgomery, 1991; Munshi, 2003), they can contribute to the reduction of information asymmetry about the general characteristics of applicants.[5] This way firms may avoid low-quality workers and, on average, hire better-quality applicants, even if they are not better-matched ones (Saloner, 1985; Ullman, 1966).[6]

Selection into high-wage firms, on the other hand, is mainly driven by information transmission. Former co-workers can be good sources of job offers (Calvo-Armengol and Jackson, 2004, 2007; Granovetter, 1995), and their information might mitigate the job-seekers' uncertainties about the possible employers (Tate, 1994; Wanous, 1980). By choosing from a larger set of vacancies, the expected quality of one's new firm could be higher. However, we note that positive firm selection could be also observed if, on average, higher-wage firms rely on the use of referrals.

We suppose that the above selections and the role of presence effects relate to information transmission and referral mechanisms in the following way. Firm selection is mainly driven by information transmission, but employee referral might also account for such gains if it dominantly happens in high-wage firms. Individual selection, we believe, is only present if employee referral happens either through direct (recommendation) or through indirect signals (homophily). Match selection could be a product of both mechanisms but is probably much more prominent in cases of active referral (Dustmann et al., 2016). Finally, presence effects emerge only when the referral is followed by other, continuous actions on the referrer's side as well. When decomposing the wage gains attributable to former co-workers, we will rely on the above framework to interpret the results.

2.2 **Empirical evidence**

In this section, we survey recent empirical evidence from papers that are based on matched employer–employee administrative data and focus on wage effects of various social contacts. While some papers aim to estimate the direct effects of employee referral or provide evidence on information transmission through networks, others are especially after the selections in the labor market produced by referral and job information networks. Our paper is related to both lines of research, both in theoretical approach and the utilized methods as well.

4　We suppose that information transmission in itself cannot be accountable for such selection. When their contribution remains hidden to the firm, workers rather share work-related information either to all of their relevant acquaintances or to only those who would be a good fit for the specific opening.

5　An employer could also assume that homophily is present not only regarding general skills but also match-specific ones. Wage premium paid based on this assumption would enhance the previously discussed match selection.

6　While employers could share gains from the reduced screening costs with the applicants through higher wages, this scenario is rather unlikely, as firms usually only incentivize their referrers, by one-time bonuses. Based on industry interviews we conducted, even these practices were not yet commonly utilized during the time frame of our study.

To study the role of employee referrals, Glitz and Vejlin (2019) constructed an indicator of events when former co-workers have reunited at a new firm with one of them arriving earlier. After showing that the number of such events in Denmark is higher than what random network forming would suggest, they interpreted these instances as potential cases of referral. They found a 4.6% wage advantage attributable to the presence of former co-workers after controlling for firm fixed effects, but not accounting for individual heterogeneity.[7] Besides, they also demonstrated that the initial wage gains of the referred workers decline over time, and in the long run, they eventually end up with lower wages than those who were hired through the external market.

Earlier, Hensvik and Skans (2016) provided similar evidence on former co-workers' effects on wages and assessed the role of homophily in terms of abilities of workers as a potential driver of individual selection. Using Swedish administrative data, including military test scores as a proxy for individual productivity, they showed that linked workers can earn 3.6% more compared to other new hires in the same establishment. Additionally, they demonstrated that the wage premium of the connected employees increases as the incumbent workers' abilities improve. This indicates that from the firms' perspective, current employees' productivity might unintentionally signal the quality of their acquaintances. The results also support the idea that network inbreeding might contribute to the generation of wage inequalities.

Dustmann et al. (2016) investigated the effects of referral on wages and turnover rates by using German data. They used the share of workers with the same ethnicity at the firms at the time of hiring as a proxy and also a direct indicator of referral coming from survey data. Their model of wages incorporated both individual and firm fixed effects, which account for the nonrandom sorting patterns of workers to firms alongside unobserved worker and firm characteristics. Their findings suggest a 3.3% wage gain by directly measured referral, potentially generated by the better matches among employers and linked hires.

Focusing more on the role of information transmission, Cingano and Rosolia (2012), Glitz (2017) and Saygin et al. (2019) investigated the co-worker network's capability of generating job offers and its impact on the reemployment outcomes of displaced workers based on the model of Calvo-Armengol and Jackson (2004, 2007). Their results demonstrated that an increase in the share of employed former co-workers comes with a higher re-employment rate of displaced workers, suggesting information transmission through the co-worker networks. Furthermore, Saygin et al. (2019) also found a significant difference between the displaced workers' pre- and post-displacement wage outcomes when the share of employed former co-workers in high-wage firms was high. This result is in line with our notion about information transmission's effect on firm selectivity.

Additionally, some papers provided evidence for the presence of individual and firms selections. Using US data, Schmutte (2015) showed that job-seekers are more likely to become co-workers of their neighbors from the same block as the individual than those from their broader neighborhood. After estimating an AKM (after Abowd, Kramarz and Margolis, 1999) decomposition of wages, he also demonstrated that referrals are more likely to happen when the applicants have better skills or when the referrers work at high-wage firms. He also argued

---

7   This difference also includes gains related to the superior unobserved quality of workers hired this way.

that employee referral in itself cannot explain this set of results, and that information transmission over the job information network also has to play a critical role.

Besides additional evidence on selection patterns and homophily, inequality consequences are also documented in the study of Eliason et al. (2019). The authors constructed a proxy of the local labor market for displaced workers by linking their closing firm to workplaces where the former co-workers of displaced workers were employed at the time of the plant closure. Comparing the role of social links in increasing hiring probabilities by levels of previously obtained AKM-style individual and firm fixed effects, they found that social ties might induce positive sorting. High-wage job-seekers tend to have links with high-wage workers who more likely to work at high-wage firms. The combination of homophily and positive assortative matching could then increase inequalities. However, they also showed that the causal impact of ties on hiring probability is the strongest for low-wage firms, which eventually leads to a lower level of sorting inequality. As directly assessing assortativeness is out of the scope of our paper, our main takeaway from their work is that referral may be more prominent in low-wage firms, attenuating the firm selection patterns generated by information transmission.

In this article, we focus on former co-worker contacts' effect on entry wages by relying on a proxy like Hensvik and Skans (2016) and Glitz and Vejlin (2019), and using multiway fixed effects approach similar to Dustmann et al. (2016). However, we utilize a framework that can also capture selections induced by co-workers. To do this, we improve upon and use the decomposition of Woodcock (2008) to assess selection mechanisms both in absolute and relative terms. In the process, we rely on AKM firm and person effects as measures of employer and worker quality, similarly to Schmutte (2015) and Eliason et al. (2019). Therefore, our proposed framework attempts to assess the direct and indirect consequences of co-worker networks at the same time. We find evidence for both wage gains after controlling for individual and firm heterogeneity like Dustmann et al. (2016)—which, we add, could still incorporate match selection and presence effects as well—and also for the presence of individual and firm selections as Hensvik and Skans (2016) and Schmutte (2015), respectively. Furthermore, we show that selections are mainly driven by their respective within components: linked workers get access to higher premium firms compared to where they usually work, and firms can increase the quality of their worker pool with referral hires.

3 Model and Empirical Strategy

To investigate the mechanisms discussed in Section 2.1, we estimate differences in specific wage components. We start by introducing an AKM model of wage-setting (Abowd et al., 1999), augmented with match effects similar to Woodcock (2008). Our wage equation also includes the effect of the presence of a referrer, θ, as a wage-determining factor.

$$w_{ijt} = \alpha + \theta T_{ijt} + \beta_X X_{it} + \beta_Y Y_{jt} + \beta_Z Z_{ijt} + \delta_i + \gamma_j + \mu_{ij} + \pi_t + \varepsilon_{ijt} \tag{1}$$

In Eq. (1), w_{ijt} denotes the starting wage earned by person i at firm j at time t. X_{it} contains the observable characteristics of the individual, such as age and education. Y_{jt} comprises the properties of the firm, such as sector and ownership. Finally, Z_{ijt} includes variables corresponding to the actual employment spell of individual i at firm j, among other occupation and form of contract. One such factor is an indicator of whether the given worker has obtained the job

through a social contact: T_{ijt}. However, this latter variable is rarely observed directly and is usually substituted by a proxy, which indicates whether an individual has a co-worker at a new firm upon entry with whom they had worked together earlier.

Besides these observable characteristics, many unobservable factors can alter an employee's starting wage at a new job. We suppose that these features, namely, the latent quality of the individual (δ_i), the wage levels of firms (γ_j), and the quality of the employer–employee match (μ_{ij}) are constant over time. Seasonal and trend effects (π_t) may also affect wages over a longer period. All other factors make up the independent error term with zero expected value (ε_{ijt}).

3.1 **Identification of match and presence effects**

The proper estimation of the full model is, however, infeasible. To obtain the match effects, we would have to compare multiple entries to the same firm by the same person. Although such a scenario occurs sometimes, gains estimated from comparing these observations could also reflect, for instance, the presence of firm-specific knowledge. Therefore, we prefer to omit these cases from the estimation sample. This way, and by focusing only on entry wages, we have only one observation for each employer–employee match. Besides, as in every match, someone either has a contact or not, there is no variation in T_{ijt} within the ij groups. These limitations induce that there will be no way to distinguish the match effects, μ_{ij}, from the idio-syncratic residual terms, ε_{ij}, and to identify the parameter on presence effects, θ, which could reflect lowered monitoring costs, knowledge transfer, or favoritism.

Therefore, we have to rely on a second-best estimator in which we cannot control for the match effects. To present the resulting implications, let us introduce the following matrix notation, based on Woodcock (2008), as an alternative for Eq. (1).

$$w = \theta T + \beta X + D\delta + F\gamma + G\mu + \varepsilon \tag{2}$$

In this form, w is the vector of wages, X is the matrix of observables, with T being the indicator for the presence of a co-worker link, and D, F, and G the design matrices of individual, firm, and match fixed effects, respectively. Without accounting for match effects, the two-way fixed effects estimator would be biased in the following way.

$$E[\theta_{TWFE}] = \theta + (T'M_{XDF}T)^{-1}T'M_{XDF}G\mu \tag{3}$$

The matrix M_{XDF} is a projection matrix, taking out the within-firm (F), within-individual (D), and observables-specific (X) variation from both the indicator (T) and match effects ($G\mu$).[8] Therefore, by omitting the match fixed effects and controlling only for separable and additive person and firm effects, the estimator will also incorporate the average difference of match effects among the two groups, controlled for firm and person effects and observables. $\hat{\theta}_{TWFE}$ would estimate θ without bias only if the match effects were, conditionally on X, F, and D, independent of the presence of contacts.[9]

---

8   The whole formula $(A'M_{abc}A)^{-1}A'M_{abc}B$ is the OLS estimator of A's effect on B, controlling for factors a, b and c. If A is a dummy variable, it reflects the conditional expectation of the difference in B between the two groups defined by A.

9   We note that the omission of match effects will lead to a biased estimation of both individual and firm effects. We discuss the implications later in the article.

While the independence of the idiosyncratic error term from the match effects seems to be a plausible assumption, the use of social contacts and match effects might be related. According to the literature, the selection into or creation of superior matches is one of the main mechanisms of referral activity. The second term in Eq. (3), $(T'M_{XDF}T)^{-1}T'M_{XDF}G\mu$, which, in the above setting, is an omitted variable bias, actually captures the magnitude of this selection. Therefore, by using two-way fixed effects regressions, we can only estimate the total of the match selection term and gains related to referrer presence, but we cannot separate them. In Section 5.3, however, we attempt to bypass this limitation.

3.2 **Individual and firm selections**

Besides the presence effects and the match selection induced by contacts, we are interested in the individual and firm sorting patterns related to co-worker networks as well. To pursue this goal, we rely on the following decomposition, based on Woodcock (2008), to compare the overall gain, θ_{OLS} with θ_{TWFE}.

$$E[\theta_{OLS}] = E[\theta_{TWFE}] + \underbrace{(T'M_XT)^{-1}T'M_XD\delta}_{\psi_{ind}} + \underbrace{(T'M_XT)^{-1}T'M_XF\gamma}_{\psi_{firm}} \tag{4}$$

This decomposition suggests that by not accounting for person and firm effects, we introduce two additional, distinct omitted variable biases. The first, ψ_{ind}, is the controlled difference between the unobserved skills among linked and nonlinked employees measured in (nominal) wage terms. That is, how much wage difference is implied by the linked employees' different latent qualities. A positive bias term suggests that good quality employees are more liable to be referred for jobs or more prone to applying to firms with their acquaintances present.

The bias created by omitting firm effects (ψ_{firm}) is the difference between the premium paid by firms where linked hires or referral activity occur and where they are not present, implicitly weighted by the number of new hires. A positive value suggests that linked employees can, on average, end up receiving higher wages as they can enter better quality firms, which pay higher (starting) wages for the same job relative to similar firms.

These average selection terms, however, do not capture whether the differences can be experienced within or between workers/firms. It is possible, for example, that while the linked workers of a firm are not especially high-wage ones, they are still better *relative* to the worker pool of the given firm. To account for the possibility that such patterns are present on the aggregate level as well, we further decompose the above-introduced selection terms.[10]

$$\underbrace{(T'M_XT)^{-1}T'M_XD\delta}_{\psi_{ind}} = \underbrace{(T'M_{XF}T)^{-1}T'M_{XF}D\delta}_{\xi_{ind}} + \underbrace{(T'M_XT)^{-1}T'M_XF\gamma^S}_{\omega_{ind}} \tag{5}$$

$$\underbrace{(T'M_XT)^{-1}T'M_XF\gamma}_{\psi_{firm}} = \underbrace{(T'M_{XF}T)^{-1}T'M_{XD}F\gamma}_{\xi_{firm}} + \underbrace{(T'M_XT)^{-1}T'M_XD\delta^S}_{\omega_{firm}} \tag{6}$$

In this decomposition, γ^S denotes the vector of *firm* effects obtained from a second-stage fixed effects regression on the estimated *individual* effects from the original two-way fixed effects wage equation. A firm that tends to hire individuals with high worker effects will have a high γ^S, regardless of the value of its firm effects. Similarly, δ_i^S reflects the

---

10   For the sake of brevity, let us assume, for now, that estimated individual and firm effects are estimated without bias.

average premium of firms a given individual ever works at. If there would be no systematic differences among firms or individuals in these parameters, as in case of the total absence of assortative matching, within and average differences in estimated effects would be the same due to the lack of correlation between individual and firm effects. Hence, this decomposition would be redundant.

Equation (5), therefore, shows that the *average* difference in the worker effects between linked and nonlinked hires is the sum of the average difference *within* firms (ξ_{ind}) where linked hires present and the difference in the average level of worker effects *between* firms with and without any linked hires (ω_{ind}). The first term could signal whether given firms benefit from accessing relatively better-skilled individuals through linked hires, while the second term describes how is the average worker pool of firms with linked hires compared to firms without such.

Similarly, ξ_{firm} will reflect whether firms, where hiring linked workers is prevalent, are better compared to the work history of the linked hires. That is, whether they benefit by moving to firms of their former colleagues. Finally, the parameter ω_{firm} will characterize the firms that are generally accessed by these workers even when they are hired without links. Table 1 briefly summarizes all the introduced parameters.

Table 1 Summary of parameters in our model

	Parameter		Interpretation
0.		$\hat{\theta}_{OLS}$	The wage differential between linked and nonlinked hires, controlling for only *observed* worker and firm characteristics.
1.		$\hat{\theta}_{TWFE}$	The wage differential between linked and nonlinked hires, controlling for *unobserved* firm and worker heterogeneity (but not match heterogeneity).
	1a.	θ	The pure 'presence effects' of having a potential referrer at the firm.
	1b.		Bias arising from the possibility that linked workers are better matched with firms (match selection).
2.		$\hat{\psi}_{ind}$	The average worker effect differential between linked and nonlinked hires.
	2a.	$\hat{\xi}_{ind}$	The average worker effect differentials between linked and nonlinked hires within firms.
	2b.	$\hat{\omega}_{ind}$	The average worker effect differential between firms that tend to make linked hires and those that tend to make nonlinked hires.
3.		$\hat{\psi}_{firm}$	The average firm effect differential between linked and nonlinked hires.
	3a.	$\hat{\xi}_{firm}$	The average firm effect differentials between linked and nonlinked hires within worker careers.
	3b.	$\hat{\omega}_{firm}$	The average firm effect differential between workers that tend to be hired with and without links.

Note: $\hat{\theta}_{TWFE}$ also contains the expected difference in error terms from Eq. (1), controlling for observables and person and firm effects. Our identifying assumption is that this term is zero. Also $\hat{\theta}_{OLS}$ could contain additional differences in the error terms due to misspecification or proxy issues that are only relevant if one does not control for two-way fixed effects. This is also assumed to be zero. This way $\hat{\theta}_{OLS} = \hat{\theta}_{TWFE} + \hat{\psi}_{ind} + \hat{\psi}_{firm}$. Also $\hat{\psi}_{ind} = \hat{\xi}_{ind} + \hat{\omega}_{ind}$ and $\hat{\psi}_{firm} = \hat{\xi}_{firm} + \hat{\omega}_{firm}$.

3.3 **Estimation of decompositions**

To get the parameters of the proposed decompositions, we estimate the following set of equations. First, we estimate the wage equation introduced in Eq. (1), but without match effects.[11]

$$w_{ijt} = \alpha + \theta_{TWFE}T_{ijt} + \beta_X X_{it} + \beta_Y Y_{jt} + \beta_Z Z_{ijt} + \delta_i + \gamma_j + \pi_t + \varepsilon_{ijt} \tag{7}$$

Then using the estimated person and firm effects $\hat{\delta}_i$ and $\hat{\gamma}_j$, we estimate the following equations to get the decompositions from Eqs (4–6).

$$\hat{\delta}_i = \alpha_2 + \psi_{ind}T_{ijt} + \beta_{2X} X_{it} + \beta_{2Y} Y_{jt} + \beta_{2Z} Z_{ijt} + \pi_{2t} + \varepsilon_{2ijt} \tag{8}$$

$$\hat{\gamma}_j = \alpha_3 + \psi_{firm}T_{ijt} + \beta_{3X} X_{it} + \beta_{3Y} Y_{jt} + \beta_{3Z} Z_{ijt} + \pi_{3t} + \varepsilon_{3ijt} \tag{9}$$

$$\hat{\delta}_i = \alpha_4 + \xi_{ind}T_{ijt} + \beta_{4X} X_{it} + \beta_{4Y} Y_{jt} + \beta_{4Z} Z_{ijt} + \gamma_j^S + \pi_{4t} + \varepsilon_{4ijt} \tag{10}$$

$$\hat{\gamma}_j = \alpha_5 + \xi_{firm}T_{ijt} + \beta_{5X} X_{it} + \beta_{5Y} Y_{jt} + \beta_{5Z} Z_{ijt} + \delta_i^S + \pi_{5t} + \varepsilon_{5ijt} \tag{11}$$

$$\hat{\gamma}_j^S = \alpha_6 + \omega_{ind}T_{ijt} + \beta_{6X} X_{it} + \beta_{6Y} Y_{jt} + \beta_{6Z} Z_{ijt} + \pi_{6t} + \varepsilon_{6ijt} \tag{12}$$

$$\hat{\delta}_i^S = \alpha_7 + \omega_{firm}T_{ijt} + \beta_{7X} X_{it} + \beta_{7Y} Y_{jt} + \beta_{7Z} Z_{ijt} + \pi_{7t} + \varepsilon_{7ijt} \tag{13}$$

We note that the omission of match effects may bias the estimated values of $\hat{\gamma}_j$ and $\hat{\delta}_i$. Firm effects will contain whether the firm makes good matches on average, and individual effects will contain if someone is prone to create good (or bad) matches. These bias terms are, however, independent of observables, including our proxy, T.[12] Thus, the controlled differences in fixed effects introduced above ($\hat{\psi}$, $\hat{\xi}$ and $\hat{\omega}$) are not affected by such biases.

Another concern could be the bias arising from identifying firm effects (and therefore person effects) only from a limited number of moves between establishments. As our panel is only a 50% sample (and we have to apply further restrictions to our sample), limited mobility bias (Andrews et al., 2008) could not be neglected. On the other hand, we can use six years of data and observe within-year movements as well, which may somewhat counterbalance the potential lack of identifying mobility. The most commonly discussed consequence of this bias is the overestimation of the variation in firm effects, and the underestimation of the correlation between firm effects and worker effects, a measure of assortative matching. While there are established methods for correcting the bias in these moments (Andrews et al., 2012; Bonhomme et al., 2020; Bonhomme et al., 2019; Gaure, 2014), we face a different problem.

According to Kline et al. (2020), limited mobility bias can also affect standard errors when someone projects the estimated firm (or person) effects of an AKM model on a set of observables, as we do in our decompositions with the proxy for social links. For instance, as we estimate biased firm effects with a higher variation, we will be seemingly able to explain this variation well with observable factors and get smaller standard errors and therefore biased inference in the second-stage estimations. In our example, for instance, we could

---

11 Models with multiple fixed effects are estimated based on the method of Correia (2017). Models with one or no fixed effects also use the Stata routine of Correia (2017), as it allows for two-way clustering of standard errors.

12 The estimated effects will capture conditional average differences in match effects in the following way: $E[\delta_{TWFE}] = \delta + (D'M_{TXF}D)^{-1}D'M_{TXF}G\mu$ and $E[\gamma_{TWFE}] = \gamma + (F'M_{TXD}F)^{-1}F'M_{TXD}G\mu$ (Woodcock, 2008).

overstate the role of the firm component in the overall wage difference. The authors propose a correction method for standard errors, which accounts for this possibility, and correct inference. However, we lack the computational infrastructure required for this exercise. Hence, standard errors in our decompositions may be somewhat underestimated, and measures of statistical significance are less reliable. Results should be treated accordingly, focusing on the relative magnitude of components of the decompositions rather than their statistical significance.[13]

3.4 Identification, proxy quality, and generalizability

The regression with two separable fixed effects, if estimated, yields a parameter, which measures the additional wage individuals could earn due to being hired with a link compared to the amount implied by their latent and observed qualities, the firm's wage setting-strategy, and other characteristics. This parameter is identified from both a comparison of employees at mixed firms and the comparisons of employment spells in the working history of individuals who were linked at least once.[14]

Based on the above, it is important to note that we cannot predict what would happen in those sectors where hiring through links is not prevalent or in population sub-samples where no such events are observed.[15] Therefore, the results may not be generalized to the whole population. However, this is not a problem as we are interested in the effects of co-worker connections where they are actually relevant. Also, it is important to note that the estimated individual and firm effects are comparable only within connected sets of workers and firms. As common in such datasets, we have a giant component in the paired graph of employers and employees, consisting of 92.7% of observations. We will estimate all models on this subset.

As the actual job-finding method is not observed in the data, another issue of our approach is the reliability of the proxy variable used. Namely, the proxy, T_{ijt}, may capture different variation depending on the controls. That is, the variation of T_{ijt} on average (OLS) or around a person's or a firm's mean (one-way fixed effects regressions) may not proxy the same phenomenon. Hence, while the variation of the proxy when using both firm and person fixed effects probably captures referral activity (Dustmann 2016), the selection terms let in other aspects from a broader set of phenomena. As we discussed previously, the sorting of high-wage workers to firms, or passing information about high-wage vacancies are aspects that we consider as part of the relevant mechanisms. However, some unintended variation may still be present in the proxy, so we have to interpret the selection terms with caution. For instance, in the case of hiring constantly from the same firm, we would systematically observe the arrival of linked workers and may falsely interpret these hires as referred ones, while wage gains may be related to the familiarity with the sending firm. We also have to account for the fact that workers

---

13 As a robustness check, we estimated an AKM model on a much larger set of data, which included mostly all spells in mostly all firms available, to acquire better estimations of firm and person effects. Then, instead of using conventional fixed effects methods (within transformation), we conditioned on these "pre-estimated" fixed effects in estimating the wage equation. The correlation between the pre-estimated firm effects and those from our main estimations were 0.84, while for individual effects, it was only 0.66. Parameters estimated this way were similar in magnitude; however, standard errors have increased for firm selection and decreased for individual selection terms.

14 More precisely, firms whose linked workers are always linked and whose nonlinked workers are never linked do not contribute to the parameter estimations. People who are linked in firms where everyone is linked and nonlinked at firms where no one is linked are also omitted from the comparisons.

15 Or in sectors where everyone is linked or with persons who are always linked.

getting into the same firm randomly is more common in sectors with high fluctuation and for people who switch workplaces often. If wages are high in these sectors or skilled persons tend to move a lot or have a limited number of options fit for their skills, we would face some unintended biases. While the two-way fixed effects regression controls for these issues, in the less-controlled regressions, we aim to avoid them by some sample restrictions and the inclusion of specific control variables.[16]

After discussing our main estimation results, we present additional evidence that may further suggest that the observed individual and firm selections are mostly driven by referral and information transmission-related mechanisms, instead of empirical artifacts or unintended variation in our proxy.

4 Data and co-workers

Our empirical analysis uses the Panel of Administrative Data from the Databank of the Centre for Economic and Regional Studies (formerly part of the Hungarian Academy of Sciences). It is a large administrative, linked employer–employee dataset, covering a random 50% of the working-age Hungarian population followed from January 2003 to December 2011. The dataset combines data from the official records of the Pension Directorate, the Tax Office, the Health Insurance Fund, the Office of Education, and the Public Employment Service. The raw register data were compiled and restructured by the Databank into a monthly level panel, in which all observations refer to the employment status of individuals on the 15th day of the given month.[17] For each observation belonging to an employment spell, the dataset has anonymous individual and employer identifiers, monthly earnings data, featuring the number of days in employment, information about employment type, occupation, and balance sheet data of the employer. Variables on health expenditures and social transfers received by the individuals are also available. Using the linked nature of the dataset, we could extract all those co-worker pairs who worked at the same company in any given month.

4.1 Co-worker definition and sample restrictions

By adding additional constraints, we selected those former colleague relationships that have the potential to serve as a basis for referral activity and/or information transmission. We defined former co-workers as those pairs of employees who had worked together at the same company, which had a maximum of 50 observed employees, before their reunion at another firm. Setting a limit on the company size of the first encounter was an essential and necessary step. Not using such a restriction would have led to the overestimation of the number of real social connections among former colleagues for two reasons. First, because at medium and large companies, not everyone knows each other. Second, among these companies having multiple sites is more typical, which would have further increased the probability of misclassification since the data

---

16 We are aware of one confounding factor that we cannot capture without person fixed effects: the personal preference for working with acquaintances. We can only assume that it is independent of average wage level or general skills; therefore, it will not lead to a higher wage among those who favor working in firms with social links.

17 While the source data contain all legal employment spells, which generate social contribution obligations, in the final structure, we do not observe employment spells that are shorter than 1 month and are not present on the 15th day of any month.

contains only firm-level, but not establishment-level information. To enhance the probability that co-workers actually knew each other, we applied further conditions. Co-worker pairs were considered valid only if they had worked together for at least 12 months in the past, they had reunited at a firm with a maximum of 250 observed employees[18], and the incumbent employees had arrived at least 1 month before their former co-workers did.[19] Also, as weaker social connections usually erode over time, we restricted the time that could pass between the two encounters to 5 years.

Our variable of interest then would be a proxy indicating whether upon entering a new firm, the entrant had at least one former co-worker who met the above criteria. Among those who had no such relations, we differentiated three groups. Regarding two of these, we cannot observe any link by definition: the first group consists of the first observed employment spell of each worker, while the second one comprises those workers who had worked only in larger firms (more than 50 observed employees). The remaining observations where former co-workers could be but are not present form the most comparable control group. While this latter is the one we will compare observations to, the former two groups are also included in the sample for the proper estimation of firm fixed effects.

In our estimations, we included only those 15–65-years-old private-sector employees who had no more than 15 distinct employment spells over 9 years and were not receiving social transfers. To avoid the confounding effects of social benefits on reservation wages, we focused only on job-to-job transitions and hires after unemployment spells no longer than 12 months. Artificial changes in firm identifiers, like those resulting from mergers, could have resulted in the overestimation of the referred employees' wage premium as we would see (re-)entries with high wages during someone's real employment spell. We removed from the data all identifiable cases of such artifacts. Observations, when more than three linked newcomers arrived together at a company from the same firm, were excluded, as comobility in itself can provide a substantial wage premium (Marx and Timmermans, 2017). We removed entries where the simple majority of the receiving firm's hires in the past year came from the same sending firm, which would potentially reflect the presence of a sending firm premium. Finally, all cases of workers returning to one of their former employers were omitted to avoid capturing the effects of firm-specific knowledge. Based on the process of defining peers, we note that in the early years of the observation period, there were artificially fewer former co-worker pairs than in later years (Figure A1). Hence, we used the first three years of data as the connection-forming period and only the later years (2006–2011) for estimations. As we focus on linked entries only in small and medium firms, we dropped entries from the nonlinked groups at firms with more than 250 observed employees as well. To get comparable estimates of firm effects, we used only the largest connected mobility group, which consists of 92.7% of the sample with the above restrictions.

Despite these restrictions, there is still a chance of misclassification. If employees do not get to know all of their co-workers within a year or if the former co-worker relationships erode in less than 5 years, employees may have been incorrectly labeled as linked ones. The reverse may also occur due to database-related issues since we could not identify former colleagues who

---

18   We restricted the possible size of the receiving firm, as our data do not comprise plant-level information. Also, in such large firms, there is a higher chance that contacts will be unaware of the application of their former co-workers.

19   This restriction only enforces that one worker arrived definitely earlier. By observing, instead of detailed employment spells, the registered workforce of the firms only on the 15th of each month this 1-month gap will reflect a 2–60-day difference between the starting dates of the two workers.

were not part of the 50% sample.[20] Furthermore, as opposed to our definition, some connections may form in large companies or may not erode even after 4 years. Either allocating high-wage, linked workers to the low-wage, nonlinked group, or vice versa results in a lower observable wage difference between the two groups. Therefore, both types of misclassification have the same effect on our estimations: the measured difference between the linked and nonlinked groups will be lower than their true values and the estimated effects will be biased toward zero.

4.2 **Definition of variables**

To estimate the parameters of the model defined in Eqs (7–13), we use the first months of employment spells as observation units. We define the independent wage variable, w_{ijt}, as the logarithm of daily earnings over the national average of daily earnings to standardize over time. We prefer to use starting wages as they are determined by different processes than, for instance, wages in subsequent years in a working spell. When defining starting wages, employers usually cannot rely on actually observed performance of the workers. Hence, a referrer's contribution might be essential in the assessment of hiring risks. Inside information about the firm could also alter the initial wage expectations of new applicants. In the subsequent months of employment, wages can be adjusted according to employers' experiences with newcomers' performance.

Individual controls consist of the interaction of (quadratic) age and imputed education[21], residence[22], and gender, with the latter two only included in regressions without individual fixed effects. We also control for previous work experience with the number of former workplaces in an elastic form, as subsequent employment spells have an increasing probability to be linked even in the absence of referral. We included the indicator of work experience in the two-digit occupation category of the new job. Time-variant firm-specific characteristics include ownership (foreign or domestic private) and a two-digit industry code. To control for any possible remaining time trends, we include year dummies. To get the effect of unemployment on reservation wages, we include dummies for the length of the unemployment spell, measured in months, preceding entering the firm.

We also include a full set of controls interacted with an indicator of the observation being the first observed employment spell of an individual and another dummy indicating that the individual could not obtain proper co-worker ties due to the lack of experience at small firms. This allows us to estimate firm fixed effects properly and also to have a larger connected mobility group (Glitz and Vejlin, 2019).

Our main variable of interest is the dummy indicating the presence of any former co-worker, interacted with gender categories to capture heterogeneous effects. In regressions without individual fixed effects, we include the set of the gender-occupation category dummies, where occupation can take on five categories: manager, skilled white-collar, unskilled white-collar, skilled blue-collar, and unskilled blue-collar.

---

20 With not being able to observe 50% of the population, we lose 75% of all possible connections and, based on simulations, around 66% of the linked observations. In our sample, around 10% of the nonlinked hires would actually be linked, given this sampling issue.

21 Unfortunately, we can observe the actual level of education only in special cases. Therefore, in our estimations we include an approximate measure of education which is defined based on the occupation in the individual's overall work history which demands the highest level of education.

22 The database contains only details about individuals' residence in 2003. However, supplementary investigations show that changing residence is not common in the sample, as it affects only 5% of individuals.

4.3 **Baseline differences**

We define the control group, to which linked hires could be reliably compared, as those non-linked workers who previously worked at a small firm. The groups of workers in their first employment spells and those who previously only worked at large companies were also distinguished. Table 2 contains the mean values and distributions of some key variables in the sample by these observation groups.

When comparing raw means of outcomes, we can observe a significant wage advantage of linked hires over the control group. In nominal monthly earnings, the difference is more than

Table 2 Summary statistics: job entrants, freshly acquired jobs, and receiving firms

	Subsample		Nonlinked subgroups			Control group	
	Linked	Nonlinked	First spell	W/o small firm experience	Control group	Always nonlinked	Nonlinked and linked
No. of observations	20,227	944,579	135,818	147,600	661,161	645,253	15,908
Log of relative daily earnings	−0.470	−0.580	−0.745	−0.468	−0.571	−0.572	−0.552
Monthly earnings (HUF)	128,511	108,053	80,897	127,616	109,264	109,245	110,046
Age	38.0	32.7	25.6	32.3	34.2	34.2	36.2
Elementary education	12%	11%	21%	14%	9%	9%	13%
Secondary education	63%	66%	63%	65%	67%	67%	64%
Tertiary education	25%	23%	16%	22%	25%	25%	24%
Central Hungary	32%	34%	34%	29%	35%	35%	34%
Central Transdanubia	12%	12%	10%	15%	12%	12%	13%
Western Transdanubia	9%	10%	8%	12%	10%	10%	9%
Southern Transdanubia	8%	7%	6%	8%	8%	8%	8%
Northern Hungary	11%	9%	8%	10%	9%	9%	11%
Northern Great Plain	13%	11%	11%	13%	11%	11%	12%
Southern Great Plain	12%	11%	10%	10%	11%	11%	11%
Max. number of workplaces	5.39	5.57	2.58	4.77	6.37	6.35	7.27
Occupation specific experience	70%	47%	–	37%	59%	59%	69%
Length of prev. unemployment	1.5	2.2	–	2.4	2.1	2.2	1.6
Manager	7%	3%	2%	4%	4%	4%	5%
White-collar work	6%	6%	6%	8%	6%	6%	5%
Other white-collar work	16%	19%	24%	20%	18%	18%	14%
Skilled blue-collar work	43%	40%	38%	34%	41%	41%	44%
Unskilled blue-collar work	28%	31%	30%	35%	30%	30%	33%
Relative wage level of firm	0.907	0.877	0.816	1.028	0.856	0.857	0.815
Sector: Agriculture	3%	2%	3%	2%	2%	2%	2%
Sector: Industry	41%	34%	32%	38%	34%	34%	36%
Sector: Trade and services	56%	64%	65%	60%	64%	64%	62%
Foreign firm	18%	20%	20%	27%	19%	19%	16%
Domestic firm	82%	80%	80%	73%	81%	81%	84%
Number of employees	39.2	42.9	44.4	60.0	38.8	39.0	30.8

Note: The estimation sample consists of starting months of worker employment spells, between 2006 and 2011, which follow a maximum 12-month long job-search period. It includes those 15–65 years old, private sector employees, who had less than 15 distinct employment spells in the observation period and did not receive social transfers. The table comprises the average wage outcomes of individuals upon entry to a new firm, demonstrates the personal traits of workers, and contains the characteristics of the workers' new jobs and firms. Indented figures reflect statistically significant differences ($p < 0.05$) from the linked group, according to t-tests.

17%. However, when we use a more fine measure, in which we normalize by the number of days worked and the national average wage, we see only a 0.1 log point difference, suggesting a 10% wage advantage of linked hires over market ones. It is also worth to note that the wage level of firms the linked group works at is 6% higher.

However, the mean difference in wages might only reflect differences in regional, occupational, or sectoral composition. While the distribution of these observable characteristics is similar in the two groups, in our estimations, we control for them. Differences in a few specific factors especially have to be accounted for, as they may be structurally connected to how links are generated in the data. For instance, if social contacts would have no effect on job search, we would expect that people who change jobs more often, are older, or work at larger firms have a higher chance of ending up in the same firm as a former co-worker. We observe significant age differences, as linked workers are on average 4 years older. However, we find that there is no difference in firm size and actually linked hires are the ones who have fewer employers in the observation period. This latter may suggest another beneficial effect of links, longer expected tenure. Finally, this descriptive comparison also suggests that contacts reduce the average length of job search by around a half month.[23]

Regarding the other nonlinked groups, we see that those who spend their first working spell in the estimation period are typically younger, earn less compared to linked hires, a higher share of them works in trade and services, and a lower share is working in the industrial sector. Workers without small firm experience on average earn approximately the same amount as linked workers while having somewhat fewer employment spells in the period and being on average younger than linked workers. Their wage advantage compared to the other nonlinked groups might originate from working at large firms who, especially multinational employers, pay significantly higher wages in Hungary than their domestic counterparts (Köllő et al., 2020).

5 Results

5.1 Main results

To understand the wage gains related to co-worker networks, we start by estimating the model described in Eq. (7), and then we decompose the gains according to Eqs (8–13). Additionally, we calculate a pooled OLS panel regression (Eq. (7) without any fixed effects) as well. The main results are presented in Tables 3 and 4, of which the former shows the results for the estimations in which the variable of interest is interacted with gender.

While the descriptive statistics (Table 2) demonstrate that there is a significant difference in raw earnings between linked and nonlinked entrants, the OLS results indicate that even after controlling for observable characteristics, the difference between the two groups is still present.[24] We can observe a 4.65% wage gain for linked male workers and 3.13% for linked

---

23 Workers in the identifying sample for person effects, that is, those who have variation in the proxy for contact presence, make up almost 2% of the estimation sample. On average, they earn 5% more, are 3 years older, and work at 1.5 more workplaces than workers in our estimation sample. These differences mostly come from the requirement of observing more hiring events for these workers.

24 We ran the OLS specification on the sample used for the TWFE estimates to assess whether sample distortions could account for differences in parameters. We found reasonably similar OLS parameters. The parameter on males turned out to be 0.051 ($t = 7.7$), and for female workers, we observed a small decrease to 0.028 ($t = 2.4$). It seems that sample differences account for only limited part of the differences between the OLS and other models, which control for unobserved heterogeneity.

Table 3　Decomposition of co-worker gains by gender

	$\hat{\theta}_{OLS}$	$\hat{\theta}_{TWFE}$	$\hat{\psi}_{ind}$	$\hat{\psi}_{firm}$	$\hat{\xi}_{ind}$	$\hat{\xi}_{firm}$	$\hat{\omega}_{ind}$	$\hat{\omega}_{firm}$
Male	0.0465***	0.0213***	0.0167***	0.0086*	0.0125***	0.0118*	0.0041	−0.0032
	(0.0055)	(0.0051)	(0.0038)	(0.0041)	(0.0034)	(0.0049)	(0.0023)	(0.0034)
Female	0.0313***	−0.0024	0.0254***	0.0083	0.0265***	0.0148	−0.0010	−0.0065
	(0.0082)	(0.0096)	(0.0063)	(0.0064)	(0.0055)	(0.0080)	(0.0038)	(0.0057)
N	964,806	501,200	964,806	964,806	943,643	571,441	964,806	964,806
N_i	616,386	197,435	616,386	616,386	616,365	223,021	616,386	616,386
N_j	105,818	61,121	105,818	105,818	84,655	105,778	105,818	105,818
R^2	0.327	0.860	0.204	0.200	0.453	0.612	0.052	0.087

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry (Eq. (7)), without any and with two-way fixed effects, and the consecutive decomposition regressions on estimated firm and individual effects (Eqs 8-13)), respectively. The selection parameters in the columns reflect overall, within unit and between unit differences in individual and firm effects, respectively. Our variable of interest, the proxy for links, is interacted with two gender categories. Additional controls (if the corresponding fixed effects are not included) consist of gender, quadratic age interacted with imputed education, residence, the number of workplaces and job search length in an elastic form, five levels of occupation, two-digit industry codes, firm ownership, a dummy for occupation-specific experience, and dummies for calendar years. All controls are interacted with the indicators for first employment spells and for non-linked workers without small firm experience. These observations contribute only to the proper estimation of firm effects. Standard errors are in parentheses and clustered at both firm level and individual level. *Statistically significant at 0.05 level; **at 0.01 level; and ***at 0.001 level.

Table 4　Decomposition of co-worker gains by occupations—male results

	$\hat{\theta}_{OLS}$	$\hat{\theta}_{TWFE}$	$\hat{\psi}_{ind}$	$\hat{\psi}_{firm}$	$\hat{\xi}_{ind}$	$\hat{\xi}_{firm}$	$\hat{\omega}_{ind}$	$\hat{\omega}_{firm}$
Manager	−0.0988***	−0.0030	−0.0775***	−0.0183	−0.0699***	0.0226	−0.0076	−0.0409**
	(0.0259)	(0.0310)	(0.0211)	(0.0142)	(0.0196)	(0.0241)	(0.0092)	(0.0127)
Skilled$_W$	0.0924***	0.0551*	−0.0006	0.0378*	−0.0049	0.0094	0.0044	0.0285
	(0.0274)	(0.0235)	(0.0187)	(0.0179)	(0.0169)	(0.0250)	(0.0101)	(0.0156)
Unskilled$_W$	0.0627***	0.0409*	0.0167	0.0051	0.0153	−0.0113	0.0013	0.0164
	(0.0182)	(0.0183)	(0.0134)	(0.0129)	(0.0120)	(0.0172)	(0.0076)	(0.0108)
Skilled$_B$	0.0584***	0.0228**	0.0217***	0.0140*	0.0128**	0.0123	0.0089**	0.0016
	(0.0070)	(0.0077)	(0.0050)	(0.0055)	(0.0044)	(0.0065)	(0.0034)	(0.0047)
Unskilled$_B$	0.0475***	0.0118	0.0340***	0.0017	0.0326***	0.0161	0.0014	−0.0144*
	(0.0081)	(0.0075)	(0.0048)	(0.0069)	(0.0045)	(0.0085)	(0.0030)	(0.0056)
N	964,806	501,200	964,806	964,806	943,643	571,442	964,806	964,806
N_i	616,386	197,435	616,386	616,386	616,365	223,021	616,386	616,386
N_j	105,818	61,121	105,818	105,818	84,655	105,778	105,818	105,818
R^2	0.327	0.860	0.190	0.200	0.443	0.612	0.052	0.086

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry (Eq. (7)), without any and with two-way fixed effects, and the consecutive decomposition regressions on estimated firm and individual effects (Eqs (8–13)), respectively. The selection parameters in the columns reflect overall, within unit and between unit differences in individual and firm effects, respectively. Our variable of interest, the proxy for links, is interacted with 10 categories based on gender and five occupational categories: managers, skilled white-collar, unskilled white-collar, skilled blue-collar, and unskilled blue-collar workers. Only the parameters for male workers are presented. For the list of additional controls, see Table 3. Standard errors are in parentheses and clustered at both firm level and individual level. *Statistically significant at 0.05 level; **at 0.01 level; ***at 0.001 level.

female workers compared to their nonlinked counterparts. This gross premium is, however, composed of various elements.

By estimating the two-way fixed effects model from Eq. (7), we get the wage premium which is attributable to either match selection or referrer-dependent explanations. The $\hat{\theta}_{TWFE}$ parameter is only significant for male workers. Among them, those who have co-worker links upon their arrival at a new workplace earn 2.13% more compared to similar workers, even considering the workers' employment history and other hires of the same firm. As established in Section 3.1, due to the lack of variability of the proxy within worker-firm pairs, the above two elements are empirically indistinguishable using the present methodology and data. Therefore, we cannot tell whether this gain is driven by selection into better matches or effects related to the presence of a referrer and the rent sharing of the firm. However, we know that for male workers, the sum of the two results in a significantly positive wage advantage. The magnitude of this estimation is in line with the literature, especially with Dustmann et al. (2016), who measured a 3.3% gain in a model with two-way fixed effects and direct information on referral.

We use the first decomposition to account for the average selection of high-wage individuals and high-wage firms into linked hire events. Based on the parameter $\hat{\psi}_{ind}$, linked male workers earn 1.67% more than nonlinked workers due to their higher individual effects. Accordingly, more than one-third of the overall wage gains originates in linked workers having better unobservable qualities. As a result of the decreased screening costs, due to direct or indirect signaling, the firm may be able to hire better quality workers whose skills would be appreciated by other firms as well in terms of higher wages. Moreover, approximately one-sixth of the male wage difference (0.86%) is explained by the higher premium of firms linked individuals work at when they are linked.[25] This may suggest a certain level of information transmission through the co-worker network or employees obtaining access to better-quality firms that would not be accessible to them in the absence of their connections. For women, this channel, and the parameter $\hat{\psi}_{firm}$ is of the same absolute magnitude, although it is not statistically significant. The most dominant element of their overall wage difference is the individual selection term.

Although $\hat{\psi}_{ind}$ and $\hat{\psi}_{firm}$ provide some insight into the average difference between linked and nonlinked workers and employers in unobservable wage components, we are interested in how the latent qualities of linked hires compare to their peers or competing firms. To achieve this, we further decompose the average differences in worker and firm effects into within ($\hat{\xi}$) and between ($\hat{\omega}$) unit components.

The $\hat{\omega}_{firm}$ parameter shows that those male individuals who ever become linked have somewhat ordinary firm pools. They typically work at firms that provide average or slightly below-average wages. However, if these workers start their new job at companies where they have links, they can easily get into higher premium firms compared to their own work history as the positive parameter $\hat{\xi}_{firm}$ suggests. Concerning linked women, even though they can get into better premium firms compared to their employment histories, on average, this gain is dampened by the fact that they usually work in inferior establishments, resulting in a nonsignificant overall difference.

Parameter $\hat{\omega}_{ind}$ demonstrates that linked male workers are typically admitted to companies where the worker pool is on the average slightly better than in similar firms without links.

25　Due to limited mobility bias, this parameter might not be significant. In our robustness check, using pre-estimated firm effects, the standard error of $\hat{\psi}_{firm}$ was somewhat higher.

However, even compared to this slightly better pool, they are still better in terms of their unobserved qualities. As $\hat{\xi}_{ind}$ suggests, there is a 1.25% advantage in starting wages, attributable to higher person effects of linked male workers compared to the firms' other employees. Similarly, for women, only this within term is dominant, with the between-firm difference being very close to zero. These results are comparable with findings by Hensvik and Skans (2016) and Glitz and Vejlin (2019), who relied on controlling for firm fixed effects, hence capturing the total of presence effects, match selection, and the within-firm selection of individuals. They found 3.6% and 4.6% wage gains, respectively.[26]

All things considered, it seems that both male workers and employers profit from co-worker networks. Workers can get into high-wage firms (both on average and in relative terms) through their contacts' information, while firms can find and select better-quality workers (averagely and compared to their own workforce) through relying on referrers. In addition, the creation of better matches and/or the referrer-related gains might benefit both parties. Regarding female workers, the only relevant channels are the selection of better workers into firms, and some, weak sorting into better firms relative to the working history of these women.

Next, we investigated the effect of links in interaction with gender and occupation. Table 4 comprises the parameters for male workers.[27] Ignoring, for the moment, the managerial category, we observe that both the OLS and the two-way fixed effects parameters are smaller in less prestigious occupations. For the unskilled blue-collar workers, the $\hat{\theta}_{TWFE}$ parameter is not even significant.[28] Regarding this latter group, individual selection is the most relevant: the differences in worker effects, coming mostly from the within term, account for 72% of the observed average gap. This channel is also important for skilled blue-collar workers, and no other groups, where individual differences (both within and between firms) contribute to almost half of the difference between OLS and two-way fixed effects results. The results suggest that match or presence-related gains are high in occupations where firm-specific or job-specific knowledge is more essential, and, therefore, the match-specific component is a more important determinant of wages. Accordingly, in less demanding categories, we observe selection with respect to general skills and productivity of workers ($\hat{\psi}_{ind}$), which is presumably more important in these occupations.

Selection into higher premium firms seems to be a dominant factor only in the two skilled occupational categories that demand specific qualifications. For skilled blue-collar jobs, the within element of firm selection is dominant, while for skilled white-collar positions, the better firm pool of linked workers drives the results. It also looks like that in these skilled occupations, linked workers get into firms with generally high-wage worker pools. Compared to these pools, skilled blue-collar workers can be somewhat better, while skilled white-collar workers are slightly worse. Finally, managers who get into firms where their former co-workers (mostly subordinates) work are usually employed in firms with lower wages. However, relative to their worse firm pool, they still get into better firms when they are hired with links, but initially earn less than other nonlinked managers. Added together, these elements result in a lower expected wage for linked managers.[29]

---

26 These parameters should be compared to the sum of $\hat{\theta}_{TWFE}$ and $\hat{\xi}_{ind}$, the overall within person gain in our model, which is around 3.38%.

27 Parameters for the female occupation categories coming from the same regression are in App. Table A1, while App. Table A2 presents the model with only occupation categories not differentiated by gender.

28 We note that we lose a lot of statistical power when we work with these categories, as the identification of the parameters rely on within-firm and within-person comparisons of workers of a given occupation-gender category only.

29 The identifying sample for this parameter is quite specific and small as firms need to hire both linked and nonlinked managers.

The patterns we observed are consistent with the predictions about how employee referral and information transmission should affect the different wage components of linked workers. The observed strong match-specific wage differences (especially for more specialized occupations) and individual selection of better workers (more in general occupations) suggest a strong role of the signaling power of employee referral. On the other hand, selection into higher-wage firms, even if weak, suggests a better opportunity pool provided by contacts through information transmission. In Section 5.3, we aim to reinforce this interpretation through alternative specifications focusing on scenarios where one or more of the mechanisms are expected to exert stronger effects.

5.2 **Exogenous job mobility**

A concern that scholars often face in this literature is that employee movements are most often endogenous, especially job-to-job transitions. Papers focusing on re-employment outcomes through contacts naturally focus on exogenous job loss (e.g., plant closures, mass layoffs), while the ones about wages typically do not make this restriction as (multiple) fixed effects are ought to take care of selection issues. However, as we interpret the selection terms as well, it is worth assessing whether the selection patterns we document may be different when switching jobs is just an option for workers and when they have to find work due to job loss. To do so, we labeled cases where more than one-third of a firm's workforce left within a 3-month long period as exogenous job losses.[30] Then, we interacted our original proxy variable with mobility type (and gender). The results are presented in Table 5.

The parameters are fairly similar to the ones we have seen before, although the relative importance of some patterns changed. The overall gains of linked male workers are even higher

Table 5 Endogenous and exogenous job mobility

	$\hat{\theta}_{OLS}$	$\hat{\theta}_{TWFE}$	$\hat{\psi}_{ind}$	$\hat{\psi}_{firm}$	$\hat{\xi}_{ind}$	$\hat{\xi}_{firm}$	$\hat{\omega}_{ind}$	$\hat{\omega}_{firm}$
Endog.	0.0436***	0.0162**	0.0188***	0.0086*	0.0161***	0.0071	0.0027	0.0015
	(0.0058)	(0.0055)	(0.0041)	(0.0044)	(0.0037)	(0.0054)	(0.0025)	(0.0037)
Exog.	0.0620***	0.0423***	0.0103	0.0095	−0.0003	0.0366***	0.0106**	−0.0272***
	(0.0119)	(0.0117)	(0.0080)	(0.0087)	(0.0072)	(0.0107)	(0.0041)	(0.0073)
N	964,806	501,200	964,806	964,806	943,643	571,442	964,806	964,806
N_i	616,386	197,435	616,386	616,386	616,365	223,021	616,386	616,386
N_j	105,818	61,121	105,818	105,818	84,655	105,778	105,818	105,818
R^2	0.327	0.860	0.203	0.200	0.453	0.612	0.052	0.087

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry (Eq. (7)), without any and with two-way fixed effects, and the consecutive decomposition regressions on estimated firm and individual effects (Eqs (8-13)), respectively. The selection parameters in the columns reflect overall, within unit and between unit differences in individual and firm effects, respectively. Our variable of interest, the proxy for links, is interacted with four categories based on gender and whether the hire was preceded by an exogenous job loss event (*Exog.*). Only the parameters for male workers are presented. For the list of additional controls, see Table 3. Standard errors are in parentheses and clustered at both firm level and individual level. *Statistically significant at 0.05 level; **at 0.01 level; ***at 0.001 level.

---

30 We applied this definition only to firms with at least 15 observed employees, and cases when the majority laid off workers did not appear again under the same firm identifier.

after exogenous job losses compared to conventional movements, with the main difference coming from a substantial and significant increase in $\hat{\theta}_{TWFE}$. This may suggest that referring someone after a job loss happens either when referrers are willing to take more responsibility (e.g., in voluntary monitoring) or when better signals can be provided. Signals, however, seem to be match-specific, instead of those of general skills, as the individual selection term is rather small, with its within component being virtually zero. The creation of better matches is consistent with the finding of Eliason et al. (2017), who show that companies often create new positions to acquire good workers experiencing layoffs. The composite effect $\hat{\psi}_{firm}$ is driven by linked workers getting into higher-wage companies compared to their averagely lower-wage firm pool. The strong within component could suggest the importance of information transmission. However, the inferior pool of the linked workers is puzzling. The overall wage gain of linked workers, nevertheless, may mitigate the long-term disadvantages of displaced individuals (Eliason and Storrie 2006).

5.3 **Supplementary specifications**

In this section, we aim to provide further suggestive evidence that reinforces our claim that the wage gains we observed are mostly driven by information transmission and/or referral activity—as opposed to, for instance, some empirical artifacts. To do so, we focus on scenarios where one or more of the (sub-)mechanisms are anticipated to exert stronger effects on wages, for instance, when referrers have larger bargaining power at their employer and expect to observe an increase in the corresponding wage gain components. First, we focus on such cases where referral-related gains should be larger, but information transmission is not necessarily more prevalent. Then, we present two exercises aimed at distinguishing between the presumably small referral-related presence effects and gains originating in match selection. Finally, we focus on job entries where information transmission in itself could be a dominant factor in generating high wage opportunities.

First, we are interested in whether the relative position of the former co-worker in the entry firm affects the estimated wage effects. We differentiate three broad levels of occupations: managers, occupations with either vocational or general higher education requirement and those without such prerequisites. We then refine the proxy from the main estimations and create three new ones, showing whether a former colleague is present at the firm in a more demanding, a similar, or a lower requirement occupation. We expect that better peers, that is managers for everyone and skilled positions for unskilled entrants, will have larger bargaining power at the firm and hence may have a larger effect on referral-related wage gains upon entry. Inferior peers may not be able to recommend the applicants at all and moving to places with such contacts are more probably random reunions.[31] Information flows, on the other hand, may be actually less common between different occupational levels.

Table 6 comprises the results of the regression, in which we used the alternative proxies.[32] If the occupation of the links is similar compared to the job entrants, we find gains of a similar magnitude as in our main estimations. Firm selection, especially compared to the

---

31 At the same time, we do not expect homophily in worker quality to be a stronger factor in the superior or inferior cases.

32 In the upcoming estimations, as before, we interact our key variables with gender, but report the parameters for only male workers.

Table 6 Heterogeneity of co-worker gains by relative position of contact

	$\hat{\theta}_{OLS}$	$\hat{\theta}_{TWFE}$	$\hat{\psi}_{ind}$	$\hat{\psi}_{firm}$	$\hat{\xi}_{ind}$	$\hat{\xi}_{firm}$	$\hat{\omega}_{ind}$	$\hat{\omega}_{firm}$
Superior	0.0572***	0.0210*	0.0331***	0.0030	0.0444***	−0.0073	−0.0113*	0.0104
	(0.0110)	(0.0106)	(0.0075)	(0.0084)	(0.0070)	(0.0110)	(0.0051)	(0.0069)
Similar	0.0463***	0.0163**	0.0167***	0.0133**	0.0096**	0.0177**	0.0071**	−0.0044
	(0.0058)	(0.0058)	(0.0040)	(0.0044)	(0.0036)	(0.0055)	(0.0024)	(0.0037)
Inferior	−0.0125	−0.0002	−0.0177	0.0054	−0.0142	0.0115	−0.0035	−0.0061
	(0.0127)	(0.0130)	(0.0100)	(0.0083)	(0.0088)	(0.0109)	(0.0054)	(0.0074)
N	964,806	501,200	964,806	964,806	943,643	571,441	964,806	964,806
N_i	616,386	197,435	616,386	616,386	616,365	223,021	616,386	616,386
N_j	105,818	61,121	105,818	105,818	84,655	105,778	105,818	105,818
R^2	0.327	0.860	0.204	0.200	0.453	0.612	0.052	0.087

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry (Eq. (7)), without any and with two-way fixed effects, and the consecutive decomposition regressions on estimated firm and individual effects (Eqs (8-13)), respectively. The selection parameters in the columns reflect overall, within unit and between unit differences in individual and firm effects, respectively. Our variables of interests reflect the presence of contacts in occupational positions that are superior, similar, or inferior compared to the job entrant's occupation in terms of skill requirements. The indicators are interacted with gender, the table presents the coefficients for male workers. For the list of additional controls, see Table 3. Standard errors are in parentheses and clustered at both firm level and individual level.
*Statistically significant at 0.05 level; **at 0.01 level; ***at 0.001 level.

entrants' work history, is somewhat stronger, suggesting that relevant information could be passed about vacancies that the incumbent worker has experience with. This channel seems negligible for superior peers, as $\hat{\psi}_{firm}$ suggests. However, the individual selection parameter, $\hat{\psi}_{ind}$, is twice as large in the latter scenario than in the baseline case, with the point estimate of $\hat{\theta}_{TWFE}$ being roughly similar. This may reflect the fact that higher-position peers may provide better quality, more reliable signals about the match-specific and general productivity of applicants, enhancing the corresponding aspects of the selection. The effect of inferior peers is insignificant regarding all wage components, being mostly near zero or slightly negative.

Next, we check whether the tenure of contacts can also affect the wage gains of newcomers similarly to the (relative) occupation of the links at the firm. It seems reasonable that as the working experience of the potential referrers increases, they will establish more trust and bargaining power, so they can generate more reliable signals about the productivity of newcomers. Therefore, it is more likely that they can meaningfully affect the hiring probabilities and wages of the applicants. We also investigate heterogeneity by tie-specific characteristics as well, such as the length of the common working spell and the time that has elapsed between the two encounters of the worker pair. We assume that while a longer co-working spell could enhance the creation of stronger links, the elapsed time between the co-working spells might weaken those links. Therefore, changes in these features can strengthen or moderate the probability of referral and information transmission and might affect the observable wage gains. To estimate the effect of the introduced features, we interacted them with the referral proxy and included these interactions in the same regression.[33]

In line with our expectations, both the tenure of the links and the length of the common working experience enhance the individual and the firm selections, although we see no effect

---

33 We demeaned the three characteristics by their sample means in order to estimate their slopes in their usual range.

Table 7 Heterogeneity of co-worker gains by link and tie characteristics

	$\hat{\theta}_{OLS}$	$\hat{\theta}_{TWFE}$	$\hat{\psi}_{ind}$	$\hat{\psi}_{firm}$	$\hat{\xi}_{ind}$	$\hat{\xi}_{firm}$	$\hat{\omega}_{ind}$	$\hat{\omega}_{firm}$
Linked	0.0382***	0.0202***	0.0173***	0.0008	0.0142***	−0.0002	0.0031	0.0009
	(0.0054)	(0.0054)	(0.0038)	(0.0041)	(0.0034)	(0.0052)	(0.0023)	(0.0034)
Seniority	0.0020***	−0.0005	0.0012***	0.0013***	0.0013***	0.0009	−0.0001	0.0004
	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0003)	(0.0005)	(0.0002)	(0.0003)
Since	−0.0003	−0.0001	−0.0000	−0.0001	−0.0003	0.0001	0.0002	−0.0002
	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0004)	(0.0002)	(0.0002)
Common	0.0013***	0.0000	0.0006*	0.0007**	0.0005*	0.0005	0.0001	0.0003
	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0002)	(0.0004)	(0.0001)	(0.0002)
N	964,806	501,200	964,806	964,806	943,643	571,441	964,806	964,806
N_i	616,386	197,435	616,386	616,386	616,365	223,021	616,386	616,386
N_j	105,818	61,121	105,818	105,818	84,655	105,778	105,818	105,818
R^2	0.345	0.860	0.203	0.232	0.453	0.628	0.069	0.059

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry (Eq. (7)), without any and with two-way fixed effects, and the consecutive decomposition regressions on estimated firm and individual effects (Eqs (8-13)), respectively. The selection parameters in the columns reflect overall, within unit and between unit differences in individual and firm effects, respectively. Our variable of interest is the proxy for links, which is interacted with both gender and contact or tie-related characteristics. *Seniority* refers to the tenure of the links with the longest working spell at the entry firm. Variable *Since* indicates the time elapsed since the latest common working spell with the link(s). *Common* denotes the length of the longest common co-working spell in the past. The coefficients show the effect of positive deviation in months of all three characteristics from their mean value among linked male workers. For the list of additional controls, see Table 3. Standard errors are in parentheses and clustered at both firm level and individual level.
*Statistically significant at 0.05 level; **at 0.01 level; ***at 0.001 level.

on the $\hat{\theta}_{TWFE}$ parameter (see Table 7). This implies more intense information transmission to both applicants and firms, but only about general qualities. It also seems that the age of the tie is not a relevant factor regarding the presence of such selections. Nevertheless, the fact that some of the gains are larger when social links tend to be stronger suggests that the selection terms are driven by the participation of peers.

In an attempt to capture the relevance of referral gains that depend on the continuous presence of the referrer, we leverage that peers may leave the firm earlier than their new, referred colleagues. Although we believe that expected voluntary monitoring of the peer and knowledge sharing are already evaluated in starting wages, the separation of the referrer will probably weaken the bargaining position of the worker in the firm due to the loss of productivity-enhancing features, reducing the wage advantage in the long run. Even if this does not lead to a decrease in wages, it may impede further wage increase and dissolve the advantage of referred workers over market hires. We estimate and plot the two-way fixed effects wage gains over the first three years at the firm for those who at the given time still have their former referrer at the firm and those whose peers have left by the time.[34]

We observe that referral gains, similarly as documented in Dustmann et al. (2016), disappear over time as actual productivity of all workers gets revealed, and workers of inferior quality leave the firm (Figure 1). However, there is a modest, although statistically insignificant difference in the point estimates of the gain-tenure path depending on the presence of the

---

34 We match on the pre-estimated individual and firm effects from the equations to enforce comparing similar individuals and firms, while also maintaining the feasibility of the estimation.

Figure 1 Wage-tenure profiles and referrer presence.
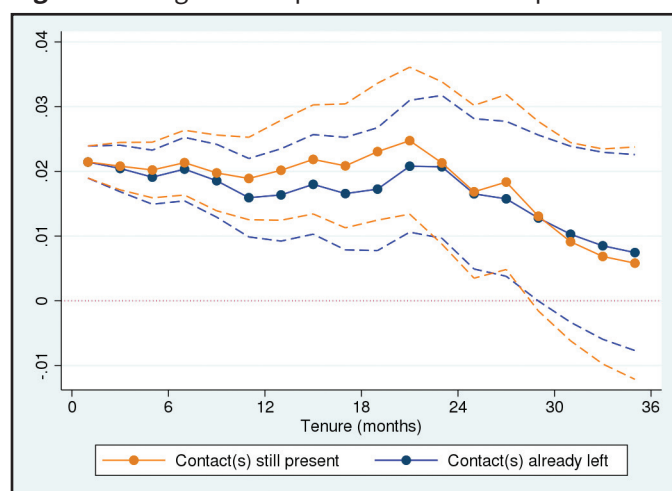
Note: The figure displays point estimations and 95% confidence intervals for two sets of estimations. The first set uses the proxy for contact presence upon entry, while in the other one the same indicator is set to zero if the original contact(s) left the firm by the given month. Both graphs present the parameter of eighteen separate regressions on the logarithm of daily earnings of male workers in the given month, for all odd months of the first three year of the employment spells. Female workers are included in the estimations, with a constant gender difference being assumed. Controlling for fixed effects is achieved by including pre-estimated individual and firm effects from the entry month equation. Additional controls are the same as listed in Table 3. Standard errors used for the confidence intervals are clustered at both firm level and individual level.

original contact(s). For those workers who do not have their peers present anymore, gains start to dissolve earlier, but even this difference disappears over time.

As an additional endeavour to separate the gains related to referrer presence from the match-specific ones, we interacted occupation-specific skill variables with the proxy of links. We assumed that regarding certain occupations, the role of monitoring, knowledge sharing and various off-cv elements will be more valuable. Therefore, larger referral gains could be observed in occupations where such related skills are dominant. For instance, knowledge sharing may have a larger role and be more valuated by the employer in jobs requiring more independence. We obtained various skill and ability measures from the O*NET 24.2 Database by the U.S. Department of Labor, Employment and Training Administration.[35] However, we could not find any skill requirement that would significantly alter the $\hat{\theta}_{TWFE}$ parameter in those occupations, which demand certain unobserved skills (see Table A3 in Appendix). This may suggest that the gains we would like to measure are rather modest or cannot be effectively captured by occupation-related skills. The signs of the parameters, however, have a pattern similar to what we observed in the specification with occupational categories (Table 4). The interaction terms are positive for job traits that reflect the need for specific knowledge (like innovation or analytical thinking), and negative for those skills which can be considered more generally applicable (like stamina and stress tolerance). While these exercises are not conclusive, they suggest match selection as the main driver of $\hat{\theta}_{TWFE}$.

---

35 Although the database is based on US occupation surveys, the scores could provide some insights for Hungary as well. Among others, Handel (2012) confirmed that US and European survey-based occupation measures typically lead to comparable results. Using Hungarian job descriptions made by experts ÉLETPÁLYA (2020) yielded similar results.

In our final exercise, in contrast with the previously introduced cases, we look at specific scenarios where the presence of actual recommendation is unlikely. To do so, we incorporate two additional indicators in the general model from Eq. (7). The first one indicates the presence of those nonlinked individuals who have at least one former firm in common with the applicant but did not share a common working spell together at that firm, hence did not have the chance to make actual personal contact. The other indicator marks the presence of second links in the co-worker network. These individuals are former co-workers of the applicants' previous peers. For this dummy, we considered only those second links who did not share a former, common firm with the applicant.[36] While information transmission about vacancies across this network is rather reasonable, actual recommendation is unlikely due to the lack of these links' personal experience and knowledge about the applicant. We expect to observe negligible recommendation-related gains from the presence of both second links and of those whose firm histories overlap—but not their employment spells.[37]

The results presented in Table 8 are only partially in line with our expectations. Concerning those who got workers at their new firms with similar working histories, we cannot observe a significant $\hat{\theta}_{TWFE}$, which is reassuring, as this parameter is ought to capture mostly referral-related wage gains. However, we see individual selection which is almost as strong as the one in our baseline case. This is somewhat unexpected, but not unreasonable. The similarity in working history might function as an indirect signal for the productivity of the entrant worker, as the employers

Table 8 Gains from links, second links, and similar workers

	$\hat{\theta}_{OLS}$	$\hat{\theta}_{TWFE}$	$\hat{\psi}_{ind}$	$\hat{\psi}_{firm}$	$\hat{\xi}_{ind}$	$\hat{\xi}_{firm}$	$\hat{\omega}_{ind}$	$\hat{\omega}_{firm}$
Linked	0.0484***	0.0194***	0.0188***	0.0102*	0.0152***	0.0141**	0.0036	−0.0039
	(0.0056)	(0.0055)	(0.0038)	(0.0042)	(0.0035)	(0.0052)	(0.0023)	(0.0035)
Similar	0.0131**	0.0072	0.0150***	−0.0091*	0.0194***	0.0013	−0.0044*	−0.0104**
	(0.0050)	(0.0050)	(0.0035)	(0.0041)	(0.0032)	(0.0047)	(0.0023)	(0.0034)
Second	0.0489**	0.0125	0.0062	0.0302**	−0.0074	0.0238	0.0136*	0.0064
	(0.0157)	(0.0173)	(0.0104)	(0.0111)	(0.0098)	(0.0142)	(0.0056)	(0.0091)
N	938,791	479,919	938,791	938,791	917,835	550,362	938,791	938,791
N_i	603,975	189,756	603,975	603,975	603,955	215,546	603,975	603,975
N_j	105,061	60,367	105,061	105,061	84,105	105,022	105,061	105,061
R^2	0.326	0.860	0.199	0.198	0.452	0.611	0.050	0.088

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry (Eq. (7)), without any and with two-way fixed effects, and the consecutive decomposition regressions on estimated firm and individual effects (Eqs (8-13)), respectively. The selection parameters in the columns reflect overall, within unit and between unit differences in individual and firm effects, respectively. Our parameters of interest are estimated with distinct indicators for the presence of former co-worker links, workers with similar working histories (those who share a common, former workplace with applicants), and second links (the former peers of the job-entrants' former co-workers who did not fall into the similar working history group). The indicators marked in the table as *Linked*, *Similar* and *Second* respectively and were interacted with gender. Results for male workers are presented. For the list of additional controls, see Table 3. Standard errors are in parentheses and clustered at both firm level and individual level. *Statistically significant at 0.05 level; **at 0.01 level; ***at 0.001 level.

36　Workers whom one shared a common workplace with, but at a different time tend to mechanically become second links.

37　As we cannot see all contacts (due to having a 50% sample), we cannot make sure that there are no first-order contacts at the new workplace. This will lead to one-sided misclassification between the groups with first links and only second links, attenuating the difference between the two set of estimated parameters, as effects estimated for second-order contacts would be contaminated by the effect of unobserved first contacts.

might assume homophily in terms of skills between those workers who have similar working histories. A more puzzling finding is the presence of a significant negative firm selection, which as $\hat{\omega}_{firm}$ suggests, can be attributed to the fact that these individuals typically work at low-paying firms. Regarding those individuals who have only second links upon entry, we observe more consistent patterns. As expected, we see no recommendation-related individual or match selections. On the other hand, a rather strong selection into high-wage firms is associated with these weak ties. This might suggest that there is indeed actual information transmission about high-paying jobs through the extended networks of co-workers.

The introduced specifications aimed to provide additional evidence that our parameters are driven by nonrandom sorting of workers and capture the effects of information transmission and referral mechanisms. When we utilized scenarios that would theoretically imply the increase of referral-related gains (such as the better position of peers at the applicants' new firm) or the dominance of information transmission-related gains (e.g., the presence of second links), our results followed the patterns we anticipated. However, we failed to infer a conclusion about the relative importance of gains strictly dependent on the presence of the referrer versus match selections already present at hiring. This could be the focus of future research.

6 Discussion

Taken together, our findings suggest that the reliance on links is beneficial for both firms and workers. Regardless of whether it is driven by referral or just information transmission, the use of contacts induces the selection of better workers into firms and selection of workers into better firms. What we deem important to highlight is the fact that these aggregate selections predominantly happen within units. That is, on the one hand, people get into superior firms compared to their working history. This way these mechanisms might contribute to the individuals' upward mobility. On the other hand, firms can enhance the quality of their worker pools through referral as referred hires are generally better workers compared to the firm's own average worker pool. In addition to these one-sided advantages, the effect on the average match quality is beneficial for both parties. By increasing the overall productivity in the labor market, referral can be socially desirable.

Nevertheless, the effect on individuals who cannot rely on social links should be considered as well. If workers with worse career prospects also have inferior co-worker networks, their initial disadvantages will be magnified by being crowded out from high-paying firms. Being trapped in inferior workplaces may hinder the development of network quality, reinforcing the path dependence in career paths. Referral may also lead to the increase of sorting inequality if it helps allocating the best workers to the best firms as shown by Eliason (2019). While the direct assessment of assortativity was beyond the scope of this study, the between terms of our detailed decomposition suggest a weak sorting pattern: firms relying on referral generally employ slightly better than average quality workers, while on average they themselves are high-paying firms. Thus, the presence of productivity gains from the generation of better matches could be counterbalanced by the crowding-out effect of disadvantaged workers and the effect on sorting inequality, resulting in unclear implications about overall welfare.

References
Abowd, J. M.; F., Kramarz; D. N. Margolis (1999): High Wage Workers and High Wage Firms. Econometrica 67(2), 251-333. doi:10.1111/1468-0262.00020

Andrews, M. J.; L. Gill; T. Schank; R. Upward (2008): High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias? Journal of the Royal Statistical Society 171(3), 673-697. doi:10.1111/j.1467-985X.2007.00533.x

Andrews, M. J.; L. Gill; T. Schank; R. Upward (2012): High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias. Economics Letters 117(3), 824-827. doi:10.1016/j.econlet.2012.07.033

Bartus, Tamás (2001): Social capital and earnings inequalities: The role of informal job search in Hungary. [S.l.]: s.n., 176 p.

Bian, Y.; X. Huang; L. Zhang (2015): Information and Favoritism: The Network Effect on Wage Income in China. Social Networks 40, 129-138. doi:10.1016/j.socnet.2014.09.003

Bonhomme, S.; K. Holzheu; T. Lamadon; E. Manresa; M. Mogstad; B. Setzler (2020): How Much Should We Trust Estimates of Firm Effects and Worker Sorting? https://drive.google.com/_le/d/1vzRRps-nOyAp2HIXRfYjVmldc8jfOdKr/view

Bonhomme, S.; T. Lamadon; E. Manresa (2019): A Distributional Framework for Matched Employer Employee Data. Econometrica 87(3), 699-739. doi:10.3982/ECTA15722

Calvo-Armengol, A.; M. O. Jackson (2004): The Effects of Social Networks on Employment and Inequality. American Economic Review 94(3), 426-454. doi:10.1257/0002828041464542

Calvo-Armengol, A.; M. O. Jackson (2007): Networks in Labor Markets: Wage and Employment Dynamics and Inequality. Journal of Economic Theory 132(1), 27-46. doi:10.1016/j.jet.2005.07.007

Castilla, E. J. (2005): Social Networks and Employee Performance in A Call Center. American Journal of Sociology 110(5), 1243-1283. doi:10.1086/427319

Cingano, F.; A. Rosolia (2012): People I know: Job Search and Social Networks. Journal of Labor Economics 30(2), 291-332. doi:10.1086/663357

Corcoran, M.; L. Datcher; G. J. Duncan (1980): Information and Influence Networks in Labor Markets, in: Duncan, G. J.; J. N. Morgan (eds.), Five Thousand American Families - Patterns of Economic Progress. Ann Arbor, MI: Institute of Social Research, 1-37.

Correia, S. (2017): Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator (tech. rep.) [Working Paper]. http://scorreia.com/research/hdfe.pdf

Dhillon, A.; V. Iversen; G. Torsvik (2015): Employee Referral, Social Proximity and Worker Discipline: Theory and Evidence from India (tech.rep.) [CMI Working Paper]. https://www.cmi.no/publications/5386-employee-referral

Dustmann, C., Glitz, A., Schönberg, U., Brücker, H. (2016): Referral-Based Job Search Networks. The Review of Economic Studies 83(2), 514-546. doi:10.1093/restud/rdv045

Ekinci, E. (2016): Employee referrals as a screening device. The RAND Journal of Economics 47(3), 688-708. doi:10.1111/1756-2171.12141

ÉLETPÁLYA. (2020): http://eletpalya.munka.hu. Accessed 4 April 2020.

Eliason, M.; L. Hensvik; F. Kramarz; O. N. Skans (2017): The Causal Impact of Social Connections on Firm's Outcomes. IFAU Working Paper, 2017(11).

Eliason, M.; L. Hensvik; F. Kramarz; O. N. Skans (2019): Social Connections and the Sorting of Workers to Firms. IZA Discussion Papers, Institute for the Study of Labor (IZA), No. 12323.

Eliason, M.; D. Storrie (2006): Lasting or Latent Scars? Swedish Evidence on the Long-Term Effects of Job Displacement. Journal of Labor Economics 24(4), 831-856. doi:10.1086/506487

Fernandez, R. M.; E. J. Castilla; P. Moore (2000): Social Capital at Work: Networks and Employment at A Phone Center. American Journal of Sociology 105(5), 1288-1356. doi:10.1086/210432

Gaure, S. (2014): Correlation Bias Correction in Two-Way fixed effects linear regression. Stat 3(1), 379-390. doi:10.1002/sta4.68

Glitz, A. (2017): Coworker Networks in the Labour Market. Labour Economics 44(1), 218-230. doi:10.1016/j.labeco.2016.12.006

Glitz, A.; R. M. Vejlin (2019): Learning through Coworker Referrals. IZA Discussion Papers, Institute for the Study of Labor (IZA), No. 12270.

Granovetter, M. S. (1995): Getting A Job: A STUDY of CONTACTS and Careers (2nd). The University of Chicago Press.

Handel, M. J. (2012): Trends in Job Skill Demands in OECD Countries. OECD Social, Employment and Migration Working Papers, No. 143. doi:10.1787/5k8zk8pcq6td-en

Hensvik, L.; O. N. Skans (2016): Social Networks, Employee Selection and Labor Market Outcomes. Journal of Labor Economics 34(4), 825-867. doi:10.1086/686253

Holzer, H. J. (1988): Search Method Use by Unemployed Youth. Journal of Labor Economics 6(1), 1-20.

Kline, P.; R. Saggio; M. Sølvsten (2020): Leave-out estimation of variance components. https://eml.berkeley.edu/_pkline/papers/KSS2020.pdf

Köllő, J.; I. Boza; L. Balázsi (2020): Wage Gains from Foreign Ownership: Evidence from Linked Employer-Employee Data. IZA Discussion Papers, Institute for the Study of Labor (IZA), No. 13040.

Kugler, A. (2003): Employee referrals and efficiency wages. Labour Economics 10(5), 531-556. doi:10. 1016/S0927-5371(03)00047-2

Marx, M.; B. Timmermans (2017): Hiring Molecules, Not Atoms: Comobility and Wages. Organization Science 28(6), 965-1167. doi:10.1287/orsc.2017.1155

Montgomery, J. D. (1991): Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. American Economic Review 81(5), 1407-1418.

Munshi, K. (2003): Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. Quarterly Journal of Economics 118(2), 549-599. doi:10.1162/003355303321675455

Saloner, G. (1985): Old Boy Networks as Screening Mechanisms. Journal of Labor Economics 3(3), 255-267. doi:10.1086/298055

Saygin, P.; A. Weber; M. A. Weynandt, (2019): Coworkers, Networks, and Job-Search Outcomes among Displaced Workers. Industrial and Labor Relations Review (ILR Review). doi:10.1177/0019793919881988

Schmutte, I. M. (2015): Job Referral Networks and the Determination of Earnings in Local Labor Markets. Journal of Labor Economics 33(1), 1-32. doi:10.1086/677389

Smith, S. S. (2005): "Don't put my name on it": Social capital Activation and Job-Finding Assistance among the Black Urban Poor. American Journal of Sociology 111(1), 1-57. doi:10.1086/428814

Tate, D. S. (1994): Restructuring Agency Job Descriptions Using Realistic Job Previews. Administration and Policy in Mental Health 22(2), 169-173. doi:10.1007/BF02106550

Ullman, J. C. (1966): Employee Referrals: Prime Tools for Recruiting Worker. Personnel 43, 30-35.

Wanous, J. P. (1980): Organizational Entry: Recruitment, Selection, and Socialization of Newcomers. Addison-Wesley (Reading, MA).

Woodcock, S. D. (2008): Wage Differentials in the Presence of Unobserved Worker, Firm, and Match Heterogeneity. Labour Economics 15(4), 771-793. doi:10.1016/j.labeco.2007.06.003

Appendix

Figure A1　Number of linked workers over time.
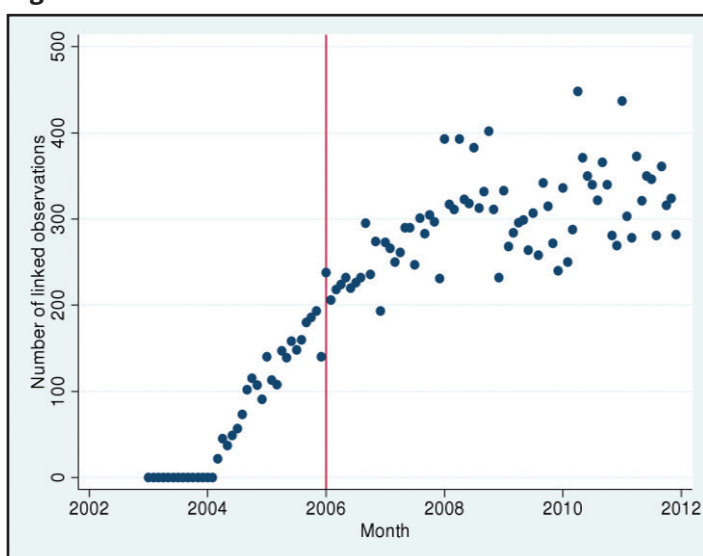
Note: The figure displays the number of hires with co-worker links present in each calendar month from 2003 until 2011. The subsample used for the estimation is the same as in Table 2.

Table A1　Decomposition of co-worker gains by occupations - female results

	$\hat{\theta}_{OLS}$	$\hat{\theta}_{TWFE}$	$\hat{\psi}_{ind}$	$\hat{\psi}_{firm}$	$\hat{\xi}_{ind}$	$\hat{\xi}_{firm}$	$\hat{\omega}_{ind}$	$\hat{\omega}_{firm}$
Manager	−0.0415	−0.0809	0.0168	0.0226	0.0372	0.0804**	−0.0204	−0.0579**
	(0.0416)	(0.0463)	(0.0345)	(0.0216)	(0.0320)	(0.0289)	(0.0141)	(0.0198)
Skilled$_W$	0.1648***	0.0784	0.0157	0.0706**	0.0015	0.0208	0.0142	0.0498*
	(0.0425)	(0.0403)	(0.0293)	(0.0253)	(0.0264)	(0.0352)	(0.0152)	(0.0227)
Unskilled$_W$	0.0503**	0.0081	0.0296*	0.0126	0.0357**	0.0150	−0.0061	−0.0024
	(0.0158)	(0.0194)	(0.0128)	(0.0113)	(0.0115)	(0.0157)	(0.0074)	(0.0098)
Skilled$_B$	0.0275*	−0.0107	0.0343***	0.0039	0.0298***	0.0170	0.0044	−0.0132
	(0.0111)	(0.0140)	(0.0090)	(0.0100)	(0.0074)	(0.0124)	(0.0062)	(0.0089)
Unskilled$_B$	−0.0097	−0.0116	0.0158	−0.0138	0.0178*	−0.0115	−0.0020	−0.0023
	(0.0140)	(0.0190)	(0.0102)	(0.0125)	(0.0090)	(0.0158)	(0.0057)	(0.0109)
N	964,807	501,200	964,807	964,807	943,643	571,443	964,807	964,807
N_i	616,386	197,435	616,386	616,386	616,365	223,021	616,386	616,386
N_j	105,818	61,121	105,818	105,818	84,655	105,778	105,818	105,818
R^2	0.327	0.860	0.190	0.200	0.443	0.612	0.052	0.086

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry (Eq. (7)), without any and with two-way fixed effects, and the consecutive decomposition regressions on estimated firm and individual effects (Eqs (8-13)), respectively. The selection parameters in the columns reflect overall, within unit and between unit differences in individual and firm effects, respectively. Our variable of interest, the proxy for links, is interacted with ten categories based on gender and five occupational categories: managers, skilled white-collar, unskilled white-collar, skilled blue-collar, and unskilled blue-collar workers. Only the parameters for female workers are presented. For the list of additional controls, see Table 3. Standard errors are in parentheses and clustered at both firm level and individual level.
*Statistically significant at the 0.05 level; **at the 0.01 level; ***at the 0.001 level.

Table A2　Decomposition of co-worker gains by occupations

	$\hat{\theta}_{OLS}$	$\hat{\theta}_{TWFE}$	$\hat{\psi}_{Ind}$	$\hat{\psi}_{Firm}$	$\hat{\xi}_{Ind}$	$\hat{\xi}_{Firm}$	$\hat{\omega}_{Ind}$	$\hat{\omega}_{Firm}$
Manager	−0.0849***	−0.0262	−0.0508**	−0.0080	−0.0397*	0.0378	−0.0111	−0.0458***
	(0.0223)	(0.0259)	(0.0183)	(0.0121)	(0.0171)	(0.0196)	(0.0079)	(0.0109)
Skilled$_W$	0.1146***	0.0620**	0.0046	0.0480**	−0.0026	0.0130	0.0072	0.0349*
	(0.0240)	(0.0202)	(0.0158)	(0.0156)	(0.0141)	(0.0205)	(0.0088)	(0.0136)
Unskilled$_W$	0.0558***	0.0244	0.0223*	0.0091	0.0249**	0.0022	−0.0026	0.0069
	(0.0126)	(0.0137)	(0.0096)	(0.0089)	(0.0086)	(0.0116)	(0.0056)	(0.0076)
Skilled$_B$	0.0507***	0.0159*	0.0236***	0.0112*	0.0156***	0.0132*	0.0079**	−0.0020
	(0.0060)	(0.0068)	(0.0044)	(0.0049)	(0.0038)	(0.0058)	(0.0030)	(0.0043)
Unskilled$_B$	0.0349***	0.0079	0.0287***	−0.0017	0.0278***	0.0117	0.0009	−0.0134**
	(0.0072)	(0.0070)	(0.0044)	(0.0063)	(0.0041)	(0.0076)	(0.0028)	(0.0051)
N	964,807	501,200	964,807	964,807	943,643	571,443	964,807	964,807
N$_i$	616,386	197,435	616,386	616,386	616,365	223,022	616,386	616,386
N$_j$	105,819	61,121	105,819	105,819	84,655	105,779	105,819	105,819
R^2	0.327	0.860	0.204	0.200	0.453	0.612	0.052	0.087

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry (Eq. (7)), without any and with two-way fixed effects, and the consecutive decomposition regressions on estimated firm and individual effects (Eqs (8-13)), respectively. The selection parameters in the columns reflect overall, within unit and between unit differences in individual and firm effects, respectively. Our variable of interest, the proxy for links, is interacted with five occupational categories: managers, skilled white-collar, unskilled white-collar, skilled blue-collar, and unskilled blue-collar workers. For the list of additional controls, see Table 3. Standard errors are in parentheses and clustered at both firm level and individual level.
*Statistically significant at the 0.05 level; **at the 0.01 level; ***at the 0.001 level.

Table A3　Co-worker gains and skill requirements

	Linked	**Skill**	**Interaction**
Baseline	0.0172*** (0.0046)	−	−
Manual Dexterity	0.0172*** (0.0047)	−0.0449*** (0.0020)	−0.0026 (0.0052)
Stamina	0.0172*** (0.0047)	−0.0445*** (0.0019)	−0.0066 (0.0055)
Persistence	0.0167*** (0.0046)	0.0441*** (0.0014)	0.0037 (0.0052)
Stress Tolerance	0.0174*** (0.0046)	0.0308*** (0.0014)	−0.0024 (0.0050)
Analytical Thinking	0.0164*** (0.0046)	0.0469*** (0.0015)	0.0053 (0.0050)
Complex Problem Solving	0.0159*** (0.0046)	0.0546*** (0.0016)	0.0056 (0.0048)
Active Learning	0.0167*** (0.0046)	0.0528*** (0.0015)	0.0016 (0.0052)
Coordination	0.0174*** (0.0046)	0.0398*** (0.0014)	−0.0025 (0.0044
Cooperation	0.0171*** (0.0046)	0.0203*** (0.0015)	−0.0022 (0.0052)
Adaptability/Flexibility	0.0170*** (0.0046)	0.0329*** (0.0015)	0.0026 (0.0050)
Originality	0.0163*** (0.0046)	0.0436*** (0.0015)	0.0066 (0.0052)
Innovation	0.0158*** (0.0046)	0.0329*** (0.0014)	0.0081 (0.0048)
Independence	0.0172*** (0.0046)	0.0152*** (0.0015)	0.0020 (0.0049)

Note: Estimation results from the main regression on the logarithm of daily earnings upon job entry with two-way fixed effects (Eq. (7)). Our variable of interest, the proxy for links, is interacted with the demeaned values of skill requirement measures from the O*Net database. For the list of additional controls, see Table 3. Standard errors are in parentheses and clustered at both firm level and individual level. All regressions are based on 483 418 observations and have an *R^2* between 0.860 and 0.861.
*Statistically significant at the 0.05 level; **at the 0.01 level; ***at the 0.001 level.