IFAU
Institute for Evaluation of Labour Market and Education Policy

The strength of the weakest link: sickness absence, internal substitutability and worker-firm matching

Lena Hensvik
Olof Rosenqvist

WORKING PAPER 2015:28

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

The strength of the weakest link: sickness absence, internal substitutability and worker-firm matching[a]

by

Lena Hensvik[b] and Olof Rosenqvist[c]

November 23, 2015

Abstract

We study how employee sickness absence affects worker-firm matching. We build on the idea that firms are sensitive to absence in jobs with few substitutes (*unique positions*). Consistent with this, we show that unique employees are less absent conditional on individual characteristics, establishment fixed effects and detailed occupational information. Half of this association is explained by *sorting* of low-absence workers into unique positions but sorting is less pronounced under imperfect information. Finally, job separations respond more to realized sickness absence in unique positions. The findings suggest that the cost of production disruptions is an important aspect of firms' hiring choices.

Keywords: Sickness absence, production disruption, coworker substitutes, hiring strategies
JEL-codes: J23, J24, L23, M51

[a] We thank Kevin Lang, Uta Schönberg, Karl-Oskar Lindgren, Illoong Kwon, Oskar Nordström Skans, Marcus Eliason and audiences at IFAU, UCLS, CReAM and EEA 2015 in Mannheim for helpful comments and suggestions.
[b] Institute for Evaluation of Labour Market and Education Policy (IFAU), lena.hensvik@ifau.uu.se
[c] Uppsala University, Economics Department and Institute for Evaluation of Labour Market and Education Policy (IFAU), olof.rosenqvist@nek.uu.se

Table of contents

1 Introduction

Recent reports in many countries document high sickness absence rates and large associated costs for firms.[1] Yet, little is known about how sickness absence affects key labor market outcomes, such as access to jobs, worker mobility and career trajectories. In addition, we know next to nothing about which strategies firms use to minimize the costs of employee absence. In this paper we examine whether firms use workers' sickness absence history as a sorting criterion when hiring. More specifically, we test the idea that firms seek workers with a low sickness absence behavior in jobs where absence leads to particularly costly production disruptions.

The idea that firms should be eager to find the right employees for the right jobs is motivated by the fact that both workers and jobs are heterogeneous, which can lead to match-specific gains in productivity.[2] Despite the theoretical notion of match-specificity there is still little empirical evidence on cross-firm differences in hiring and the importance of worker-firm complementarities. One reason is that it is difficult to measure how well a worker matches a particular job, which has forced researchers to infer match effects based on how wages and separations vary with tenure and job mobility (Nagypál, 2007, Lazear and Oyer, 2012).[3]

In addition, most of the discussion about match quality is focused on complementarities in terms of worker skills (or human capital) and the skill requirements (or technology) of different jobs.[4] But it is well possible that there may be important complementarities in other dimensions of employee attributes and firm technology that determine how workers select into jobs and, in turn, their labor market outcomes.[5] In this paper we expand the notion of match-quality, by focusing on complementarities between workers' sickness absence behavior and job attributes.

[1] It is reported that 131 million working days were lost due to sickness absence in the UK in 2013 (Office for National Statistics [UK], 2014). Another report from the UK estimates that employers pay GBP 9 billion a year in sick pay and associated costs (Black and Frost, 2011). In Germany it is reported that employers spend about EUR 25 million per year on sick pay. This number is more than 1 percent of the total GDP in Germany (German Federal Statistical Office, 2011). Numbers for Sweden suggest that employers spent SEK 21 billion on sick pay and associated costs in 2012 (Previa, 2013).

[2] See Sattinger (1975), and Tinbergen (1956) for the original work on the problem of assigning heterogeneous workers to heterogeneous jobs.

[3] Two exceptions are Jackson (2013) who shows that teacher-school match effects explain a quarter of the variation in teacher quality and Fredriksson et al. (2015) who show that wages and job separations depend on how well worker's cognitive abilities and personality traits match the abilities of the existing workforce.

[4] See, for example, Abowd et al. (2007) on how different components of skills are related to firms' technological inputs; Andersson (2009) on the relation between firms' product market segment and the demand for worker innovation skills in the software industry or Lazear (2009) on firm-level heterogeneity in skill-weights.

[5] For example, Lazear (1998) argues that the match-quality for a given firm depends on the riskiness of workers, and firm-level characteristics such as expected time-horizon and the degree of private information.

We use a simple search and matching model to illustrate the idea that firms avoid high-absence workers in jobs with few substitutable workers, as the costs associated with production disruptions from work absence should be decreasing in the number of employees performing the same tasks (as captured by occupational classifications).[6] The extent to which worker-job sorting occurs via hiring choices naturally depends on the ability of firms to differentiate high absence types from low absence types ex ante. Because employee sickness absence is more difficult to observe for employers compared to, for example, formal credentials it is not clear that it will have a significant impact on actual hiring outcomes.

Empirically, we start by documenting a very strong relationship between employee absence and the number of coworker substitutes in the private sector, conditional on detailed occupation and establishment fixed effects. This finding is robust to a number of variations of the empirical model (controlling for individual characteristics, changing the level of detail of the occupational codes and including public sector employees). The difference in sickness absence between workers with few/several substitutes is economically non-trivial, and is approximately similar in magnitude to the difference in absence rates between labor market entrants and workers in the middle of their career.

Importantly, about half of the association remains when we account for worker heterogeneity by including worker fixed effects. This suggests that the relationship between sickness absence and internal worker substitutability is partly due to assignment of low sickness absence types into unique positions and partly due to endogenous adjustments of sickness absence when employees are assigned to jobs with low internal substitutability. We corroborate the first mechanism by showing that workers hired for unique positions have lower *pre-hire* sickness absence compared to workers hired for jobs with higher internal substitutability. In addition, we show evidence of systematic job turnover among recently hired employees who become absent within jobs with few substitutes. Conditional on staying in the job, realizations of absence is instead related to the probability of receiving assignment more coworker substitutes.

[6] Conceptually, consider a workplace with complementarities between two occupational groups where one group is small and one is large, e.g. a newspaper with many journalists and one IT-expert. Intuitively the employer should care more about keeping sickness absence low in the IT-expert position given the potential production disruption that absence in that position could create. This notion is related to the O-ring production function described by Kremer (1993) where the marginal value of improving the weakest link in the production is very high.

Overall, these findings suggest that firms manage to reduce absence in sensitive positions by both selective hiring and ex post adjustments. The fact that job separations respond to realized sickness absence in addition suggests that employment relationships are formed under uncertainty as suggested in the seminal work of Jovanovic (1979b). We further explore the role of information available in the hiring stage using three different (but related) proxies for the amount of information in the hiring stage derived from the employees' employment history: strong pre-hire employment record, previous firm connection and common workplace history with incumbent employees.

Our findings suggest that the level of information about worker sickness absence type is an important determinant of the sorting patterns. First, we show that less precise information (using the proxies mentioned above) generally is associated with less pronounced sorting. Second, information is more strongly related to pre-hire sickness absence in jobs with few substitutes than in jobs with relatively many substitutes, suggesting that screening for low-absence workers is primarily important for key (unique) positions. Third, we find that the separation response due to realized sickness absence among workers in unique positions is negatively related to the amount of information in the hiring stage.

The paper contributes to the literature in several ways. First and foremost, employee selection and hiring strategies is still somewhat of a black box (Oyer and Schaefer, 2011). Limited evidence suggests that employers are reluctant to hire applicants with a history of sickness absence, but remain uninformative of why (Eriksson et al., 2012). The strong association between employee absence and the number of substitutes suggest that the cost of disruptions associated with sickness absence is a key aspect of the hiring decision, as the direct costs of absence (e.g. co-payments) are the same irrespectively of the number of substitutes. In that sense, high-sickness absence workers constitute *weak links* in jobs with few substitues.

Our findings also relate to the literature documenting a positive association between sickness absence rates and firm size in the cross-section, which is consistent with the argument that production in small firms should be particularly sensitive to individual sickness absence.[7] However, it is possible that this relationship also reflects other

[7] See Allen (1981), Barmby and Stephan (2000), Dionne and Dostie (2007) and Ose (2005). Lindgren (2012) also provides an interesting contribution showing that the higher sickness absence rates in larger firms primarily are driven by a higher incidence of sickness spells rather than longer durations.

between-firm differences related to size. By exploiting variation in the number of substitutes within narrowly defined job cells, the present paper provides a more credible assessment of the direct relationship between sickness absence and employee substitutability.

We also provide some evidence on the role of information in the hiring decision. Here, some recent studies show that firms rely on signals or informal search channels in order to screen for the right workers.[8] Our findings suggest that both pre-hire screening and post-hire separations induced by realizations of absence types serve as means to achieve an allocation of low-absence workers in unique positions within firms.

Our findings could, finally, also have implications for other issues related to worker sorting, such as the reasons behind the gender gaps in career outcomes. It is well-documented that women are underrepresented in key positions (such as managerial positions) in most countries. At the same time, women are absent from work more frequently than men (in particular after childbearing). In the light of our findings it would thus be interesting to address whether there is a direct relationship between the difference in real and perceived absenteeism across genders and the job opportunities of men and women in key positions.[9] [10]

The remainder of the paper is structured as follows. In Section 2 we explain the theoretical framework that we rely on. Section 3 describes the data and clarifies crucial definitions. The empirical specifications and the results are presented in Section 4. Section 5 concludes.

2 Theoretical framework

The core idea of this paper is that firms are particularly sensitive to absence in jobs that have a smaller number of substitutable workers (i.e. unique positions). This hypothesis builds on the assumption that the costs associated with insuring against production disruption from sickness absence should be decreasing in the number of employees

[8] Empirical studies in this literature suggest that employers use observable signals such as education (Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lange, 2007; and Schönberg, 2007), unemployment status (Eriksson and Rooth, 2014), and referral ability (Hensvik and Skans, forthcoming) to form expectations about prospective workers productivity.

[9] Note that this selection can be both demand- and supply-driven.

[10] A similar argument is made by Ichino and Moretti (2009) who attribute part of the gender earnings gap to gender gaps in absenteeism induced by the menstrual cycle. They do not however speak specifically about women's access to key positions.

performing a specific task.[11] Essentially, we argue that the combination of a job with few substitutes and a worker with a high proneness to sickness absence is a bad match in terms of generated payoff. We use a search and matching framework developed by Eeckhout and Kircher (2011) to formalize this idea and its implications for worker-job sorting patterns. The model should primarily be seen as a pedagogical tool to illustrate our main hypotheses.

2.1 The model

Assume that we have two worker types ($x^H$and x^L) where H (L) stands for high (low) proneness to sickness absence and two types of jobs, one with high (H) and one with low (L) internal substitutability.[12] The payoff for a given match is $f(x, y)$. Jobs are sensitive to sickness absence if there are few substitutable workers. Thus, the production function is asymmetric in the sense that type H jobs always produce the same payoff irrespectively of worker type but type L jobs are sensitive to worker absence. Thus, high-sickness absence types constitute *weak links* when having few substitutes.[13]

The matching procedure contains two stages. In the first stage workers and jobs meet randomly. Each worker is matched to one job. Each worker-job pair then decides if they want to continue together and establish a wage or if they want to separate. If a pair separates both the worker and the firm organizing the job incurs a cost *c*. Then, in stage 2 all the agents who have not been matched yet end up with their optimal (frictionless) production partner.

It should be noted that type L jobs are not inherently more productive than type H jobs; they are only more sensitive to sickness absence (as matching with a type H workers generates lower payoff compared to a match with a type L worker). Thus, all other pairings than (x^H, y^L) represent business as usual situations whereas (x^H, y^L) is a bad match both for the job and the worker.[14]

A match in period 1 will be accepted if the period 1 surplus over the period 2 value for the agents is positive. First, consider the case of a type L worker meeting a type L

[11] This idea is related to theoretical work by Weiss (1985), Coles and Treble (1996) and Barmby and Stephan (2000) suggesting that the cost associated with the risk of sickness absence among the employees is larger for small firms than for large firms.
[12] In our data a job is defined as the combination of an establishment and an occupation. The number of employees within such a combination will determine the internal substitutability of a job. A job in a large (small) establishment * occupation group will have a high (low) internal substitutability.
[13] This production process is similar to the one discussed in Kremer (1993).
[14] Formally we have: $f(x^L, y^L) = f(x^L, y^H) = f(x^H, y^H) > f(x^H, y^L)$.

job. Then, it is impossible to get a higher payoff than $f(x^L, y^L)$. The same holds when a type L worker meets a type H job and when a type H worker meets a type H job. However, when a type H worker meets a type L job both agents can find a better match in terms of payoffs .

What are the probabilities that the different matches are accepted in the first stage? The amount of information available for the agents in the first stage will be important for this derivation and thus we will consider varying information levels. Here we find it reasonable to assume that there is perfect information about the job types, i.e. the firms know what type of job that is posted and the workers can observe the type of the job they meet in the first stage. Instead we introduce uncertainty about the worker type and suggest that the degree of proneness to sickness absence is partly unobservable both for the workers and for the jobs. A worker's proneness to sickness absence is arguably hard to observe for an employer and it also seems possible that some workers (e.g. inexperienced and/or unemployed) are not completely sure about their own type.

Since all matches involving type H jobs are accepted (i.e. $f(x^L, y^H) = f(x^H, y^H)$) the information level about worker type will not matter for the acceptance probability in those cases. Formally we can write the acceptance probability for matches involving type H jobs as:

$$P_{jH}^A = 1 \,, \; j \in \{L, H\} \tag{1}$$

where P_{jH}^A represents the acceptance probability for a match between a type j worker and a type H job. For meetings involving type L jobs, however, the agents need to take into account that they face the risk of forming suboptimal matches (i.e. matches between workers of type H and jobs of type L). The payoff loss from forming a suboptimal match can be written as:

$$\Delta HL = f(x^H, y^L) - f^*(x, y) \tag{2}$$

where $f^*(x, y)$ represents the value of an optimal match. By assumption ΔHL is a negative number. A match involving a type L job is accepted if the *expected* payoff loss is less than or equal to the aggregate search cost (i.e. $2c$). Thus, the general expression for acceptance probability for matches involving type L jobs can be written as:

$$P_{jL}^{A} = P\left(\left[-\gamma_j^H \Delta HL\right]/2 \leq c\right), \; j \in \{L, H\} \tag{3}$$

where γ_j^H represents the likelihood that a worker whose true absence type is j is considered to be a type H worker. γ_j^H is hence a measure of the degree of available information about worker-absence-type where $\gamma_H^H = 1$ and $\gamma_L^H = 0$ correspond to a situation with perfect information. With no information at all on the worker-absence-type we get $\gamma_H^H = \gamma_L^H = 0.5$, and assuming that information never can be misleading γ_H^H will vary between 0.5 and 1 and γ_L^H will vary between 0 and 0.5 depending on the preciseness of the information.

Assuming that the search cost c is a stochastic variable Eq. (3) can be further developed and we can express P_{jL}^{A} in the following way:

$$P_{jL}^{A} = 1 - F_c\left(\left[-\gamma_j^H \Delta HL\right]/2\right), \; j \in \{L, H\} \tag{4}$$

where $F_c(\cdot)$ is the cumulative distribution function of the search cost. This expression is useful since we now can take the derivative of P_{jL}^{A} with respect to the information level. An increase in the information level corresponds to an increase in γ_H^H which makes it intuitive to take the derivative of P_{HL}^{A} with respect to γ_H^H to evaluate the impact of better information about worker-absence-type on match acceptance probability. This derivative takes the following form:

$$\partial P_{HL}^{A}/\partial \gamma_H^H = (-1)F_c\left(\left[-\gamma_j^H \Delta HL\right]/2\right)[-\Delta HL]/2 = (-)(+)(+) < 0 \tag{5}$$

The probability of match acceptance for a meeting between a type H worker and a type L job is thus decreasing in the information quality. Correspondingly, P_{LL}^{A} is increasing in the information quality. This result will be important later on.

Ultimately we are interested in the sorting pattern of workers over jobs when all agents have been matched (be it in the first stage or in the second stage). This is also what we can observe empirically. But first we will look at the number of matches of each sort after the first round. Given that workers and jobs meet randomly in the first stage, that the number of jobs and workers is *N*, that the proportion of workers of type *H* is δ^H and that the proportion of jobs of type *H* is η^H (where $\delta^H = \eta^H$) we can write the expected number of accepted matches after the first round between type *H* workers and type *H* jobs as:

$$E(M^1_{HH}) = N\delta^H\eta^H P^A_{HH} = N\delta^H\eta^H \quad (6)$$

We can write the expected number of matches after the first round for the three other type combinations in a corresponding way:

$$E(M^1_{HL}) = N\delta^H\eta^L P^A_{HL} \quad (7)$$

$$E(M^1_{LH}) = N\delta^L\eta^H P^A_{LH} = N\delta^L\eta^H \quad (8)$$

$$E(M^1_{LL}) = N\delta^L\eta^L P^A_{LL} \quad (9)$$

The interesting action then comes from the meetings which did not result in matches in the first stage, i.e. some of the meetings between type *H* workers and type *L* jobs and some of the meetings between type *L* workers and type *L* jobs. If a type *L* worker and a type *L* job reject the meeting in the first stage they will still end up with each other in the second stage since they really are optimal partners. But type *H* workers and type *L* jobs that met each other in the first stage and continued into the second stage need to find new partners. In expectation $N\delta^H\eta^L(1 - P^A_{HL})$ suboptimal meetings continue to the second stage and given the assumption of optimal matching in the second stage, these meetings will generate equally many additional matches both between type *H* workers and type *H* jobs, and between type *L* workers and type *L* jobs. We make the simplifying assumption that agents in unmatched *HL* (type *H* worker, type *L* job) pairs from the first stage can get their optimal partners in the second stage by breaking up already matched *LH* pairs. We allow this simplifying procedure since *L* type workers are insensitive to job type and since *H* type jobs are insensitive to worker type. Given this assumption we can write the expected number of matches for the four type combinations after the two stages as:

$$E(M_{HH}^2) = N\delta^H\eta^H P_{HH}^A = N\delta^H\eta^H + N\delta^H\eta^L(1 - P_{HL}^A) \quad (10)$$

$$E(M_{HL}^2) = N\delta^H\eta^L P_{HL}^A \quad (11)$$

$$E(M_{LH}^2) = N\delta^L\eta^H P_{LH}^A = N\delta^L\eta^H - N\delta^H\eta^L(1 - P_{HL}^A) \quad (12)$$

$$E(M_{LL}^2) = N\delta^L\eta^L P_{LL}^A + N\delta^L\eta^L(1 - P_{LL}^A) + N\delta^H\eta^L(1 - P_{HL}^A)$$
$$= N\delta^L\eta^L + N\delta^H\eta^L(1 - P_{HL}^A) \quad (13)$$

Given the initial assumption of suboptimality of the match between a type H worker and a type L job we expect to see some degree of sorting of L type workers to L type jobs. We measure this sorting by putting the quotient between $E(M_{LL}^2)$ and $E(M_{HL}^2)$ in relation to the quotient between $N\delta^L$ and $N\delta^H$. We do this by forming the following expression:

$$\Omega_{LL} = [E(M_{LL}^2)/E(M_{HL}^2)]/[N\delta^L/N\delta^H] = [E(M_{LL}^2)/E(M_{HL}^2)]/[\delta^L/\delta^H] =$$
$$\left[\left(N\delta^L\eta^L + N\delta^H\eta^L(1 - P_{HL}^A)\right)/(N\delta^H\eta^L P_{HL}^A)\right]/[\delta^L/\delta^H] =$$
$$\left[\left(\delta^L + \delta^H(1 - P_{HL}^A)\right)/(P_{HL}^A)\right]/[\delta^L/1] = [\delta^L + \delta^H(1 - P_{HL}^A)]/[\delta^L P_{HL}^A] =$$
$$[1/P_{HL}^A] + [\delta^H/(\delta^L P_{HL}^A)] - [(\delta^H P_{HL}^A)/(\delta^L P_{HL}^A)] \quad (14)$$

where Ω_{LL} represents the relative overrepresentation of type L workers in type L jobs. Obviously, if this measure is greater than 1, we have that type L workers are overrepresented in type L jobs. If $P_{HL}^A = 1$, we will have no overrepresentation and hence Ω_{LL} will be equal to 1. But from Eq. (4) and the fundamental assumption that ΔHL is a negative number we get that P_{HL}^A always is lower than 1 and subsequently that Ω_{LL} always is greater than 1. Thus, we have reached our main prediction (**Prediction 1**) which states that type L workers should be overrepresented in type L jobs. Empirically, this means that employees in jobs with relatively few substitutes should exhibit lower sickness absence than employees in jobs with relatively many substitutes.

How does the preciseness of the information about worker-absence-type affect the overrepresentation rate? We can draw conclusions about this by taking the derivative of Ω_{LL} with respect to γ_H^H:

$$\partial\Omega_{LL}/\partial\gamma_H^H = (-1)[1/(P_{HL}^A)^2](\partial P_{HL}^A/\partial\gamma_H^H)$$
$$+(-1)[\delta^H/(\delta^L(P_{HL}^A)^2)](\partial P_{HL}^A/\partial\gamma_H^H) = (-)(+)(-) + (-)(+)(-)$$
$$= (+) + (+) > 0 \qquad (15)$$

where $\partial P_{HL}^A/\partial\gamma_H^H < 0$ comes from Eq. (5). Thus, more precise information about worker absence type tends to increase the degree of overrepresentation of type *L* workers in type *L* jobs. This is the second main prediction (*Prediction 2*) that comes out from the model. Empirically, this implies that the relation between the number of coworker substitutes and sickness absence should be less clear among workers carrying relatively worse information about their absence type. Correspondingly, carrying relatively good information about absence type should be associated with having low sickness absence within the pool of type *L* jobs while good information about absence type should be associated with high sickness absence in the pool of type *H* jobs. This is because type *H* workers that carry precise information about their absence type ultimately end up in type *H* jobs while type *H* workers that carry imprecise information might stay in type *L* jobs.

2.2 Extension: updated information about absence type

The model outlined above is a two-period model, where all agents receive their optimal match in the second stage. In reality, however, it is reasonable that the agents reassess the usefulness of the match once the worker‘s absence type is revealed (i.e. when the worker gets some tenure). Since the worker-absence-type does not matter for type *H* jobs in the model we do not expect that matches involving type *H* jobs should respond to the revelation of the true absence type. However, matches between type *H* workers and type *L* jobs that were formed because of bad information conditions might be resolved when the agents are updated about the true type of the worker. Therefore we expect that high levels of realized sickness absence (i.e. absence in the new job) should be associated with match separation in the pool of matches involving type *L* jobs (*Outside-of-model prediction 1*). Within the pool of matches involving type *L* jobs we also expect this association to be particularly pronounced among matches formed under relatively worse information about worker-absence-type, as actual absence behavior should contain more new information when the precision of the prior was poor (*Outside-of-model prediction 2*).

2.3 Summary of predictions

The model outlined in section 2.1 generates the following two main *predictions* that we take to the data:

- *Prediction 1*: We expect that type *L* workers should be overrepresented in type *L* jobs. Empirically, this means that employees in jobs with relatively few substitutes should exhibit lower sickness absence than employees in jobs with relatively many substitutes.
- *Prediction 2*: We expect that more precise information about worker absence type should increase the degree of overrepresentation of type *L* workers in type *L* jobs. Empirically, this implies that the relation between the number of coworker substitutes and sickness absence should be less clear among workers carrying relatively worse information about their absence type. Correspondingly, carrying relatively good information about absence type should be associated with having low sickness absence within the pool of type *L* jobs while good information about absence type should be associated with high sickness absence in the pool of type *H* jobs.

As an extension to the model we also expect the following relation between realized sickness absence and job separation when allowing for updated information about worker's absence type:

- *Outside-of-model prediction 1*: We expect that high levels of realized sickness absence (i.e. absence in the new job) primarily should be associated with match separation in the pool of matches involving type *L* jobs.
- *Outside-of-model prediction 2*: Within the pool of matches involving type *L* jobs we expect the above association to be particularly pronounced among matches formed under relatively worse information about workers' absence type.

3 Data

3.1 Data sources and definitions

We use Swedish register data for the private sector between the years 1997–2007. These data are drawn from registers administered by Statistics Sweden that follow all Swedish workers from 1985–2010, with unique person, firm and establishment identifiers. To these data we add socioeconomic background characteristics from a population-wide dataset and information on occupational codes, which is available from 1997–2010 for a sample of private establishments covering almost 50% of private sector workers.[15]

We use the occupational codes to define *substitutes* as the number of other workers within the same combination of establishment and occupation (ISCO-88, 3d) a given year. For example, an administrator at a workplace that in total employs four administrators will have three substitutes[16] In order to focus on regular workers, we drop employees in managerial positions. We also drop employees employed at very small establishments (less than three employees).

We add sickness absence spells from the Swedish Social Insurance Agency. These data include all spells longer than two weeks.[17] Sickness absence will generally be defined as an indicator for having at least one such spell in a given year. The fact that we cannot observe shorter spells is obviously a shortcoming of the data, and we will therefore complement our analysis with short-term work absence due to caring for sick children as an alternative absence measure. In Sweden, parents with small children (0–10 years old) can be absent from work to care for sick children (that are too sick to be in school or in daycare). The parent that stays home will then receive benefits from the Social Insurance system from day one meaning that these benefits data also pick up short term absence spells.[18]

[15] We start the observation period in 1997 since this is the first year that we can observe occupations. The reason for ending already in 2007 is that we in some cases want to follow workers for a 3-year follow-up period.

[16] In our main sample of new hires, the shares of employees having 0 substitutes in our 8 broad occupation groups (based on the first digit in the occupational code) are: Professionals (5.3 percent), Technicians and associate professionals (4.7 percent), Clerks (8.2 percent), Service workers and shop sales workers (1.4 percent), Skilled agricultural and fishery workers (5.9 percent), Craft and related trades workers (2.7 percent), Plant and machine operators and assemblers (0.8 percent) and Elementary occupations (3.4 percent). This indicates that unique positions are present, to roughly the same extent, in all occupational skill levels.

[17] The data include all spells for which the individual was entitled to sickness benefits from the social insurance system. Since spells shorter than two weeks are paid by the employers, these are not available in our data.

[18] Parents may claim benefit compensation for up to 120 days per year. The replacement rate is 80 percent of lost earnings up to a monthly wage ceiling of SEK 37,000. The benefit compensation data contains information on the total amount of child sick benefits received each year, from which we construct an indicator for having at least one child sick spell in a given year.

We examine the role of worker sorting in more detail using a dataset consisting only of new hires. We define new hires as employees observed in a workplace in a given year, but not in the same workplace or in the same firm in any of the five preceding years. For each hire, we determine their *pre-hire* sickness absence. The term pre-hire sickness absence is defined as the average incidence of having at least one sickness absence spell longer than two weeks per year in the three years prior to employment. In order for all new hires to have three proper pre-hire years, we restrict the sample to workers with at least 4 years of labor market experience.[19] Pre-hire sickness absence is potentially correlated with employment status during the relevant time period. To deal with this issue we control for pre-hire employment probability in our empirical model. We will also examine the probability of job separation when the worker-absence-type is revealed. To this end we study the relation between realized sickness and job separation, where realized sickness is defined as the average sickness absence probability in t and t+1.

3.2 Worker type and job type measures

In section 2.1 we outlined a model with two worker types ($x^H$and x^L) and two job types ($y^H$and y^L). When we study the relationship between sickness absence and the number of substitutes at work in Section 4.1, using the sample of all workers, we measure worker-absence-type in terms of present sickness absence (i.e. sickness absence in year t). When we study new hires, however, we are more interested in their sickness behavior before they were recruited and hence worker-absence-type will here be measured in terms of pre-hire sickness absence (see definition in Section 3.1). Generally, we consider a job to be a type L job (few substitutes) if the number of substitutes is less than five,[20] but in Section 4.1 we do exercises that contain variations of this definition.

3.3 Information measures

One part of our empirical analysis aims to contrast realized matches between workers and jobs that took place under more or less information about worker-absence-type. To

[19] In other words, labor market entrants with less than 4 years since they graduated from their highest education are excluded.
[20] Jobs with more than five substitutes constitutes the left-out reference category in the empirical model.

this end, we use three different proxies for the amount of information about worker-absence-type that was available to both agents when they matched:

- *Pre-hire employment:* an indicator variable taking on the value 1 if the new employee was employed in *t*-3, *t*-2 and *t*-1.
- *Firm connection:* an indicator variable taking on the value 1 if the new employee was employed by the firm, but in another establishment, during *t*-5 to *t*-1.
- *Network connection:* an indicator variable taking on the value 1 if the new employee has a common workplace history with at least one of the incumbent employees.[21]

These information proxies are all based on the notion that a worker's employment history can provide both sides (prospective hires and employers) with updated information about the employee's sickness absence probability. Hence, when new hires fulfill one of the three criteria above, we assume the hiring decision was based on a more precise signal about future absence behavior. Although it is clear that these measures are far from perfect, several studies support our choice of information proxies in the hiring decision. Recent work by Eriksson and Rooth (2014) shows that employers are reluctant to hire people from non-employment which indicates that non-employment is associated with some degree of uncertainty about worker type. Thus, when matches involving workers with a weak pre-hire employment record actually takes place we find it reasonable to assume that they are less informed.[22] Schönberg (2007) further provides evidence that the traces of information that employees leave about themselves to some extent only can be accessed by the employer where the employee works. Under this assumption, we expect that matches involving workers with an earlier connection to the recruiting firm are better informed about the worker-absence-type. Finally, there is also recent evidence that incumbent employees with previous labor market links to entering employees can provide useful information about the properties of the entering workers,

[21] For each new hire we construct dyads consisting of the new hire him-/herself and each incumbent worker (i.e. if a new worker comes to a workplace with 10 incumbent workers we create 10 dyads). For each dyad we add information on the workplace of the incumbent and the new worker, respectively, for all years going back to 1985. If the workplaces match for at least one previous year the new hire is defined as being connected to the incumbent worker. If the new hire is connected to at least one incumbent worker at the recruiting workplace, the new hire is defined as having a network connection to the workplace.

[22] Farber and Gibbons (1996), Altonji and Pierret (2001) and Lange (2007) show that employers overprice formal credentials (and underprice "hidden talents") among inexperienced workers, which suggests that there is less information about worker type for employees with weaker labor market experience.

which is why we consider matches that are formed in the presence of network connections to be more informed (see Hensvik and Skans [forthcoming]).

3.4 Data description

Table A1 and Table A2 in the Appendix contain descriptive statistics on the sample of all workers and on the sample of new hires respectively. We have about 6 million observations in the full sample (Table A1) and 400,000 new hires (Table A2). In the full sample about 4 percent of the workers have truly unique jobs (0 substitutes) and about 11 percent of the employees have at least one sickness absence spell that is longer than two weeks in a given year.[23] Consistent with our model, the incidence of sickness absence is lower for workers in unique positions, but these workers differ in other aspects as well: they are for example employed in smaller establishments and in more skilled professions with higher wages suggesting that they have key positions within the firms.[24] Workers in relatively unique positions are in addition older and more often women, although education levels appear similar as to other employees.

The image of the new hires is very much in line with the full sample. Worth noting, however, is that contrary to our model predictions workers hired into relatively unique positions have somewhat higher pre-hire sickness absence whereas entry wages are about the same. But as noted before, it is important to account for other aspects that differ systematically between more/less unique positions before we can draw conclusions about the relationship between employee absence and internal substitutability.

4 Empirical relations between employee absence and coworker substitutes

In this section we provide empirical tests of the predictions generated in Section 2. We estimate the relationship between having a job with few substitutes and sickness absence while controlling for individual characteristics, establishment size, establishment and occupational fixed effects. Formally we use the following general model:

$$Y_{ijpt} = \gamma S_{ijpt} + \alpha_j + \alpha_p + \theta_t + \beta X_{it} + \delta Z_{jt} + \varepsilon_{ijpt} \qquad (16)$$

[23] The figure on sickness absence is confirmed by estimates from Statistics Sweden (Statistics Sweden, 2007).
[24] The summary statistics show the distribution of workers/hires across a broader set of occupations (1-digit level). When defining the number of substitutes we use more detailed occupation codes (3-digit level).

where the outcome (Y_{ijpt}) is a sickness absence measure for worker (i) in establishment (j) and profession (p) in year (t); S_{ijpt}, is an indicator for the number of substitutes within each "job", defined by the interaction between the establishment and the 3-digit occupational code (see section 3.1). α_j and α_p are workplace and profession fixed effects respectively. We also include year fixed effects θ_t to further account for, for example, business cycle swings potentially correlated with firms' organization of work and individual sickness absence. The worker characteristics (X_{it}) consist of gender, age, education, country of origin and an indicator for having children under the age of three.[25] Finally we include workplace size (Z_{jt}). ε_{ijp} is the error term.

The parameter of interest is γ, which aims to capture the influence from the number of internal substitutes on employee absence. In general we treat jobs with more than five substitutes as the reference category and hence consider a job to be a type L job (few substitutes) if the number of substitutes is less than five. We will, however, also employ subcategories within this category in order to investigate potential patterns.

It should be noted that the model is fairly rich as it identifies the relationship between the *job-specific* number of substitutes and employee absence, while accounting for any spurious correlation between the two caused by unobserved differences between particular occupations or establishments through α_j and α_p.

In order to test the predictions of our model, we also want to disentangle to which extent γ captures behavioral responses and/or employee selection. As a first step we therefore include worker fixed effects to Eq. (16), which will account for the selection of workers over jobs with few/many substitutes. In addition, we estimate the model separately for new hires using their pre-hire sickness absence as the outcome of interest.

4.1 Results for all workers

We start our empirical exercise by exploring the association between present sickness absence and the number of coworker substitutes among all private sector workers. While Prediction 1 in Section 2.3 concerns selection of low-absence workers into jobs with few substitutes and therefore is more relevant for new hires, it still implies that the number of substitutes should be associated with sickness absence in the existing workforce. This association therefore constitutes a preliminary test of Prediction 1, but

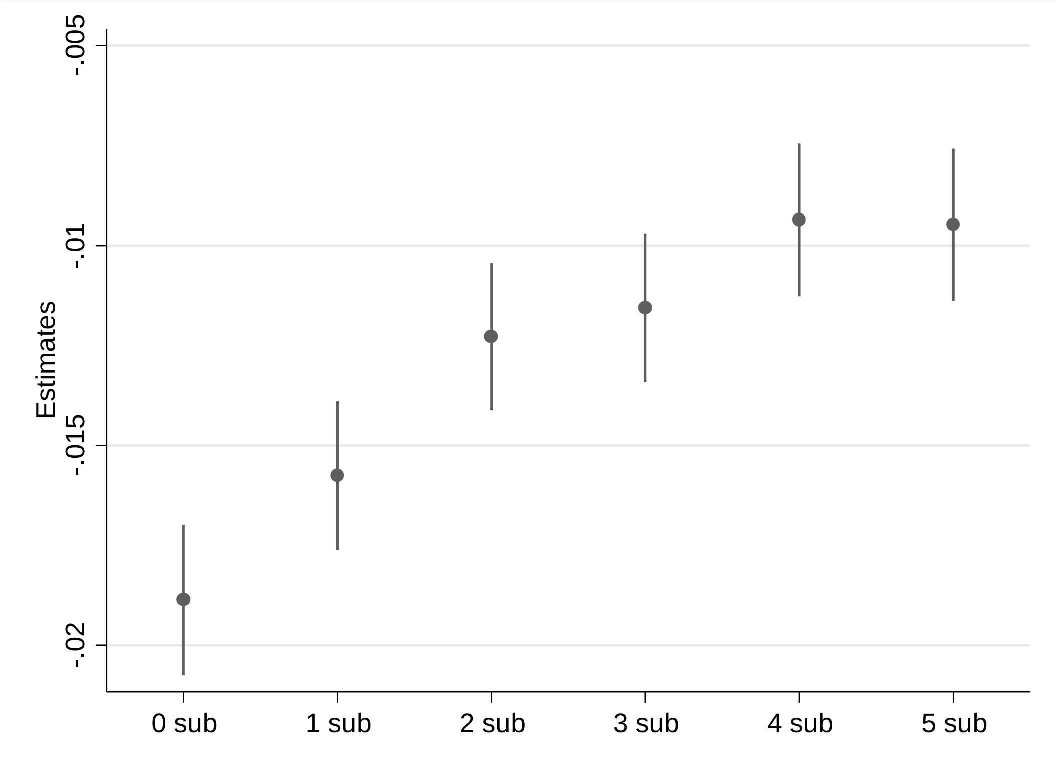
[25] We group individuals by their country of origin into the following six categories: Sweden, rest of the Nordic countries, rest of Europe, North America, South America, and the rest of the world.

the estimates will also pick up effects that go beyond pure selection, e.g. potential causal effects on sickness absence of working in a job with few substitutes.

Holding this in mind we start with a graphical representation of the results in Figure 1. We obtain the estimates by estimating Eq. (16) with separate indicators for having 0–5 substitutes (employees with more than 5 substitutes constitute the reference category). The estimates in Figure 1 are all statistically negative on the 1-percent level, ranging between 1 and 2 percentage points, supporting our general hypothesis that jobs with few substitutes are more sensitive to sickness absence. Interestingly, the estimates become smaller in absolute value as the number of employees performing the same job increases, which is consistent with the idea that the costs associated with sickness absence increase, in terms of production disruptions, as the number of substitutes decreases. In the Appendix (Figure A1) we show the same relationship for 0–10 substitutes. Interestingly, these results suggest that there is a a significant “jump” in the absence probability when the number of substitutes increases from 0 to 1. Beyond that, there is a fairly linear relationship between employee absence and the number of substitutes.[26] The magnitudes of the estimates are substantial, especially for the coefficients on 0 and 1 substitutes: the difference in sickness absence between jobs with more than 5 substitutes and jobs with 0 substitutes, conditional on the model, is roughly equivalent to the estimated difference in absence rates between workers in their 20s and 40s, or between workers with and without small children (0–3 years of age).

[26] The difference in absence probability between jobs with ten and more than ten substitutes is around 0.5 percentage points. This remaining difference may seem surprising but could be due to the fact that we have measurement error in the possibilities of substituting an absent employee which is likely to decrease with the size of the job-cell (in large cells there is a greater chance that at least some workers are perfect substitutes for each other).

Figure 1 Present sickness absence and the number of coworker substitutes



Notes: The standard errors are clustered on the workplace level. The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children and establishment size. The model also includes year,, occupational, and workplace fixed effects.

Table 1 shows the point estimates (with and without worker characteristics) when we only use two categories for having few substitutes (0–1 substitutes and 2–5 substitutes). Like before the reference category is employees with more than 5 substitutes. Overall, these results are consistent with the prediction that there should be selection of low-absence-type workers into jobs with few substitutes, but these results can potentially also be partly explained by endogenous adjustments taking place once in a job with few substitutes. To separate between these two explanations we exploit variation in the number of substitutes for the same worker over time by including individual fixed effects (see column [3]). The point estimates are roughly halved by the inclusion of individual fixed effects but remain significantly negative on the 1-percent level, suggesting that workers that become relatively unique reduce their sickness absence. Taken together, the results suggest that the correlation between internal substitutability and sickness absence entails both a selection component and an endogenous component that appear to be of similar importance.[27]

While the finding that the effect partly survives the inclusion of individual fixed effects is interesting, data limitations restrict us from exploring the foundations behind the result that having a job with few substitutes seem to change employees' sickness absence behavior. Instead the available data is better equipped at investigating the part of the relation that pertains to selection of low absence workers into jobs with few substitutes. In Sections 4.3 and 4.4 we extend the analysis in this dimension to various outcomes for new hires in order to deepen our understanding of the processes involved.

Table 1 Present sickness absence and the number of coworker substitutes

Outcome: Sickness absence in *t*	(1)	(2)	(3)
0–1 substitutes	-0.0132***	-0.0172***	-0.0089***
	(0.0008)	(0.0008)	(0.0010)
2–5 substitutes	-0.0089***	-0.0107***	-0.0048***
	(0.0006)	(0.0006)	(0.0007)
Number of observations	5,863,497	5,863,497	5,863,497
Mean of dependent variable	0.109	0.109	0.109
Background controls	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Occupation fixed effects	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes
Worker fixed effects	No	No	Yes

Notes: The standard errors are clustered on the workplace level (on the worker level in column [3]). The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children and establishment size. *, **, and *** denote statistical significance at the 10-,5-, and 1-percent level.

4.2 Additional results/robustness checks

The strong association between sickness absence and coworker substitutability naturally raises the relevant question whether working in a unique position is associated with a higher wage. Table A4 in the Appendix suggests that is indeed the case. We obtain these estimates by replacing sickness absence as the outcome in Eq. (16) with the monthly full-time wage. The results in column (2) suggest that employees with 0–1

[27] We have also estimated the effects in Table 1 separately for different sizes of the workplace. We have results for small (3–59 employees), medium-sized (60–339 employees) and large (340–6815 employees) workplaces. The results are presented in Table A3 in the Appendix. The results for small and medium-sized workplaces are similar to the ones presented in Table 1. The results for large workplaces stand out in the sense that the estimates for 0–1 substitutes are lower than the estimates for 2–5 substitutes. However, we should bear in mind that these are really large workplaces and it is unclear what exactly having few substitutes represents here.

substitutes have 1.6 percent higher wage on average relative to employees with more than 5 substitutes.[28]

This finding suggests that unique jobs are more productive in general, and/or that the employees having unique jobs have higher (unobserved) skills. When we account for individual heterogeneity in column (3) the wage premium reduces to 0.3 percent. Thus, most of this relationship reflects that high wage types sort into jobs with few substitutes. The result raises a potential concern against the interpretation of our results, namely that sickness absence may be correlated with other employee characteristics that are more productive in unique positions (e.g. ability). In Panel A of Table A5 we therefore show the estimated relationship between substitutes and absence holding the wage constant. Even if it is potentially problematic to control for the wage (as this is likely to be endogenous to the number of substitutes), it is reassuring to see that this has a minor impact on the main estimates.

Panel B of Table A5 shows the estimates when we add the public sector employees to our sample. These estimates are somewhat smaller, but still significant and of important magnitude both in the model with and in the model without individual fixed effects suggesting that the relationship between the number of coworker substitutes and sickness absence probability holds in the full economy. In Panel C of Table A5 we use data on private sector employees for the years 2005–2007. In these years the occupational code is available on a 4-digit level and thus we can test if our main results in Table 1, which are based on a 3-digit occupational code, are robust to finer definitions of the occupations. Since we only have data covering three years a model with individual fixed effects seems inappropriate and we therefore restrict the analysis to a model corresponding to column (2) in Table 1. The estimates we find are very similar to the ones in column (2) in Table 1 and thus the results based on the 4-digit occupational code confirm the general picture of low sickness absence in jobs with few substitutes.

Finally, we test the robustness of our results using two alternative absence measures: *probability of caring for sick child* (see section 3.1), which also includes short-term work absence[29] and *log of sickness benefits* which picks up the duration of the sickness absence in excess of the two first weeks of each sick period. With respect to caring for sick child we restrict the sample to individuals with at least one child between 0 and 10

[28] Estimating the same model for new hires we find an identical wage premium.
[29] The reason is that parents receive benefits from the Social Insurance System from day one.

years old (these are the children that parents are entitled to be at home with) and use an indicator for having positive sick-child benefits in a given year as the outcome. 65 percent of the parents have at least one absence spell according to our definition, which suggests that caring for sick children is a first-order source of work absence among parents. With respect to log of sickness pay we restrict our sample to observations with positive values on sickness benefits, i.e. we measure the length of the absence spell conditional on having at least one absence spell longer than two weeks.

In Table 2 we present the results from this exercise focusing on a model without worker fixed effects. The results in column (1) clearly show that workers in jobs with few substitutes are significantly less likely to be home caring for sick children. Thus, the results are in line with our general findings in Table 1, although compared to the baseline they are smaller in magnitude.[30] Also when we use the log of the received sickness benefits (column [2]) among those having absence spells longer than two weeks as the outcome of interest we find similar results. Conditional on being absent, employees in positions with 0–1 substitutes have absence spells that are roughly two percent shorter (the received benefits are closely related to the length of the absence spell) than employees in positions with more than 5 substitutes.[31]

Table 2 Alternative absence measures and the number of coworker substitutes

	(1)	(2)
Outcome:	**Pr(sick child absence)**	**Log of sickness pay**
0–1 substitutes	-0.0182***	-0.0227**
	(0.0024)	(0.0110)
2–5 substitutes	-0.0081***	-0.0179**
	(0.0019)	(0.0087)
Number of observations	1,911,734	638,409
Mean of dependent variable	0.654	4.822
Background controls	Yes	Yes
Year fixed effects	Yes	Yes
Workplace fixed effects	Yes	Yes
Occupation fixed effects	Yes	Yes
Worker fixed effects	No	No

Notes: The standard errors are clustered on the workplace level. The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children and establishment size. In column (1) we restrict the sample to individuals with at least one child less than 11 years of age. We further control for the number of children in the following categories: 0–3 years, 4–6 years and 7–10 years. In column (2) we restrict the sample to observations with positive values on sickness benefits. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

[30] When we include worker fixed effects in the model the estimates go to zero indicating that being in a job with few substitutes does not decrease the likelihood of caring for sick children per se. Instead the difference in the estimates between the models with and without worker fixed effects respectively indicates that the negative association displayed in column (1) mainly is driven by selection of low-absence-types into jobs with few substitutes.

[31] When we include worker fixed effects in the model the estimate for 0–1 substitutes goes up to about -0.07. This result is, however, hard to interpret since the variation comes from a very selected group of individuals.

4.3 Results for new hires

4.3.1 Pre-hire sickness absence and selection into jobs

The results in Table 1 and Table 2 are consistent with the prediction that there should be selection of low-absence-type workers into jobs with few substitutes. In this section we provide more direct empirical tests of Prediction 1 from Section 2.3 using a sample of new hires.

First, we use the pre-hire sickness absence described in Section 3.1 as the outcome in Eq. (16) (using the same specification as in column [2] of Table 1). To gain precision, we group entrants according to whether they are hired into jobs with more/less than five coworker substitutes. Consistent with Prediction 1 and our earlier results, Table 3 shows that workers hired into positions with fewer coworker substitutes have lower *pre-hire* sickness absence. Employees hired to jobs with few substitutes had 0.4 percentage points lower pre-hire sickness absence (from a baseline pre-hire sickness probability of 11.6 percent) than employees entering jobs with more than 5 substitutes. Interestingly, the magnitude of this effect is similar to the difference between the estimates in column (2) and column (3) of Table 1, which supports that workers with few sick spells sort into jobs with low substitutability.

Table 3 Pre-hire sickness absence and the number of coworker substitutes

Outcome: Pre-hire sickness absence	(1)
0–5 substitutes	-0.0043***
	(0.0014)
Number of observations	387,901
Mean of dependent variable	0.116
Background controls	Yes
Year fixed effects	Yes
Occupation fixed effects	Yes
Workplace fixed effects	Yes
Worker fixed effects	No

Notes: The standard errors are clustered on the workplace level. The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children, pre-hire employment probability and establishment size. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

4.3.2 Realized sickness absence and post-hire outcomes

We have so far focused on employee absenteeism and the selection into jobs. But in Table 4 we complement the analysis by asking how entrants' realized sickness absence (i.e. the average sickness absence probability in t and t+1) affects (i) the probability of exiting the employment relationship as well as (ii) the probability of receiving more coworker substitutes within three years after entry. We study (i) using a sample of new hires that stay on the workplace at least until *t*+1. The outcome is an indicator for not being in the workplace in *t*+2 (Panel A), i.e. exiting the workplace between *t*+1 and *t*+2.[32] We study (ii) using a sample of new hires that are observed on the workplace in *t*+1, *t*+2 and *t*+3. The outcome is an indicator for having more substitutes in *t*+3 than in *t* (Panel B).

The results suggest that higher realized sickness absence is associated with significantly higher turnover rates (Panel A, column [1]), and a higher likelihood of receiving more coworker substitutes (Panel B, column [1]).[33] Consistent with Out-of-model prediction 1 this relationship is particularly strong for workers employed in relatively unique positions (columns [2–4]).

Table 4 Realized sickness absence and post-hire adjustments in different job types

	(1)	(2)	(3)	(4)
Panel A. **Outcome: Not on the workplace in *t*+2**	*All*	*Jobs with subs.≤5*	*Jobs with subs.>5*	*Difference*
Realized sickness	0.1100***	0.1284***	0.1072***	0.1072***
	(0.0037)	(0.0097)	(0.0040)	(0.0040)
Realized sickness * 0–5 substitutes				0.0212**
				(0.0105)
Observations	336,026	63,624	272,402	336,026
Mean of dependent variable	0.270	0.280	0.267	0.270
Panel B. **Outcome: More substitutes in *t*+3**	*All*	*Jobs with subs.≤5*	*Jobs with subs.>5*	*Difference*
Realized sickness	0.0152**	0.0582**	0.0134*	0.0134*
	(0.0072)	(0.0248)	(0.0075)	(0.0075)
Realized sickness * 0–5 substitutes				0.0448*
				(0.0258)
Observations	121,195	20,254	100,941	121,195
Mean of dependent variable	0.485	0.444	0.494	0.485
Background controls	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Occupation fixed effects	Yes	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes	Yes
Worker fixed effects	No	No	No	No

Notes: In Panel A the sample is restricted to new hires that are observed on the workplace also in *t*+1. In Panel B the sample is restricted to new hires that are observed on the workplace in *t*+1, *t*+2 and *t*+3. The standard errors are clustered on the workplace level. The background controls are gender, age, education, birth country, having small children and establishment size. In column (4) all variables are interacted with the 0–5 substitutes indicator. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

[32] As a robustness check we have also used an indicator for not being observed on the workplace in either *t*+2 or *t*+3. This does not substantially change the results.

[33] Interestingly, when we condition on being observed on the workplace in *t*+1 and *t*+2 and use the average sickness absence probability in *t*+1 and *t*+2 as an explaining variable for leaving the workplace in *t*+3 the estimate in Panel A, column (1), is substantially lower. This is consistent with the notion that the marginal effect of exhibiting bad properties (in this case high sickness absence), in relation to the job, on job separation probability should decrease with tenure (see Kwon, 2005 for an interesting contribution on this topic).

4.4 The role of information

The fact that job separations respond to realizations of sickness absence suggests that matches are formed under some remaining uncertainty. In this section we examine the direct importance of the amount of information about worker-type available when the matches were formed for the selection into and out of jobs.

First, we expect sorting of low-absence-type workers into jobs with few substitutes to be less pronounced among matches formed under relatively worse information conditions (Prediction 2 in Section 2.3). In Table 5 we estimate Eq. (16) for matches formed under more/less precise information about worker-absence-type. We use the information proxies described in Section 3.3: an indicator for strong pre-hire employment (Panel A), an indicator for previous firm connection (Panel B) and an indicator for previous labor market connection to incumbent employees (Panel C).[34] Consistent with Prediction 2 the estimates in column (1) (more informed matches) are generally more negative than the corresponding estimates in column (2) (less informed matches). In column (3) we compare the more informed to the less informed matches in a joint framework in order to be able to say something about the statistical significance of the difference in the estimates between the two. The estimates of the interaction effect between good information and a job with few substitutes are generally negative but fail to exhibit significance on conventional levels.

[34] When previous firm connection is used as the information proxy we relax the new hire definition and include new hires on the workplace with a history within the firm, which explains why the sample size is larger in Panel B.

Table 5 Pre-hire sickness absence and the number of coworker substitutes under precise respectively imprecise information on worker-absence-type

	Outcome: Pre-hire sickness		
	(1)	(2)	(3)
Panel A. Information proxy: Pre-hire employment	*More informed*	*Less informed*	*Difference*
0–5 substitutes	-0.0060***	-0.0010	-0.0010
	(0.0014)	(0.0060)	(0.0060)
0–5 substitutes * Pre-hire employment			-0.0050
			(0.0062)
Number of observations	321,641	66,260	387,901
Mean of dependent variable	0.107	0.159	0.116
Panel B. Information proxy: Firm connection	*More informed*	*Less informed*	*Difference*
0–5 substitutes	-0.0070***	-0.0043***	-0.0043***
	(0.0019)	(0.0014)	(0.0014)
0–5 substitutes * Firm connection			-0.0027
			(0.0023)
Number of observations	199,093	387,901	586,994
Mean of dependent variable	0.113	0.116	0.115
Panel C. Information proxy: Network connection	*More informed*	*Less informed*	*Difference*
0–5 substitutes	-0.0048*	-0.0050***	-0.0050***
	(0.0026)	(0.0016)	(0.0016)
0–5 substitutes * Network connection			0.0002
			(0.0031)
Number of observations	135,698	252,203	387,901
Mean of dependent variable	0.110	0.119	0.116
Background controls	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Occupation fixed effects	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes
Worker fixed effects	No	No	No

Notes: The standard errors are clustered on the workplace level. The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children, pre-hire employment probability and establishment size. In column (3) all variables are interacted with Pre-hire employment/Firm connection/Network connection. In Panel B, we relax the new hire definition and include new hires on the workplace with a history within the firm, which explains why the sample size is larger than in Panels A and C. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

Second, we examine Prediction 2 in section 2.3 using an alternative empirical approach. The prediction implies that carrying relatively good information about absence type should be associated with having low sickness absence within the pool of type *L* jobs while good information about absence type should be associated with high sickness absence in the pool of type *H* jobs. This is because type *H* workers that carry precise information about their absence type ultimately end up in type *H* jobs (because they are blocked from entering type *L* jobs) while type *H* workers that carry imprecise information might stay in type *L* jobs. In column (1) of Table 6 we study the overall relation between pre-hire sickness absence and the amount of information about the worker-absence-type that was available at the time of the match. For the pre-hire employment information measure (Panel A) and the firm connection information measure (Panel B) the relationship is statistically negative as predicted while the network information measure (Panel C) delivers a zero result. In column (2) we restrict the analysis to jobs with less than 6 substitutes and in column (3) to jobs with more than 5 substitutes. Consistent with Prediction 2 the estimates are generally more negative in column (2) indicating that type H workers and type L jobs rarely match under good information conditions. Contrary to Prediction 2, however, we see a significantly negative relation also in column (3) (at least in Panels A and B). This is arguably due to the fact that our empirical counterpart to a type H job is imprecise. Thus, it is possible to imagine that good information, to some extent, can lead to blocking of high absence type workers also among jobs in this category. Instead we find it more relevant to examine the differences between the estimates in columns (2) and (3), which is what we do in column (4). The interaction effects in Panels A and B are significantly negative while the interaction effect in Panel C is negative but not significant. Overall the results suggest that there are significant differences in the relation between information about worker-absence-type and pre-hire sickness depending on the category of the job (i.e. few substitutes or many substitutes) which is in line with Prediction 2.

Table 6 The relation between precise information on worker-absence-type and pre-hire sickness in different job types

	Outcome: Pre-hire sickness			
	(1)	(2)	(3)	(4)
Panel A. Information proxy: Pre-hire employment	*All*	*Jobs with subs.≤5*	*Jobs with subs.>5*	*Difference*
Pre-hire employment	-0.0251***	-0.0457***	-0.0208***	-0.0208***
	(0.0014)	(0.0040)	(0.0015)	(0.0015)
Pre-hire employment * 0–5 substitutes				-0.0249***
				(0.0043)
Number of observations	387,901	73,366	314,535	387,901
Mean of dependent variable	0.116	0.125	0.114	0.116
Panel B. Information proxy: Firm connection	*All*	*Jobs with subs.≤5*	*Jobs with subs.>5*	*Difference*
Firm connection	-0.0108***	-0.0159***	-0.0093***	-0.0093***
	(0.0010)	(0.0022)	(0.0011)	(0.0011)
Firm connection * 0–5 substitutes				-0.0065***
				(0.0024)
Number of observations	586,994	116,023	470,971	586,994
Mean of dependent variable	0.115	0.119	0.114	0.115
Panel C. Information proxy: Network connection	*All*	*Jobs with subs.≤5*	*Jobs with subs.>5*	*Difference*
Network	-0.0014	-0.0052*	-0.0008	-0.0008
	(0.0009)	(0.0028)	(0.0010)	(0.0010)
Network * 0–5 substitutes				-0.0045
				(0.0029)
Number of observations	387,901	73,366	314,535	387,901
Mean of dependent variable	0.116	0.125	0.114	0.116
Background controls	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes	Yes
Occupation fixed effects	Yes	Yes	Yes	Yes
Worker fixed effects	No	No	No	No

Notes: The standard errors are clustered on the workplace level. The background controls are gender, age, education, birth country, having small children, pre-hire employment probability (not included in Panel A) and establishment size. In column (4) all variables are interacted with the 0–5 substitutes indicator. In Panel B, we relax the new hire definition and include new hires on the workplace with a history within the firm, which explains why the sample size is larger than in Panels A and C. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

Finally, Table 7 shows how job separations induced by realized sickness absence are related to the information available in the hiring stage. Intuitively, separations should respond more strongly to high levels of realized sickness absence if matches were formed under relatively worse information conditions. We restrict this analysis to jobs with less than 6 substitutes and interact realized absence with our information proxies. Consistent with Out-of-model prediction 2, we find that the effect of realized sickness absence on leaving the workplace generally is larger for less informed matches (columns [1–2]). Depending on the exact information proxy, the point estimates are about 4–7 percentage points larger for less informed matches. The results in column (3) show that the difference between more/less informed matches is insignificant using strong pre-hire employment as the information proxy (Panel A), highly significant (1-percent level) using previous firm connection as the information proxy (Panel B) and marginally significant (10-percent level) using network connection as the information proxy (Panel C). These findings suggest that agents in less informed matches are more likely to experience negative surprises that affect the continuation of matches as the worker-absence-type is revealed.

Table 7 Realized sickness, post-hire separation and information

	Outcome: Not on the workplace in t+2		
	(1)	(2)	(3)
Panel A. Information proxy: Pre-hire employment	*More informed*	*Less informed*	*Difference*
Realized sickness	0.1133***	0.1575***	0.1575***
	(0.0108)	(0.0392)	(0.0390)
Realized sickness * Pre-hire employment			-0.0441
			(0.0404)
Number of observations	55,268	8,356	63,624
Mean of dependent variable	0.269	0.352	0.280
Panel B. Information proxy: Firm connection	*More informed*	*Less informed*	*Difference*
Realized sickness	0.0507***	0.1284***	0.1284***
	(0.0123)	(0.0097)	(0.0097)
Realized sickness * Firm connection			-0.0777***
			(0.0157)
Number of observations	31,612	63,624	95,236
Mean of dependent variable	0.267	0.280	0.276
Panel C. Information proxy: Network connection	*More informed*	*Less informed*	*Difference*
Realized sickness	0.0837***	0.1318***	0.1318***
	(0.0249)	(0.0113)	(0.0113)
Realized sickness * Network connection			-0.0481*
			(0.0273)
Number of observations	13,318	50,306	63,624
Mean of dependent variable	0.231	0.293	0.280
Background controls	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Occupation fixed effects	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes
Worker fixed effects	No	No	No

Notes: The standard errors are clustered on the workplace level. The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children and establishment size. In column (3) all variables are interacted with Pre-hire employment/Firm connection/Network connection. In Panel B, we relax the new hire definition and include new hires on the workplace with a history within the firm, which explains why the sample size is larger than in Panels A and C. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

5 Conclusions

We examine if workers' sickness absence predicts sorting over jobs, wages and retention rates. The key idea is that firms should be sensitive to employee absence in jobs with few internal substitutes, and therefore reluctant to hire workers with high absence probabilities to these jobs. Based on this notion we outline a stylized search and matching model (building on Eeckhout and Kircher [2011]) for the matching between workers and jobs with the main assumption that matches between workers with high proneness to sickness absence and jobs with few substitutes generate particularly bad pay-offs. We show that our data generally are consistent with the predictions of the model:

First, employees performing a particular task (as captured by occupational classifications) are less absent if they perform this task in an environment where they have few substitutes. This finding is consistent with the prediction that there should be sorting of low-absence type workers into jobs with few substitutes, but it is not sufficient to prove a selection effect since the cross-sectional association also can be driven by endogenous adjustments of the sickness absence once in a position with few substitutes. By contrasting the estimates without and with individual fixed effects we can, however, get additional information about the relative importance of selection. This exercise reveals that about half of the predictive effect survives after the inclusion of individual fixed effects indicating that selection and endogenous adjustments are of equal importance for the observed relationship. The fact that we find a significant negative relation between having few substitutes and sickness absence also in a model with individual fixed effects is a clearly interesting venue for future research.

Our second set of results pertains to further investigations of the selection mechanism using a sample of new hires. With this data we can perform more direct tests of the predictions from the theoretical model. We find that new hires in jobs with few substitutes have significantly lower pre-hire sickness absence, conditional on establishment and occupational fixed effects. Workers in unique positions also display higher turnover rates from realizations of absence. Together these two results suggest that both ex ante sorting and ex post separations are means to achieve low absence in key positions within the firm.

Additionally, we find suggestive evidence that the sorting pattern is more pronounced among matches formed under relatively more precise information about worker-absence-type, where the precision is proxied by strong pre-hire employment, previous connection to the recruiting firm and coworker network connection to the recruiting establishment. We further show that these information proxies are better predictors of pre-hire sickness absence among matches involving jobs with few substitutes suggesting, as the theoretical model predicts, that screening for low-absence types is more pronounced for positions with low internal substitutability. Finally, we find that the separation response due to realized sickness absence among workers in jobs with few substitutes is negatively related to the amount of information in the hiring stage, suggesting that learning about match quality is an important determinant of turnover rates as in Jovanovic (1979b).

Overall these results suggest that sickness absence is a key factor in the sorting process as employers seek low-absence workers for jobs where absence is associated with production disruption costs. However, the inability to perfectly observe the absence propensity of prospective employees leads to mismatch between workers and jobs and in turn job separations. The findings thus validate previous theoretical and empirical work on the importance of sorting and point at sickness absence as a previously unexplored dimension of match quality.

References

Abowd, J.M., Haltiwanger, J., Lane, J., McKinsey, K.L. and Sandusky, K. (2007), "Technology and the demand for skill: an analysis of within and between firm differences", NBER Working Paper 13043.

Allen, S. (1981), "Compensation, safety, and absenteeism: Evidence from the paper industry", *Industrial & Labor Relations Review*, 34(2), 207–218.

Altonji, J. and Pierret, C. (2001), "Employer Learning and Statistical Discrimination", *The Quarterly Journal of Economics*, 116(1), 313–350.

Andersson, F., Freedman, M., Haltiwanger, J., Lane, J. and Shaw, K. (2009), "Reaching for the Stars: Who Pays for Talent in Innovative Industries?", *The Economic Journal*, 119(538), F308–F332.

Barmby, T. and Stephan, G. (2000), "Worker absenteeism: why firm size may matter", *The Manchester School*, 68(5), 568–577.

Black, C. and Frost, D. (2011), "Health at work – and independent review of sickness absence", Presented to the UK parliament in November, 2011.

Coles, M. and Treble, J. (1996), "Calculating the price of worker reliability", *Labour Economics*, 3(2), 169–188.

Dionne, G. and Dostie, B. (2007), "New evidence on the determinants of absenteeism using linked employer-employee data", *Industrial & Labor Relations Review*, 61(1), 108–120.

Eeckhout, J. and Kircher, P. (2011), "Identifying sorting—in theory", *The Review of Economic Studies*, 78(3), 872–906.

Eriksson, S., Johansson, P. and Langenskiöld, S. (2012), ”What is the right profile for getting a job? A stated choice experiment of the recruitment process", IFAU Working Paper 2012:13.

Eriksson, S. and Rooth, D. (2014), "Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment", *The American Economic Review*, 104(3), 1014–1039.

Farber, H. and Gibbons, R. (1996), “Learning and Wage Dynamics”, *The Quarterly Journal of Economics*, 111(4), 1007–1047.

Fredriksson, P., Hensvik, L. and Skans, O.N. (2015), “Mismatch of Talent: Evidence on Match Quality, Job Mobility and Entry Wages”, mimeo.

German Federal Statistical Office. (2011), “Statistical Yearbook 2011 for the Federal Republic of Germany”, Stuttgart: Metzler-Poeschel.

Hensvik, L. and Skans, O.N. (2015), “Social Networks, Employee Selection and Labor Market Outcomes”, *Journal of Labor Economics*, forthcoming.

Ichino, A and Moretti E. (2009), “Biological Gender Differences, Absenteeism and the Earnings Gap”, *American Economic Journal*: Applied Economics", 2009, 1:1, 183–218.

Jackson, K. (2013), “Match quality, worker productivity, and worker mobility: Direct evidence from teachers”, *Review of Economics and Statistics*, 95(4), 1096–1116

Jovanovic, B. (1979), “Job matching, and the theory of turnover“, *Journal of Political Economy*, 87, 972–990.

Kremer, M. (1993), “The O-ring theory of economic development”, *The Quarterly Journal of Economics*, 108(3), 551–575.

Kwon, I. (2005), “Threat of dismissal: incentive or sorting?”, *Journal of Labor Economics*, 23(4), 797–838.

Lange, F. (2007), “The speed of employer learning”, *Journal of Labor Economics*, 25(1), 1–35.

Lazear, E. (1998), “Hiring Risky Workers”, Internal Labour Market, Incentives, and Employment, eds. Isao Ohashi and Toshiaki Tachibanki. New York: St. Martin’s Press.

Lazear, E. (2009), “Firm-specific human capital: a skill-weights approach”, *Journal of Political Economy*, 117(5), 914–940.

Lazear, E. and Oyer, P. (2012), “Personnel Economics”, Handbook of Organizational Economics, 479–519.

Lindgren, K. (2012), “Workplace size and sickness absence transitions”, IFAU Working Paper 2012:26.

Nagypál, É. (2007), “Learning by doing vs. learning about match quality: Can we tell them apart?”, *The Review of Economic Studies*, 74(2), 537–566.

Office for National Statistics (UK). (2014), “Full report: Sickness Absence in the Labour Market, February 2014”.

Ose, S. (2005), “Working conditions, compensation and absenteeism”, *Journal of Health Economics*, 24(1), 161–188.

Oyer, P. and Schaefer, S. (2011), “Personnel Economics: Hiring and Incentives”, Handbook of Labor Economics, vol. 4, 1769–1823.

Previa (private company in the Swedish health industry). (2013), “Arbetsgivarnas kostnader för sjukfrånvaro ökade med en miljard”, Press release, April 23, 2013.

Sattinger, M. (1975), “Comparative advantage and the distribution of earnings and abilities“, *Econometrica*, 43, 455–468.

Schönberg, U. (2007), “Testing for asymmetric employer learning”, *Journal of Labor Economics*, 25(4), 651–691.

Statistics Sweden. (2007), “Labour income and public transfers in the Nordic countries in the 1990’s”, Economic Welfare 2007:1.

Tinbergen, J. (1956), “On the theory of income distribution”, *Weltwirtschaftliches Archiv*, 77, 156–173.

Weiss, A. (1985), “Absenteeism and wages”, *Economics Letters*, 19(3), 277–279.

Appendix: Descriptive statistics and additional results

Table A1 Descriptive statistics for all employees by degree of uniqueness

	All	0–5 sub	5+ sub
Workplace characteristics			
Unique position	0.036	0.204	0.000
Establishment size	570.9	90.8	675.7
Demographics			
Age	40.6	42.6	40.2
Male	0.629	0.495	0.658
Number of children age 0–17	0.829	0.836	0.828
Country of origin			
Sweden	0.913	0.939	0.908
Rest of Nordic countries	0.032	0.026	0.033
Rest of Europe	0.026	0.018	0.027
North America	0.001	0.001	0.001
South America	0.006	0.003	0.007
Rest of the world	0.022	0.012	0.024
Education			
Pre high school education (< 9 years)	0.025	0.021	0.026
Pre high school education ((≥ 9 years)	0.054	0.042	0.057
High school education max 2 years	0.413	0.386	0.418
High school education 2–3 years	0.225	0.233	0.223
Post high school education (< 3 years)	0.143	0.169	0.138
Post high school education (≥ 3 years)	0.131	0.143	0.128
Postgraduate education	0.008	0.004	0.008
Wage and Benefits			
Monthly wage in *t* (SEK)	23,657	22,824	23,839
Sickness benefit recipient in *t*	0.109	0.098	0.111
Professions			
Professionals	0.165	0.191	0.159
Technicians	0.245	0.315	0.229
Clerks	0.124	0.212	0.105
Service workers and shop sales	0.088	0.093	0.087
Skilled agricultural and fishery	0.005	0.010	0.004
Craft and related trades workers	0.116	0.090	0.121
Plant and machine operators	0.196	0.043	0.230
Elementary occupations	0.062	0.047	0.065
Number of observations	5,863,497	1,050,017	4,813,480

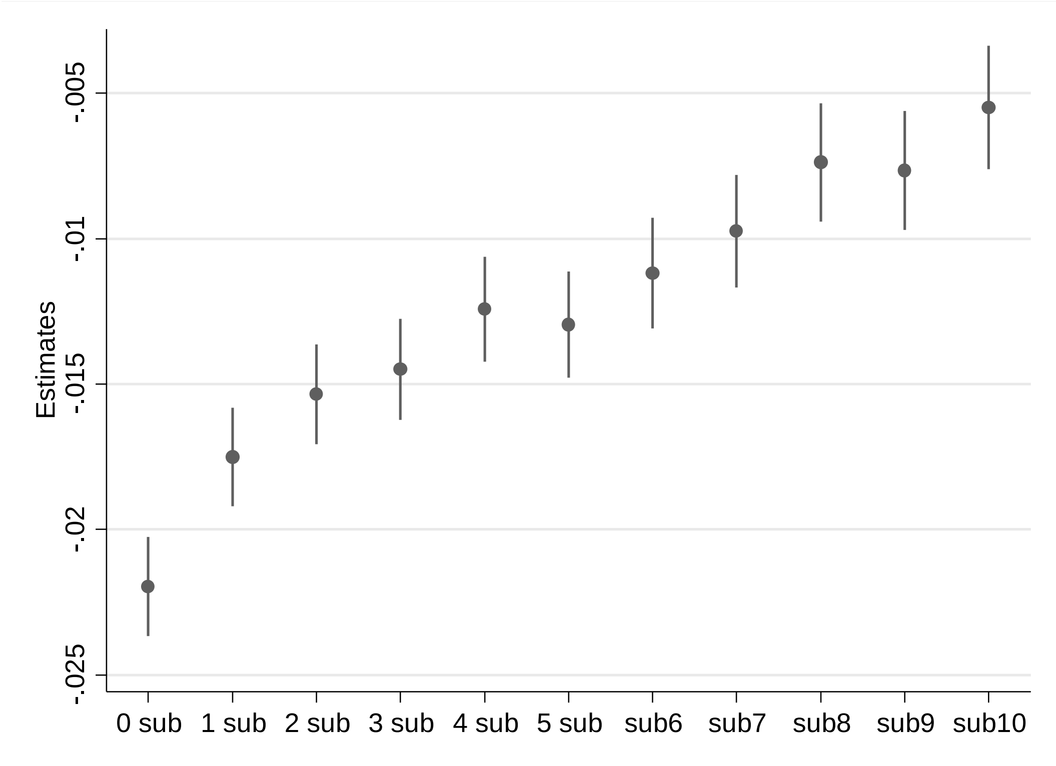
Notes: The sample is based on private sector employees in Sweden in 1997–2007. Managers and labor market entrants are excluded. The distribution across occupations is reported at the 1-digit level of the occupation code.

Table A2 Descriptive statistics for new hires by degree of uniqueness

	All	0–5 sub	5+ sub
Workplace characteristics			
Unique position	0.039	0.207	0.000
Establishment size	406.8	77.0	483.7
Demographics			
Age	35.7	37.4	35.3
Male	0.599	0.501	0.622
Number of children age 0–17	0.807	0.890	0.787
Country of origin:			
Sweden	0.898	0.928	0.891
Rest of Nordic countries	0.024	0.023	0.025
Rest of Europe	0.030	0.021	0.032
North America	0.001	0.001	0.001
South America	0.010	0.006	0.011
Rest of the world	0.065	0.041	0.071
Education			
Pre high school education (< 9 years)	0.008	0.006	0.008
Pre high school education ((≥ 9 years)	0.089	0.059	0.097
High school education max 2 years	0.343	0.332	0.346
High school education 2–3 years	0.270	0.270	0.270
Post high school education (< 3 years)	0.133	0.158	0.127
Post high school education (≥ 3 years)	0.150	0.172	0.145
Postgraduate education	0.006	0.004	0.006
Wages and Benefits			
Monthly wage in *t* (SEK)	22,413	22,226	22,457
Pre-hire sickness benefit recipient	0.116	0.125	0.114
Professions			
Professionals	0.171	0.186	0.168
Technicians	0.227	0.293	0.212
Clerks	0.128	0.207	0.110
Service workers and shop sales	0.133	0.113	0.137
Skilled agricultural and fishery	0.006	0.010	0.004
Craft and related trades workers	0.091	0.082	0.094
Plant and machine operators	0.165	0.052	0.192
Elementary occupations	0.079	0.058	0.084
Number of observations	387,901	73,366	314,535

Notes: The sample is based on private sector new hires in Sweden in 1997–2007. Managers and labor market entrants are excluded. The distribution across occupations is reported at the 1-digit level of the occupation code.

Figure A1 Present sickness absence and the number of coworker substitutes



Notes: The standard errors are clustered on the workplace level. The reference category is employees with more than 10 substitutes. The background controls are gender, age, education, birth country, having small children and establishment size. The model also includes year fixed effects, occupational fixed effects and workplace fixed effects.

Table A3 Present sickness absence and the number of coworker substitutes – workplace size heterogeneity

Outcome: Sickness absence in *t*	(1)	(2)	(3)
Panel A: Small workplaces			
0–1 substitutes	-0.0146***	-0.0157***	-0.0109***
	(0.0010)	(0.0010)	(0.0015)
2–5 substitutes	-0.0093***	-0.0088***	-0.0043***
	(0.0009)	(0.0009)	(0.0011)
Number of observations	1,890,998	1,890,998	1,890,998
Mean of dependent variable	0.111	0.111	0.111
Panel B: Medium-sized workplaces			
0–1 substitutes	-0.0142***	-0.0187***	-0.0067***
	(0.0013)	(0.0013)	(0.0019)
2–5 substitutes	-0.0092***	-0.0116***	-0.0051***
	(0.0010)	(0.0010)	(0.0014)
Number of observations	2,015,934	2,015,934	2,015,934
Mean of dependent variable	0.115	0.115	0.115
Panel C: Large workplaces			
0–1 substitutes	-0.0060**	-0.0130***	-0.0042
	(0.0029)	(0.0028)	(0.0036)
2–5 substitutes	-0.0106***	-0.0151***	-0.0062***
	(0.0020)	(0.0019)	(0.0022)
Number of observations	1,937,411	1,937,411	1,937,411
Mean of dependent variable	0.100	0.100	0.100
Background controls	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Occupation fixed effects	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes
Worker fixed effects	No	No	Yes

Notes: The standard errors are clustered on the workplace level (on the worker level in column [3]). The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children and establishment size. Small workplaces have 3–59 employees. Medium-sized workplaces have 60–339 employees. Large workplaces have 340–6815 employees. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

Table A4 Present wage and the number of coworker substitutes

	(1)	(2)	(3)
Log of monthly wage in (t)			
0–1 substitutes	0.0194***	0.0161***	0.0030***
	(0.0014)	(0.0012)	(0.0004)
2–5 substitutes	0.0101***	0.0079***	0.0007**
	(0.0011)	(0.0010)	(0.0003)
Number of observations	5,863,497	5,863,497	5,863,497
Mean of dependent variable	10.01	10.01	10.01
Background controls	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Occupation fixed effects	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes
Worker fixed effects	No	No	Yes

Notes: The standard errors are clustered on the workplace level (on the worker level in column [3]). The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children and establishment size. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

Table A5 Present sickness absence and the number of coworker substitutes – robustness checks

	(1)	(2)
Panel A. **Dep. var: Sickness absence in *t***	*With wage control*	
0–1 substitutes	-0.0169***	-0.0089***
	(0.0008)	(0.0010)
2–5 substitutes	-0.0106***	-0.0048***
	(0.0006)	(0.0007)
Number of observations	5,863,497	5,863,497
Mean of dependent variable	0.109	0.109
Panel B. **Dep. var: Sickness absence in *t***	*Including public sector*	
0–1 substitutes	-0.0113***	-0.0090***
	(0.0006)	(0.0007)
2–5 substitutes	-0.0060***	-0.0036***
	(0.0005)	(0.0005)
Number of observations	12,160,539	12,160,539
Mean of dependent variable	0.125	0.125
Panel C. **Dep. var: Sickness absence in *t***	*4-digit occupational code*	
0–1 substitutes	-0.0158***	
	(0.0013)	
2–5 substitutes	-0.0081***	
	(0.0010)	
Number of observations	1,656,960	
Mean of dependent variable	0.105	
Background controls	Yes	Yes
Year fixed effects	Yes	Yes
Occupation fixed effects	Yes	Yes
Workplace fixed effects	Yes	Yes
Worker fixed effects	No	Yes

Notes: The standard errors are clustered on the workplace level (on the worker level in column [2]). The reference category is employees with more than 5 substitutes. The background controls are gender, age, education, birth country, having small children and establishment size. In Panel A, wage is included in the model. In Panel B we include public sector employees. In Panel C we calculate the number of substitutes based on a 4-digit occupational code which is available for the years 2005–2007. The occupational fixed effects are also based on the 4-digit occupational code. *, **, and *** denote statistical significance at the 10-, 5-, and 1-percent level.

Publication series published by IFAU – latest issues

Rapporter/Reports

2015:1 Albrecht James, Peter Skogman Thoursie and Susan Vroman ”Glastaket och föräldraförsäkringen i Sverige”

2015:2 Persson Petra ” Socialförsäkringar och äktenskapsbeslut”

2015:3 Frostenson Magnus ”Organisatoriska åtgärder på skolnivå till följd av lärarlegitimationsreformen”

2015:4 Grönqvist Erik and Erik Lindqvist ”Kan man lära sig ledarskap? Befälsutbildning under värnplikten och utfall på arbetsmarknaden”

2015:5 Böhlmark Anders, Helena Holmlund and Mikael Lindahl ”Skolsegregation och skolval”

2015:6 Håkanson Christina, Erik Lindqvist and Jonas Vlachos ”Sortering av arbetskraftens förmågor i Sverige 1986–2008”

2015:7 Wahlström Ninni and Daniel Sundberg ” En teoribaserad utvärdering av läroplanen Lgr 11”

2015:8 Björvang Carl and Katarina Galic´ ”Kommunernas styrning av skolan – skolplaner under 20 år”

2015:9 Nybom Martin and Jan Stuhler ”Att skatta intergenerationella inkomstsamband: en jämförelse av de vanligaste måtten”

2015:10 Eriksson Stefan and Karolina Stadin ”Hur påverkar förändringar i produktefterfrågan, arbetsutbud och lönekostnader antalet nyanställningar?”

2015:11 Grönqvist Hans, Caroline Hall, Jonas Vlachos and Olof Åslund ”Utbildning och brottslighet – vad hände när man förlängde yrkesutbildningarna på gymnasiet?”

2015:12 Lind Patrik and Alexander Westerberg ”Yrkeshögskolan – vilka söker, vem tar examen och hur går det sedan?

2015:13 Mörk Eva, Anna Sjögren and Helena Svaleryd ”Hellre rik och frisk – om familjebakgrund och barns hälsa”

2015:14 Eliason Marcus and Martin Nilsson ”Inlåsningseffekter och differentierade ersättningsnivåer i sjukförsäkringen”

2015:15 Boye Katarina ”Mer vab, lägre lön? Uttag av tillfällig föräldrapenning för vård av barn och lön bland svenska föräldrar”

2015:16 Öhman Mattias ”Smarta och sociala lever längre: sambanden mellan intelligens, social förmåga och mortalitet”

2015:17 Mellander Erik and Joakim Svärdh ”Tre lärdomar från en effektutvärdering av lärarstödsprogrammet NTA”

2015:18 Regnér Johan and Martin Söderström ”Tidiga insatser för arbetssökande med hög risk för långtidsarbetslöshet”

2015:19 Blind Ina, Matz Dahlberg and Olof Åslund ”Som på räls? Förbättrade pendlingsmöjligheter och individens arbetsmarknad”

2015:20 Fredriksson Peter, Lena Hensvik and Oskar Nordström Skans ” Rätt man på rätt plats? Matchning, ingångslöner och separationer”

2015:21 Fredriksson Peter, Hessel Oosterbeek and Björn Öckert ”Hur reagerar föräldrar på resursneddragningar i skolan?”

2015:22 Hensvik Lena and Olof Rosenqvist ”Sjukfrånvaro och anställningar till unika positioner”

Working papers

2015:1 Avdic Daniel “A matter of life and death? Hospital distance and quality of care: evidence from emergency hospital closures and myocardial infarctions”

2015:2 Eliason Marcus “Alcohol-related morbidity and mortality following involuntary job loss”

2015:3 Pingel Ronnie and Ingeborg Waernbaum “Correlation and efficiency of propensity score-based estimators for average causal effects”

2015:4 Albrecht James, Peter Skogman Thoursie and Susan Vroman “Parental leave and the glass ceiling in Sweden”

2015:5 Vikström Johan “Evaluation of sequences of treatments with application to active labor market policies”

2015:6 Persson Petra “Social insurance and the marriage market”

2015:7 Grönqvist Erik and Erik Lindqvist “The making of a manager: evidence from military officer training”

2015:8 Böhlmark Anders, Helena Holmlund and Mikael Lindahl “School choice and segregation: evidence from Sweden”

2015:9 Håkanson Christina, Erik Lindqvist and Jonas Vlachos “Firms and skills: the evolution of worker sorting”

2015:10 van den Berg Gerard J., Antoine Bozio and Mónica Costa Dias “Policy discontinuity and duration outcomes”

2015:11 Wahlström Ninni and Daniel Sundberg “Theory-based evaluation of the curriculum Lgr 11”

2015:12 Frölich Markus and Martin Huber “Direct and indirect treatment effects: causal chains and mediation analysis with instrumental variables”

2015:13 Nybom Martin and Jan Stuhler “Biases in standard measures of intergenerational income dependence”

2015:14 Eriksson Stefan and Karolina Stadin “What are the determinants of hiring? The role of demand and supply factors”

2015:15 Åslund Olof, Hans Grönqvist, Caroline Hall and Jonas Vlachos “Education and criminal behaviour: insights from an expansion of upper secondary school”

2015:16 van den Berg Gerard J. and Bas van der Klaauw “Structural empirical evaluation of job search monitoring”

2015:17 Nilsson Martin “Economic incentives and long-term sickness absence: the indirect effect of replacement rates on absence behavior”

2015:18 Boye Katarina “Care more, earn less? The association between care leave for sick children and wage among Swedish parents”

2015:19 Assadi Anahita and Martin Lundin “Tenure and street level bureaucrats: how assessment tools are used at the frontline of the public sector”

2015:20 Stadin Karolina “Firms’ employment dynamics and the state of the labor market”

2015:21 Öhman Mattias “Be smart, live long: the relationship between cognitive and non-cognitive abilities and mortality”

2015:22 Hägglund Pathric, Per Johansson and Lisa Laun “Rehabilitation of mental illness and chronic pain – the impact on sick leave and health”

2015:23 Mellander Erik and Joakim Svärdh “Inquiry-based learning put to test: long-term effects of the Swedish science and technology for children program”

2015:24 Norén Anna “Childcare and the division of parental leave”

2015:25 Åslund Olof, Ina Blind and Matz Dahlberg “All aboard? Commuter train access and labor market outcomes”

2015:26 Fredriksson Peter, Lena Hensvik and Oskar Nordström Skans ”Mismatch of talent: evidence on match quality, entry wages, and job mobility”

2015:27 Fredriksson Peter, Hessel Oosterbeek and Björn Öckert ”Parental responses to public investments in children: evidence from a maximum class size rule”

2015:28 Hensvik Lena and Olof Rosenqvist "The strength of the weakest link: sickness absence, internal substitutability and worker-firm matching"

Dissertation series

2014:1 Avdic Daniel "Microeconometric analyses of individual behaviour in public welfare systems"

2014:2 Karimi Arizo "Impacts of policies, peers and parenthood on labor market outcomes"

2014:3 Eliasson Tove "Empirical essays on wage setting and immigrant labor market opportunities"

2014:4 Nilsson Martin "Essays on health shocks and social insurance"

2014:5 Pingel Ronnie "Some aspects of propensity score-based estimators for causal inference"

2014:6 Karbownik Krzysztof "Essays in education and family economics"