Coworker complementarity

Frank M. H. Neffke*

February 27, 2017

Download the most recent version here

Abstract

How important is working with people who complement one's skills? Using administrative data that record which of 491 educational tracks each worker in Sweden absolved, I quantify the educational fit among coworkers along two dimensions: coworker match and coworker substitutability. Complementary coworkers raise wages with a comparable factor as does a college degree, whereas working with close substitutes is associated with wage penalties. Moreover, this coworker fit does not only account for large portions of the urban and large-plant wage premiums, but the returns to own schooling and the urban wage premium are almost completely contingent on finding complementary coworkers.

Keywords

complementarity, substitutability, job match, coworkers, skills

1 Introduction

Division of labor allows a society to reap the benefits of specialization. An obvious, yet somewhat underappreciated aspect of this statement is that division of labor does not just require that workers specialize, but that they specialize in *different things*. When know-how takes the shape of such distributed expertise,

*Harvard Kennedy School, 79 J.F.K. Street, Cambridge, MA 02138, USA. and Lund University, Department of Human Geography, frank_neffke@hks.harvard.edu. I would like to thank Alberto Abadie, Thomas Åstebro, David Autor, Richard Freeman, Andres Gomez, Matté Hartog, William Kerr, Ljubica Nedelkoska, Lant Pritchett, Muhammed Yildirim, the participants of the CID Faculty Luncheon on Economic Policy, the Geography of Innovation Conference in Toulouse, the NBER productivity lunch, the URU seminar in Utrecht, the SKEMA workshop on Spatial evolution of industries, jobs and innovative activities in Nice, the Eurkind GCW 2016 Conference in Valencia, and my colleagues at the CID Growth Lab for their advice and comments. I am furthermore indebted to the SCB MONA-support team for their excellent technical assistance. Finally, I am extremely grateful to Ricardo Hausmann and Martin Henning for their continuous advice and support throughout this project.

the value of human capital will depend on the knowledge ecosystem in which it is embedded. That is, workers' productivity will depend on whether or not they have access to coworkers with skills and know-how that complement their own. In this paper, I quantify the fit between a worker's and her coworkers' skill sets and show how this fit affects wages and career paths, as well as how it explains and moderates various well-known wage premiums.

A number of papers has shown that a worker's productivity depends on her coworkers. For instance, Mas and Moretti (2009) find significant productivity spillovers among supermarket cashiers. Similarly, Arcidiacono et al. (2013) show that an important component of a professional basketball player's value goes unnoticed when spillovers to teammates are ignored. Moreover, also Card et al.'s (2013) finding that about a third of the recent increase in West German wage dispersion can be attributed to stronger assortative matching among workers indicates that strong coworker interdependencies exist.

Coworker interdependencies are also acknowledged in theoretical work, giving rise, for instance, to the notion of "team human capital" in Chillemi and Gui (1997) and to the idea that much firm-specific human capital relates, not to a firm, but to the "network of workers" (Mailath and Postlewaite, 1990) that constitutes a firm's labor force. However, whereas the empirical literature tends to focus on spillovers and theory papers often revolve around wage bargaining, neither of these approaches addresses the interdependence that arises with the complementary specialization that occurs when different bits of know-how are held by different people. In this paper, I focus on the latter phenomenon by empirically assessing the fit among coworkers' skill sets in terms of their complementarity and substitutability.

To do so, I focus on the skills workers acquire through education in Sweden, recorded in 491 detailed educational tracks that describe the content and level of education for each individual in Sweden between 1990 and 2010. I use these data to determine which educational tracks are substitutes and which are complements to one another. To quantify substitutability, I assess which educational tracks give access to the same occupations. To quantify complementarity, I determine which educational tracks often co-occur in the workforces of economic establishments, assuming that, on average (and controlling for the substitutability among coworkers), firms hire teams in which workers complement each other. The result is a 491×491 matrix that describes how *substitutable* two educational tracks are, and a similar matrix that portrays the prevalence of educational co-occurrences in economic establishments, to which I will refer as the *coworker match* of two educations. Conditional on how substitutable a worker is within her team, this coworker match turns out to behave as an indicator of the worker's complementarity to her coworkers. I construct these two matrices using a *measurement sample* that contains 75% of the Swedish working and use them in a separate *estimation sample* to estimate the effects of the (two-dimensional) fit of a worker to her coworkers on wages, careers and wage premiums.

Improving the educational fit with coworkers proves to be important in a worker's career: Ordinary Least Squares (OLS) regressions suggest that in-

creases in average coworker match are associated with substantial increases in wages. The effects are strongest for college-educated workers: the premiums for working with complementary coworkers they receive are similar in magnitude to the returns to college education itself. However, because the coworker match is measured with error, these OLS estimates are still substantially downward biased. In fact, multiple error-correction approaches indicate that the true effect of coworker match is about twice as high as what OLS regressions suggest. These error-corrected effects are remarkably close to causal estimates obtained from instrumental variable (IV) models that exploit exogenous shifts in the local supply of graduates.

In contrast to the positive effect of having complementary coworkers, working with close substitutes is associated with lower wages. Moreover, workers who are easily substituted by their coworkers switch jobs sooner than those who don't, whereas having more complementary coworkers increases the likelihood of long tenures. As a consequence, the observed coworker fit increases for over 20 years into a worker's career and it does so along a concave curve that closely tracks the evolution of the Mincer residual.

The educational fit with coworkers also helps explain a number of well-known wage premiums. For instance, the two coworker-fit variables explain between 30% and 34% of the urban wage premium for college-educated workers, and between 50% and 75% for workers with post-graduate degrees. Similarly, for workers with post-secondary degrees or higher, the entire large-plant premium can be accounted for by the fact that large establishments employ many complementary (yet relatively few substitutable) coworkers. This suggests that these premiums exist, because large cities and large establishments facilitate the formation of teams of highly complementary coworkers. Moreover, some of these premiums only materialize in the presence of a high coworker-fit. For instance, the observed college-to-primary-school wage premium varies from a low 18% in the bottom quintile to a high 89% in the top quintile of coworker complementarity. Similarly, the wage elasticity with respect to city-size varies from below 1% for workers in the bottom complementarity quintile to above 9% for workers in the top quintile.

The findings in this paper contribute to a number of debates. First, they shed light on how the education of coworkers affects a worker's own wage through another channel than common peer effects. In particular, whereas the peer-effects literature studies the value of working with (or being surrounded by) *highly educated* or *highly productive* coworkers, this paper focuses on the extent to which coworkers' educational specializations *complement* a worker's own specialization. Consequently, this study does not suffer from the identification problems that have plagued the estimation of peer effects (e.g., Acemoglu and Angrist, 2001; Angrist, 2014), although other endogeneity concerns (related to sample-selection and capability-based sorting) do arise and are dealt with.

Second, the paper contributes to the literature on returns to education. Recent work in this area shows that returns to tertiary education differ widely by field of study. For instance, Arcidiacono (2004) finds large differences in earnings premiums across college majors in the U.S., even after controlling for

ability-based sorting. Similarly, Kirkebøen et al. (2014) find that the variation in premiums to different fields of post-secondary study is on par with the average returns to college education. The present paper shows that such returns may only fully materialize when teammates possess complementary skills. This also provides a new perspective on social returns to education. So far, evidence for such educational externalities has been mixed. For instance, exploiting exogenous variation in compulsory schooling laws, Acemoglu and Angrist (2001) don't find that a high average education yields significant spillovers in U.S. states, whereas Moretti (2004) finds substantial spillovers from a large prevalence of college graduates in U.S. cities. The present study suggests that social returns to schooling do not just depend on the average *level* of education in an economy, but rather on the available *mix* of educational specializations and that these benefits accrue only to workers whose own education fits this educational mix well.

Third, by quantifying a worker's educational fit with her coworkers, the paper contributes to our understanding of worker-job matches. Because teams change, this educational fit is time-varying. Hence, it can be used to assess how job matches evolve - highlighting, for instance, the importance of early-career job-switching in finding the right team of coworkers. Moreover, the focus on coworker relations is close in spirit, yet not substance, to recent work by Jäger (2016), who uses unexpected deaths of coworkers to assess how substitutable coworkers are to one another. Whereas Jäger focuses on the question of whether coworkers are *in general* substitutes or complements, I use an *a priori* estimate of complementarity and substitutability to estimate the impact of coworker fit on a worker's career.

Fourth, the paper contributes to debates on the urban and the large-plant wage-premiums. The fact that workers in larger cities earn higher wages has been attributed to greater learning opportunities in urban environments (e.g., Glaeser et al., 2001). However, the current paper suggests that large cities are attractive because they allow highly educated workers to find well-matching work environments (Helsley and Strange, 1990). Similarly, although a variety of mechanisms has been put forward to explain why larger establishments pay observationally equivalent workers higher wages (Troske, 1999), this paper proposes (and provides empirical evidence for) a new explanation for this phenomenon: large establishments can pay higher wages because they rely on a deeper division of labor that exploits coworker complementarities more fully.

The most important implication, however, is the fact that present-day workers' human capital is highly specific creates interdependencies among workers. These interdependencies may give rise to complex coordination problems. For instance, if returns to education depend on the availability of workers with complementary skill sets, such educational interdependencies can complicate the upgrading of human capital in an economy.

2 Coworker complementarities and wages

A central premise in Adam Smith's pin factory allegory is that specialization leads to productivity gains. Opportunities for division of labor are commonly viewed to be limited by the extent of the market. However, Becker and Murphy (1992) sketch a model of division of labor in which individuals are organized in teams. In this model, division of labor is not just limited by the size of the market, but also by coordination costs. Implicit in the model is that, given that humans only have a limited capacity to acquire knowledge and skills, depth must come at the expense of breadth. Consequently, as specialization increases, knowledge needs to be distributed across an expanding variety of experts. Such a process has much intuitive appeal. It is, for instance, reflected in the seemingly unending branching of fields of knowledge found at today's universities and research institutes. Moreover, it helps understand why occupational variety rises with city size (Bettencourt et al., 2014): the large local markets of big cities allow for a deep division of labor.

An interesting aspect of Becker and Murphy's model is that, to ensure that all tasks are carried out, specialization of one worker requires the specialization of other workers. In other words, workers' investments in task-specific human capital only pay off if they are matched by other workers' investing in *different, yet complementary* skills. Consequently, the division of labor leads to interdependencies among workers: expertise in producing pinheads is of little use if no one else knows how to make pins.

Becker and Murphy (1992) capture these interdependencies among coworkers in a Leontief production function that converts a continuum of tasks into output. Workers choose their specialization taking the specializations of other workers into consideration. Consequently, there are an infinite number of equilibria, each of which describe a different way of dividing tasks among workers. However, what would happen if workers had to choose their specialization before they know whom they will work with?[1]

To simplify issues, consider an economy with two types of workers who produce a continuum of tasks on the unit interval. Each worker type produces all tasks, but workers of type 1 are best at producing tasks close to 0 and workers of type 2 are best at producing tasks close to 1. In particular, an i-type worker's output of task $t \in [0,1]$, $T_i(t)$, is:

$$\begin{aligned} T_1(t) &= t + a \\ T_2(t) &= 1 - t + a \end{aligned}$$

[1] In reality, skills and knowledge are bundled into pre-existing educational tracks. This standardization of expertise alleviates coordination problems: schools bundle skills and firms combine this expertise, as it were, off-the-shelve. Standardization also prepares students for future coordination of their skills with those of others. That is, education does not merely (and maybe not even primarily) teach how to carry out a set of actions, but also helps recognize which of these actions are required in the light of the actions of others. For instance, a designer needs to understand the marketing team's sales plan at the same time that she needs to understand how the printer will use her designs. In part, this boils down to learning a jargon, but it also involves interpreting other experts' outputs and aligning one's own efforts accordingly.

Workers produce in teams of two and, as a team, they produce output of value π, according to a Leontief production function:

$$\pi_{ij} = \min_{0 \leq t \leq 1} \left(T_i(t) + T_j(t) \right), \quad i, j \in \{1, 2\} \tag{1}$$

where i and j index the worker types in the team. The value produced by a team is therewith either $\pi_{11} = \pi_{22} = 2a$ or $\pi_{12} = \pi_{21} = 2a + 1$. Hence, the highest output is achieved when teams mix workers of opposite types.

Workers divide output in a bargaining process, equally splitting the surplus over each other's outside options, $\bar{w}_i$. Consequently, the wage of an i-type worker in a team consisting of a type i and a type j worker, $w_{ij}^{(i)}$, equals:

$$w_{ij}^{(i)} = \frac{\pi_{ij} - \bar{w}_i - \bar{w}_j}{2} + \bar{w}_i = \frac{\pi_{ij}}{2} + \frac{\bar{w}_i - \bar{w}_j}{2} \tag{2}$$

In a team with either two type 1 workers or two type 2 workers, each worker receives a wage of a:

$$w_{ii}^{(i)} = \frac{2a}{2} + \frac{\bar{w}_i - \bar{w}_i}{2} = a$$

However, in mixed teams ($i \neq j$), the wage of a worker of type i depends on her outside option:

$$w_{ij}^{(i)} = \frac{2a + 1}{2} + \frac{\bar{w}_i - \bar{w}_j}{2}$$

Workers can disband their team and search for new partners. In particular, after paying search costs of c, they can draw a new team member at random. The probability of teaming up with a worker of type i equals p_i. The value of the outside option for an i-type worker is now the expected wage after breaking up a team:

$$\bar{w}_i = E\left[w^{(i)}\right] - c = p_i a + (1 - p_i)\left(\frac{1}{2} + a + \frac{\bar{w}_i - \bar{w}_j}{2}\right) - c \tag{3}$$

Meanwhile, the outside option of a worker of type j is valued at:

$$\bar{w}_j = p_i\left(\frac{1}{2} + a + \frac{\bar{w}_j - \bar{w}_i}{2}\right) + (1 - p_i)\, a - c \tag{4}$$

After rearranging terms, subtracting (4) from (3) yields:

$$\frac{\bar{w}_i - \bar{w}_j}{2} = \frac{1}{2} - p_i \tag{5}$$

Substituting (5) into (2) yields:

$$w_{ii}^{(i)} = a$$

and, for $i \neq j$,

$$w_{ij}^{(i)} = a + 1 - p_i$$

Consequently, an i-type worker's wage depends on the type of her coworker. In particular, there is a premium to working with a worker of the opposite type:

$$w_{ij}^{(i)} - w_{ii}^{(i)} = a + 1 - p_i - a = 1 - p_i$$

The size of this premium depends, first, on the complementarities in teamwork - here normalized to be of size 1 - and, second, on the relative prevalence of types among rematching workers. In particular, the productivity premium tends to go predominantly to the least ubiquitous worker type. Note that the exact shape of the production function is unimportant for this result, as long as it exhibits complementarities that create the payoff structure associated with (1). Moreover, the rematching process does not need to be fully random, as long as rematching is costly. In fact, models that embed intrafirm wage bargaining processes in search-and-matching models generalize these predictions. For instance, Jäger (2016) - building on work by Stole and Zwiebel (1996a,b), De Fontenay and Gans (2003) and Cahuc et al. (2008) - shows that in the presence of search frictions and intrafirm bargaining, a worker's wage will depend positively on her complementarity and negatively on her substitutability to coworkers.[2] The intuition behind this is that workers have bargaining power, because they can force employers into a time-consuming hiring process by quitting their jobs. How much bargaining power a worker has depends on the relative abundance of her skills in the labor market, on the one hand, and on the drop in output that would result from her withdrawal from the team, on the other hand. The more complementary a worker is to her teammates, the more valuable her threat to quit is. In contrast, if a worker is easily substituted by her coworkers, her bargaining power is relatively low.[3]

3 Measurement

3.1 Data

The empirical analyses in this paper are based on employer-employee linked data derived from Sweden's official registries as provided by Statistics Sweden (SCB).[4] This dataset contains yearly observations on all individuals who are

[2] Complementarities among coworkers will also be reflected in wages if employers share some of their productivity gains for reasons of fairness as in Akerlof and Yellen (1990) or to prevent shirking Bulow and Summers (1986).

[3] Jäger (2016) studies coworker relations by asking whether coworkers are substitutes or complements to one another. Using German linked employer-employee data, the author investigates the impact of unexpected deaths of workers on their coworkers' wages. Jäger finds that workers in the same occupation as the deceased tend to experience wage gains (i.e., are substitutes to the deceased worker), but that deaths of high-skill workers and managers often lead to wage drops, indicating that these workers used to complement their colleagues. In what follows, I approach the issue of coworker interdependence differently, starting from *a priori* measures on the educational fit among coworkers and then determining how this fit affects wages.

[4] These data are described in detail in SCB (2011).

living or working in Sweden. In each year, it records a number of sociodemographic characteristics, such as gender, age, and municipality of residence, together with work-related variables, such as the individual's main establishment of work and annual wage-income. Unfortunately, there is no information on total hours worked. Consequently, it is impossible to distinguish differences in annual wages that reflect differences in hourly pay from those that reflect differences in hours worked, which complicates some of the analyses in this paper.[5] From 2001 on, the data also record the occupations for about 90% of the working population.[6] Furthermore, apart from individual-level characteristics, the data contain 5-digit industry and municipality codes for all work establishments. Finally, and, from this paper's perspective, most importantly, from 1990 to 2010, the dataset contains detailed information on individuals' highest absolved education.

This educational information consists of two components. The first component is an alphanumeric code that divides educations into 351 different fields, such as "344z: Accounting and taxation," "214a: Fashion design" or "725f: Radiology nursing program." The second component is a 3-digit code that distinguishes 49 different levels of education, such as "337: Vocationally oriented programme, three years" or "640: Doctoral program." However, because the ranking of educational levels at the 2-digit level is not unambiguous, I retain only the first digit of this code and distinguish among six levels: primary school, secondary school, upper secondary school, post-secondary education, tertiary (henceforth referred to as "college") education and post-graduate ("Ph. D.") education. Next, I combine this 1-digit educational level with the 4-digit educational field identifier to create 491 educational tracks. Because primary and secondary school levels lack variation in content, I focus on workers with at least upper secondary education. Moreover, because occupational information is missing before 2001, I further restrict the analysis to the period 2001 to 2010. Finally, I exclude workers without educational or establishment information, self-employed individuals, workers with annual wages below the subsistence income, workers with rare education codes[7] and workers employed through employment agencies.[8]

The construction of measures and the regression analyses are carried out on two distinct partitions of the data. The first partition, the *measurement sample*, contains a 75% random sample of workers between the ages of 18 and 65. I use this sample to measure two kinds of relations among educational tracks: their substitutability and something I will call their *coworker match*. The remaining 25% of the population forms the *estimation sample*, which is

[5] In particular, the analyses in section 4.3 require special care.

[6] Occupations are coded into one of 335 4-digit occupation classes of the Swedish occupational classification system, *Standard för svensk yrkesklassificering 1996*, which derives from the United Nations' ISCO-88 classification. Examples include "1231: Personnel management" and "7232: Aircraft mechanics."

[7] Rare educational tracks have fewer than 500 individuals a year in the measurement sample. Dropping workers with such education ensures that the coworker fit measures are based on a reasonably large number of individuals.

[8] For these workers, it is impossible to determine their actual team of coworkers.

used to estimate how these educational relations affect careers and wages. In this sample, the same restrictions hold as before, but, to reduce problems of nonrandom selection into the labor force, I restrict this sample further to male, private-sector employees of between 20 and 60 years old, while excluding workers below the 0.5^{th} and over the 99.5^{th} wage percentile. Finally, workers in single-employee establishments are dropped, because they don't have any coworkers.

3.2 Substitutability and coworker match of educational tracks

From an employer's perspective, two educational tracks are substitutes if workers with either education can carry out the same tasks. Because occupations are essentially bundles of tasks, I quantify the substitutability of two educational tracks as the correlation between their occupational employment vectors. That is, let E_{eo} be the number of workers with education e who work in occupation o. The substitutability of education e by e' is now defined as:

$$s_{ee'} = corr\left(E_{eo}, E_{e'o}\right) \tag{6}$$

where correlations are taken over all occupations o in the economy. Note that, according to this definition, an educational track is a perfect substitute to itself: $s_{ee} = 1$.

Educational tracks can also be regarded to be related if they are often combined in teams. For lack of information about actual teamwork, I will use establishments to delineate teams.[9] In other words, to quantify how often workers work together, I assess whether particular combinations of educational tracks are often overrepresented in establishments' workforces. Let E_{ep} be the number of workers with education e in establishment p. e is present in p, denoted by $P_{ep} = 1$, if the share of workers with education e in establishment p exceeds those workers' share in the economy as a whole:

$$P_{ep} = 1\left(\frac{E_{ep}/\sum_{e'} E_{e'p}}{\sum_{p'} E_{ep'}/\sum_{e''}\sum_{p''} E_{e''p''}}\right)$$

where $1\left(.\right)$ is an indicator function that evaluates to 1 if its argument is true and to 0 otherwise. Furthermore, educational tracks *co-occur* if they are *present* in the same establishment. The number of co-occurrences in an establishment rises in proportion to the square of the number of different educations among its workers. Because larger establishments typically hire a greater variety of workers, they would quickly dominate the co-occurrence count. To avoid this, I normalize co-occurrences such that each establishment contributes a total of one co-occurrence:

$$N_{ee'} = \sum_{p} \frac{P_{ep}P_{e'p}}{\sum_{e''}\sum_{e'''} P_{e''p}P_{e'''p}}$$

[9] Because most establishments are small, this is defensible as a first approximation.

$N_{ee'}$ thus captures how often educations e and e' co-occur. However, some educations are more prevalent than others. To correct for this, I calculate the expected value of $N_{ee'}$, taking the total number of co-occurrences that involve each educational track as given, while assuming that, otherwise, co-occurrences form at random:[10]

$$\hat{N}_{ee'} = \sum_{e'} N_{ee'} \frac{\sum_{e} N_{ee'}}{\sum_{e''} \sum_{e'''} N_{e''e'''}}$$

Normalizing observed with expected co-occurrences yields:

$$R_{ee'} = \frac{N_{ee'}}{\hat{N}_{ee'}}$$

This ratio has a strongly skewed distribution: for overrepresented educational combinations, $R_{ee'}$ ranges from 1 to infinity, whereas underrepresented combinations have $R_{ee'}$ values of between 0 and 1. Therefore, I transform $R_{ee'}$ as follows:

$$c_{ee'} = \frac{R_{ee'}}{R_{ee'} + 1} \tag{7}$$

to map $R_{ee'}$ symmetrically around 0.5 onto the interval $[0, 1)$. By lack of a better term, I will refer to $c_{ee'}$ as the *coworker match* between educations e and e', while reserving the term *coworker fit* for the variable-pair of coworker match and coworker substitutability. Furthermore, in analogy to educational substitutability, I impose that an educational track is a perfect match for itself: $c_{ee} := 1$.

Tables 1 and 2 show the educational pairs with the highest coworker match and substitutability. High coworker match values often coincide with high levels of substitutability. Apparently, educations that often co-occur in establishments also tend to give access to similar occupations. However, the overlap between coworker match and substitutability is far from perfect. For instance, Table 1 shows that, although workers with a background in agricultural management often work with workers who have a degree in agricultural science, they cannot easily substitute for one another.[11]

3.3 Worker-establishment aggregation

The estimates of $c_{ee'}$ and $s_{ee'}$ are collected in educational proximity matrices of dimension $T_e \times T_e$, where T_e represents the number of different educational tracks

[10] Similar quantities are used in trade (revealed comparative advantage), economic geography (location quotient) and computer science (lift).

[11] Moreover, the relation between the two variables is not symmetric: educations that give access to the same occupations also often co-occur in coworker teams, but the reverse does not necessarily hold. Indeed, whereas for educations to be substitutes, they must typically be taught at the same level, coworker relations can form more freely. A typical example is a person with an upper secondary degree in "724d: Program in dental nursing" who works with someone with a college degree in "724a: Dental surgery." These patterns, as well as additional examples are described in greater detail in Appendix A.1.

Table 1: Top 10 educational pairs: coworker match

rank	edu. (1)	edu. (2)	match	subst.
1	3: Other forestry	5: Programme in forest management and forest engineering	0.985	0.415
2	5: Agricultural and rural management	5: Agricultural sciences	0.982	0.467
3	5: Other town planning and architecture	5: Architecture	0.981	0.960
4	5: Tactical military	5: Professional officers	0.980	0.999
5	6: Physics	5: Physics	0.969	0.826
6	5: Professional officers	5: Programme for air transport	0.969	0.016
7	5: Tactical military	5: Programme for air transport	0.966	0.019
8	3: Other agriculture	5: Agricultural and rural management	0.966	0.504
9	6: Chemistry	5: Chemistry	0.965	0.795
10	5: Programme for water transport	5: Marine vehicle engineering and aircraft engineering	0.964	0.203

Top 10 pairs of educational tracks with the highest coworker match among educational tracks with at least 100 employees a year in the estimation sample. Numbers preceding an educational track's name represent educational levels: 1: primary school; 2: secondary school; 3: upper secondary school; 4: post-secondary education; 5: college; 6: Ph. D.. Column "match" reports the coworker match of the two educational tracks as measured in (7), column "subst." their substitutability as measured in (6).

Table 2: Top 10 educational pairs: substitutability

rank	edu. (1)	edu. (2)	match	subst.
1	5: Tactical military	5: Professional officers	0.980	0.999
2	5: Teacher, after-school activities	5: Teacher, pre-school	0.931	0.999
3	5: Teacher, history/social sciences/economics	5: Teacher, mathematics/computers/nat. science	0.947	0.998
4	3: Other nursing, medical	3: Other nursing, general nursing and health care	0.864	0.995
5	5: Computing, general	4: Computing, general	0.910	0.989
6	5: Applied systems science and software engineering	5: Computing, general	0.903	0.989
7	5: Applied systems science and software engineering	4: Computing, general	0.895	0.983
8	5: Other electronics, comp. eng. and automation	5: Engineering: electronics, comp. eng. and automation	0.916	0.980
9	3: Other nursing, general nursing and health care	3: General	0.649	0.979
10	4: Applied systems science and software engineering	4: Computing, general	0.909	0.973

Idem Table 1, showing top 10 most substitutable educational pairs.

in the economy. However, what needs to be quantified is the substitutability and coworker match of a worker to her teammates. This fit of a worker w to her coworkers can be assessed in a number of ways. One way is counting how many well-matched coworkers (or close substitutes) a worker has. Using a logarithmic transformation, a worker's coworker match and substitutability to her team then becomes:

$$C^{\#}_{e(w,t)p(w,t)} = log_{10}\left(\sum_{e'} E_{e'p(w,t)t} 1\left(c_{e(w,t)e'} > \zeta_c\right)\right) \tag{8}$$

$$S^{\#}_{e(w,t)p(w,t)} = log_{10}\left(\sum_{e'} E_{e'p(w,t)t} 1\left(s_{e(w,t)e'} > \zeta_s\right)\right) \tag{9}$$

where $e(w,t)$ is w's education, and $p(w,t)$ w's work establishment in year t.[12] Furthermore, E_{ept} represents the number of workers with education e in establishment p in year t and ζ_c and ζ_s are thresholds chosen such that, if coworker relations were to form at random, the probability of having a coworker who is well-matched or a close substitute would equal 1%.[13] Also note that because a worker is both a perfect substitute and perfectly matched to herself, $S^{\#}_{e(w,t)p(w,t)}$ and $C^{\#}_{e(w,t)p(w,t)}$ are strictly greater than zero, ensuring that the logarithms in (8) and (9) are always well-defined.

As alternative measures, I also assess w's employment-weighted average coworker match and substitutability to coworkers:[14]

$$C_{e(w,t)p(w,t)} = \sum_{e'} \frac{E_{e'p(w,t)t} - 1\left(e' = e(w,t)\right)}{\sum_{e''} E_{e''p(w,t)t} - 1} c_{ee'} \tag{10}$$

Similarly, w's substitutability to her coworkers is defined as:

$$S_{e(w,t)p(w,t)} = \sum_{e'} \frac{E_{e'p(w,t)t} - 1\left(e' = e(w,t)\right)}{\sum_{e''} E_{e''p(w,t)t} - 1} s_{ee'} \tag{11}$$

Table 3 provides sample sizes for workers with different levels of education. Because 17% of workers have a college degree or higher, whereas workers with only primary school represent 3% of the sample, moving from primary school to college education amounts to an 80-percentiles rise in educational attainment. Therefore, I will often report the effect of moving from a variable's 10th to its 90th percentile as an increase comparable to moving from primary to college education. Table 4 shows this difference between 90th and 10th percentile, as well as means and standard deviations, for the main variables of interest.

[12] The argument (w,t) makes explicit that a worker can switch establishments and upgrade her education from one year to another, although the latter event is rare.

[13] In reality, the share of coworkers in an establishment that is deemed well-matched or a close substitute exceeds this value of 1% by, on average, over a factor 8. The reason is that both measures, which are constructed on the measurement sample, strongly predict coworker patterns in the estimation sample.

[14] The terms $-1\left(e' = e(w,t)\right)$ in the numerator and -1 in the denominator are added to exclude worker w herself from the group of coworkers.

Table 3: Sample sizes

	# individuals	share
primary	83,390	3.2%
sec.	348,609	13.5%
upper sec.	1,522,026	59.1%
post-sec.	184,033	7.1%
college	426,827	16.6%
Ph. D.	12,079	0.5%

Table 4: Descriptive statistics main variables

	mean	median	st. dev.	p90 - p10
cow. match	0.592	0.579	0.109	0.264
cow. subst.	0.393	0.386	0.185	0.481
# well-matched cow.	29.7	3.0	116.0	50.0
# subst. cow.	28.0	4.0	99.8	51.0
establishment size	345.0	40.0	1008.0	759.0
sh(own edu.)	0.081	0.019	0.146	0.250
age	38.8	38.0	10.4	29.0

Descriptive statistics for workers with at least upper secondary education in the estimation sample. *Cow. match* and *cow. subst.* are weighted average coworker-match and substitutability to coworkers, as in (10) and (11); *# well-matched cow.* and *# subst. cow.* count the number of coworkers with coworker match and substitutability over thresholds ζ_c and ζ_s in (8) and (9); *establishment size* in number of workers, *sh(own edu.)* the share of coworkers with the focal worker's education; *age* in years.

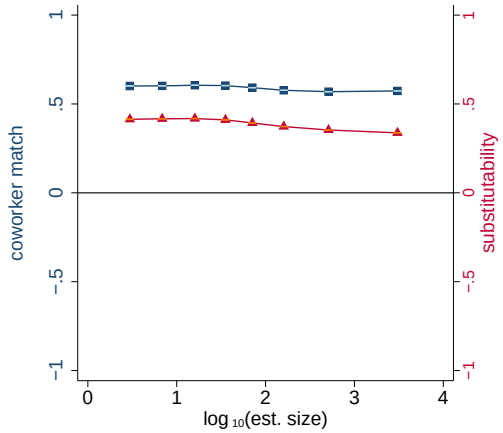


Figure 1: Coworker proximities and establishment size
Binned weighted-average coworker match (blue squares) and substitutability (red triangles) against average establishment-size in a bin, 90% confidence intervals in bright shades.

Arguably, the most relevant coworker-fit variables are the count-based measures (equations (8) and (9)). After all, these measures estimate how many coworkers are good matches or close substitutes to the focal worker. However, they are also strongly affected by an establishment's size, making it hard to disentangle establishment-size effects from effects related to coworker fit. By contrast, the weighted-average-based measures of coworker match and substitutability (equations (10) and (11)) are only weakly correlated with establishment size, as shown in Figure 1.[15] It is thus unlikely that findings using these measures will be driven by establishment-size effects. The main results of the paper are, therefore, based on the weighted averages of (10) and (11). Nevertheless, using measures based on coworker counts yields similar conclusions. In the remainder, however, *coworker match* and *substitutability* will refer to these weighted-average based measures.

What are typical work environments with high coworker matches? Appendix A.2 describes how coworker match and substitutability vary by industry. High coworker-match values are typical for industries that rely heavily on highly skilled workers, such as advanced business services and health care. By contrast, a high substitutability among coworkers is more often found in lower-skill industries, such as retail and cleaning. However, this division along skill-intensity lines is far from perfect. For instance, the construction sector consists of tightly matching teams of workers with relatively low levels of educations, who cannot easily substitute one another.

[15] In fact, the correlation between the logarithm of establishment size and weighted average coworker match (substitutability) is negative at -0.11 (-0.14).

4 Wages and coworker composition

4.1 Correlational analyses

To study the relation between a worker's coworker fit and wages, I start from the following regression model:

$$log_{10}\left(wage_{wt}\right) = X_{wt}\beta_x + Q_{p(w,t)t}\beta_p + \beta_c C_{e(w,t),p(w,t)} + \beta_s S_{e(w,t),p(w,t)} + \epsilon_{wt} \tag{12}$$

where X_{wt} is a vector of worker characteristics and $Q_{p(w,t)t}$ a vector of establishment characteristics. The main variables of interest are $C_{e(w,t),p(w,t)}$ and $S_{e(w,t),p(w,t)}$. Results are shown in Table 5.

Across all models, better coworker matches are associated with higher wages. By contrast, workers whose coworkers are close substitutes tend to earn lower wages. The estimated effects are substantial. In model (3), which estimates the effects of both variables simultaneously, without controlling for other variables, an increase from the 10th to the 90th percentile in coworker match translates into 25.3% higher wages. Moving from the 10th to the 90th percentile of substitutability to coworkers, in contrast, is associated with a wage reduction of 19.8%.

Effects of coworker match drop when adding control variables. Column (4) adds worker-level characteristics such as age and educational level. This drastically reduces the negative effect of coworker substitutability and, to a lesser extent, the positive effect of coworker match. Our preferred specification is model (5). This model controls for establishment-size and important worker characteristics, without adding variables that may partially capture coworker fit. An 80-percentiles rise in coworker match increases wages in this model by 18.1%, whereas a similar increase in substitutability lowers wages by 4.8%.

Columns (6) to (9) add a number of further control variables. These variables could capture elements of coworker fit, making conditioning on them potentially undesirable. Column (6) adds the share of coworkers who have the same education as the worker herself. This leaves effects of coworker substitutability and coworker match all but unchanged. Column (7) controls for the shares of coworkers in each of six educational levels, which reduces coworker-match effects and increases substitutability effects somewhat, but not overwhelmingly. Column (8) controls for how well a worker's education fits her occupation. This occupational match is determined in the measurement sample as:

$$M_{eo} = \frac{E_{eo} / \sum_{e'} E_{e'o}}{\sum_{o'} E_{eo'} / \sum_{e''} \sum_{o''} E_{e'',o''}} \tag{13}$$

According to this definition, an education matches an occupation if the education is *typical* for the occupation: it quantifies the degree to which an education's employment share in the occupation exceeds its share in the overall economy.[16] To reduce skewness, I map M_{eo} onto the interval $[0, 1)$, using the

[16] For instance, if workers with a college degree in "462z: Statistics" are overrepresented

Table 5: Coworker educational fit and wages

dep. var.: log(wage)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cow. match	0.126***		0.371***	0.257***	0.274***	0.282***	0.229***	0.179***	0.337***
	(0.0025)		(0.0036)	(0.0031)	(0.0030)	(0.0033)	(0.0030)	(0.0036)	(0.0032)
cow. subst.		-0.061***	-0.211***	-0.059***	-0.044***	-0.043***	-0.058***	-0.036***	-0.043***
		(0.0015)	(0.0021)	(0.0019)	(0.0018)	(0.0018)	(0.0018)	(0.0020)	(0.0018)
log(est. size)					0.044***	0.046***	0.022***	0.041***	-0.036***
					(0.0003)	(0.0005)	(0.0003)	(0.0003)	(0.0013)
log(own edu.)						-0.004***			
						(0.0007)			
edu.-occ. match								0.060***	
								(0.0013)	
log(diversity)									0.128***
									(0.0020)
4th polyn. age?				yes	yes	yes	yes	yes	yes
edu level dum?				yes	yes	yes	yes	yes	yes
edu composition?							yes		
fixed effects?	yr	yr	yr	yr	yr	yr	yr	yr	yr
R^2	0.040	0.038	0.060	0.269	0.300	0.300	0.340	0.300	0.305
# obs.	2,144,965	2,144,965	2,144,965	2,144,965	2,144,965	2,144,965	2,144,965	1,640,144	2,144,965
# clust.	364,642	364,642	364,642	364,642	364,642	364,642	364,642	323,400	364,642

***: p¡0.01; **: p¡0.05; *: p¡0.10. OLS regressions with $log_{10}(wage)$ as a dependent variable using workers with at least upper secondary education. edu.-occ. match is quantified as in (14); diversity is the logarithm (base 10) of the number of distinct educational tracks in the establishment, including the worker's own. Standard errors, clustered at the worker level, in parentheses.

Table 6: Wage regressions, fixed effects models

dep. var.: log(wage)	(1)	(2)	(3)	(4)
cow. match	0.274***	0.165***	0.160***	0.057***
	(0.0030)	(0.0054)	(0.0068)	(0.0075)
cow. subst.	-0.044***	-0.046***	-0.048***	-0.016***
	(0.0018)	(0.0029)	(0.0030)	(0.0039)
log(est. size)	0.044***	0.028***		0.051***
	(0.0003)	(0.0004)		(0.0020)
4th polyn. age?	yes	yes	yes	yes
edu level dum?	yes	yes	yes	yes
fixed effects?	yr	yr, w.	yr×est.	yr, w.×est.
R^2	0.300	0.810	0.648	0.886
# obs.	2,144,965	2,144,965	2,144,965	2,144,965
# clust.	364,642	364,642	144,371	144,371

***: p¡0.01; **: p¡0.05; *: p¡0.10. OLS regressions of log_{10} (*wage*) using workers with at least upper secondary schooling. Standard errors (in parentheses) are clustered at the worker level in columns (1) and (2) and at the establishment level in columns (3) and (4).

same transformation as in equation (7):

$$\mu_{eo} = \frac{M_{eo}}{M_{eo} + 1} \tag{14}$$

Controlling for occupational match reduces the coworker-match effect by about a third and the substitutability effect by one fifth vis-à-vis the preferred specification. Finally, column (9) controls for workforce diversity by adding the logarithm of the number of different educational tracks in an establishment. Adding this control strengthens the estimated effects of coworker match and substitutability slightly. Overall, however, none of the variables in columns (6) to (9) can explain a major part of the effects of coworker match and substitutability on wages.

One concern is that capable workers sort themselves into well-matching work environments and, therewith, create a positive correlation of wages with coworker match and a negative correlation with coworker substitutability.[17] To control for such sorting effects, Table 6 adds worker, establishment, and worker-establishment fixed effects.

Adding either worker (column (2)) or establishment-year (column (3)) fixed

in occupation "2413: Market analysts," I interpret this as evidence that statisticians are well-suited for a job as market analyst.

[17]Note that such sorting would not immediately invalidate the conclusion that a good coworker-fit is important. After all, the fact that workers sort themselves into well-fitting work environments suggests that such environments are attractive. Only if this attractiveness has other than monetary reasons, does the correlation between coworker match (or lack of substitutability) and wages derive from the fact that high-wage jobs *attract* productive workers (as opposed to a good coworker-fit *causing* higher productivity and wages).

effects reduces the estimated effect of coworker match by roughly 40%, while leaving the estimated effect of substitutability unchanged. Column (4) adds worker-establishment fixed effects such that effects are identified solely from variation that arises because the team around a worker changes, when coworkers enter or exit the establishment, while the worker herself does not change jobs. Although effect estimates are reduced substantially, even in this specification better coworker matches are associated with higher wages, whereas being easily substituted by coworkers is associated with lower wages.

4.2 Measurement error

Weakening coworker-fit effects in fixed-effects specifications are consistent with the presence of unobserved ability-based sorting. Arguably, however, the lion's share of such sorting effects should be absorbed in *worker* fixed effects. It is, therefore, surprising that the most precipitous drop in point estimates results when controlling for worker-establishment, not worker fixed-effects. An alternative explanation for why point estimates are lower in fixed-effects models is that our variables of interest are mismeasured. In particular, if a variable's actual values are more strongly autocorrelated than its measurement error, fixed effects will absorb large parts of the variable's signal, but little of its noise.[18] The consequent deterioration in signal-to-noise ratio will exacerbate existing attenuation biases. In this subsection, I will first show that there are strong reasons to believe that coworker-match is estimated with substantial error and then present two strategies to quantify the resulting errors-in-variables bias.

Symptoms of measurement error

A first sign of noise in the coworker match variable can be found in the fact that the coworker match for a given educational pair varies markedly from one year to the next. The correlation between an educational pair's coworker match in two consecutive years is, on average, 0.88. If we assume that there are no abrupt changes in the intrinsic match of two educations, a correlation below one reflects measurement error.[19] To the extent that measurement errors carry over to the worker-establishment level, the estimated effects in Table 5 will be biased towards zero. In comparison to coworker-match estimates, measurements of substitutability seem much less noisy. In fact, estimates of the substitutability of educational pairs are highly stable over time, with a correlation of 0.98 between consecutive years.[20]

[18] The addition of control variables in the models of Table 5 that are correlated with the variable's signal, but not with the noise has similar consequences.

[19] The implied measurement-error variance-share of $1 - 0.88^2 = 0.23$ underestimates true measurement error because teams typically endure several years. As a consequence, year-on-year correlations are artificially high.

[20] A plausible reason for this is that, whereas coworker match is based on co-occurrences in *establishments*, substitutability is measured by the relation between the occupation and education of *individuals*. How precisely estimated coworker match is depends therefore on the total number of establishments in which two educations are present, whereas the precision

Table 7: Fixed-effects models, within-transformation versus first-differences

dep. var.: log10(wage)	(1) OLS	(2) within	(3) 1st dif.	(4) within	(5) 1st dif.
cow. match	0.274***	0.165***	0.076***	0.057***	0.045***
	(0.0030)	(0.0049)	(0.0031)	(0.0058)	(0.0042)
cow. subst.	-0.044***	-0.046***	-0.020***	-0.016***	-0.013***
	(0.0018)	(0.0026)	(0.0017)	(0.0032)	(0.0024)
log(plant size)	0.044***	0.028***	0.015***	0.051***	0.048***
	(0.0003)	(0.0004)	(0.0003)	(0.0009)	(0.0008)
4th polyn.of age?	yes	yes	yes	yes	yes
edu. level dum.?	yes	yes	yes	yes	yes
fixed effects?	yr	yr, w.	yr, w.	yr, w.×est.	yr, w.×est.
R^2	0.300	0.277		0.242	
N	2,144,965	2,144,965	1,697,584	2,144,965	1,474,639

***: p¡0.01; **: p¡0.05; *: p¡0.10. Standard errors (in parentheses) are clustered at the worker level in model (1) and robust in columns (2) to (5).

If the reduction in point estimates in the fixed-effect models is indeed due to measurement error, it matters *how* fixed effects are removed. Doing so by means of a within-transformation typically yields less-attenuated parameter estimates than first-differencing the data (Griliches and Hausman, 1986).[21] Table 7 shows that this is indeed the case. Model (1) repeats the preferred OLS specification of Table 5. Models (2) and (3) include worker fixed-effects, models (4) and (5) worker-establishment fixed-effects. The within-transformation is used in models (2) and (4), whereas fixed-effects are eliminated through first-differencing in models (3) and (5).

The difference between the two panel-data approaches is striking. Although the within-transformed models in columns (2) and (4) should, theoretically, give the same results as the first-differenced models in columns (3) and (5), estimated effects in the latter are much lower than in the former, supporting the notion that (some of) the drop in point estimates in fixed effects models can be attributed to measurement error.

Estimating the magnitude of measurement error

To quantify the bias measurement error introduces, I derive theoretical expressions for the measurement-error variance of coworker match. To do so, let coworker-match estimates in year t be composed of two unobserved components:

with which educational substitutability is estimated depends on how many individuals absolved these educations. Because the latter number is typically much larger than the former, substitutability will be measured more accurately than coworker match.

[21] First-differencing will particularly inflate the error-to-signal ratio vis-à-vis a within-transformation if the number of observations per individual is large and if the signal is strongly and positively autocorrelated, whereas the noise is not.

the true coworker match and measurement error. That is, let:

$$\hat{c}_{ee't} = c_{ee'} + \nu_{ee't} \tag{15}$$

where $\hat{c}_{ee't}$ is the observed coworker match of educational tracks e and e', $c_{ee'}$ the underlying, actual coworker match and $\nu_{ee't}$ a measurement error that is uncorrelated with $c_{ee'}$. Let $N_{ee'}$ be the number of co-occurrences of education e with e'. Furthermore, let a subscript "." indicate a summation over the omitted category, i.e.: $N_{e.} = \sum_{e'} N_{ee'}$, $N_{.e'} = \sum_{e'} N_{ee'}$ and $N_{..} = \sum_{e}\sum_{e'} N_{ee'}$. Now assume that $N_{ee'}$ is drawn from a Binomial distribution, $N_{ee'} \sim BIN\left(\Pi_{ee'}, N_{..}\right)$, and, therefore has a variance of:

$$V\left[N_{ee'}\right] = N_{..}\Pi_{ee'}\left(1 - \Pi_{ee'}\right) \tag{16}$$

$\Pi_{ee'}$, the probability of a co-occurrence between educations e and e', is unknown, but can be estimated by the observed relative frequency of $N_{ee'}$. Denoting observed or estimated quantities by a hat ("ˆ"), we get: $\hat{\Pi}_{ee'} = \frac{\hat{N}_{ee'}}{\hat{N}_{..}}$. Consequently, (16) can be written as:

$$V\left[N_{ee'}\right] = N_{..}\frac{\hat{N}_{ee'}}{\hat{N}_{..}}\left(1 - \frac{\hat{N}_{ee'}}{\hat{N}_{..}}\right) \tag{17}$$

One problem is that, for the vast majority of educational pairs, $\hat{N}_{ee'}$ equals zero. Therefore, equation (17) leads to the implausible conclusion that coworker match is perfectly measured in these pairs. This happens, because the uncertainty in $\hat{N}_{ee'}$ is not taken into account. To address this, I re-estimate $\hat{\Pi}_{ee'}$ in a Bayesian framework. First, I determine expectations and variances for $N_{ee'}$, assuming that $N_{ee'}$ is drawn from a Hypergeometric distribution that takes the total number of co-occurrences in which educations e and e' participate as given and equal to $\hat{N}_{e.}$ and $\hat{N}_{.e'}$. Next, from these variances and expectations, I derive a prior distribution for $\Pi_{ee'}$, which is then updated with information on the actually observed number of co-occurrences, $\hat{N}_{ee'}$. The posterior expectation of $\Pi_{ee'}$, $\hat{\Pi}_{ee'}^{post}$, that results from this exercise is always strictly greater than zero.[22] Now recall equation (7), which defines the coworker match of educational pair (e, e'), $c_{ee'}$, as follows:

$$c_{ee'} = \frac{\kappa_{ee'}\hat{N}_{ee'}}{\kappa_{ee'}\hat{N}_{ee'} + 1}$$

where $\kappa_{ee'} = \frac{\hat{N}_{..}}{\hat{N}_{e.}\hat{N}_{.e'}}$. Its variance is therefore given by:

$$V\left[c_{ee'}\right] = V\left[\frac{\kappa_{ee'}\hat{N}_{ee'}}{\kappa_{ee'}\hat{N}_{ee'} + 1}\right] \tag{18}$$

[22] Details are provided in in Appendix A.3.

Approximating this variance using the delta method, we get:

$$V\left[c_{ee'}\right] \approx V\left(\hat{N}_{ee'}\right)\left(\frac{\kappa_{ee'}+\hat{N}_{ee'}\frac{d\kappa_{ee'}}{d\hat{N}_{ee'}}}{\left(\kappa_{ee'}\hat{N}_{ee'}+1\right)^2}\right)^2$$

with $\frac{d\kappa_{ee'}}{d\hat{N}_{ee'}} \approx -\hat{N}_{..}\frac{\hat{N}_{e.}+\hat{N}_{.e'}}{\left(\hat{N}_{e.}\hat{N}_{.e'}\right)^2} = k\frac{\hat{N}_{e.}+\hat{N}_{.e'}}{\hat{N}_{e.}\hat{N}_{.e'}}$.[23] Using the expectation for $\hat{N}_{ee'}$, $\hat{\Pi}^{post}_{ee'}\hat{N}_{..}$, and its variance, $\hat{N}_{ee'}\hat{\Pi}^{post}_{ee'}\left(1-\hat{\Pi}^{post}_{ee'}\right)$, results in the following (non-zero) error-variance for educational pair (e,e')'s coworker match:

$$V\left[c_{ee'}\right] \approx \hat{N}_{..}\hat{\Pi}^{post}_{ee'}\left(1-\hat{\Pi}^{post}_{ee'}\right)\left(\frac{\kappa_{ee'}\left(1-\hat{N}_{..}\hat{\Pi}^{post}_{ee'}\frac{\hat{N}_{e.}+\hat{N}_{.e'}}{\hat{N}_{e.}\hat{N}_{.e'}}\right)}{\left(\kappa_{ee'}N_{..}\hat{\Pi}^{post}_{ee'}+1\right)^2}\right)^2 \tag{19}$$

The theoretical standard deviation that follows from equation (19) correlates surprisingly well with the empirically estimated (timeseries) standard deviation: the estimated rank correlation between the two quantities is 0.826.[24] This suggests that equation (19) yields a good prediction of the heteroscedasticity in $c_{ee'}$'s measurement error.

Errors-in-variables correction: extrapolation

The coworker match between educational pairs, $c_{ee'}$, is aggregated in equation (10) to the worker-establishment level into $C_{e(w,t),p(w,t)}$, by taking its weighted average across a worker's coworkers. Assuming that measurement errors are uncorrelated across educational pairs and with an establishment's employment composition, the error-variance of $C_{e(w,t),p(w,t)}$ becomes:

$$V\left[C_{e(w,t),p(w,t)}\right] = \sum_{e'}\left(\frac{E_{e'p(w,t)t}-1\left(e'=e\left(w,t\right)\right)}{\sum_{e''}E_{e''p(w,t)t}-1}\right)^2 V\left[c_{e(w,t)e'}\right] \tag{20}$$

The smaller $V\left[C_{e(w,t),p(w,t)}\right]$ is, the more accurately measured $C_{e(w,t),p(w,t)}$ will be. Determining $V\left[C_{e(w,t),p(w,t)}\right]$ for each observation in the data allows us to focus on observations with minimal measurement error. This, in turn, should minimize attenuation biases. Following this logic, I divide all worker-year observations into ten error-variance bins with equal numbers of observations and then run separate regression analyses for each bin, $\mathscr{B}_b$, with $b \in \{1,2,...,10\}$:

$$log_{10}\left(wage_{wt}\right) = X_{wt}\beta_x + Q_{p(w,t)t}\beta_p + \gamma_b C_{e(w,t),p(w,t)} + \beta_s S_{e(w,t),p(w,t)} + \epsilon_{wt} \tag{21}$$

where $(w,t) \in \mathscr{B}_b$ denotes the set of observations in decile b, and ϵ_{wt} is an error term. The control variables collected in X_{wt} and $Q_{p(w,t),t}$ are a 4th order polynomial of worker age, year fixed-effects, educational-track fixed-effects,

[23] The approximation uses the fact that $\hat{N}_{..}\left(\hat{N}_{e.}+\hat{N}_{.e'}\right) \ggg N_{e.}N_{.e'}$.

[24] A scatter plot is provided in Appendix A.3.

coworker-shares by educational level, and the logarithm of establishment size. These characteristics are chosen to ensure that error-variances and wages are approximately conditionally independent.

As shown in Appendix A.3, effect-estimates for coworker match will be biased according to the following equation:

$$\hat{\gamma}_b = \gamma \left(1 - \frac{V[\eta]}{V\left[\hat{C}\right]} \right) \tag{22}$$

where $\hat{\gamma}_b$ is the estimated effect, γ the real effect, $V[\eta]$ represents the error-variance in bin b and $V\left[\hat{C}\right]$ the variance of $C_{e(w,t),p(w,t)}$ conditional on the other regressors of (21) in this bin.[25]

Figure 2a plots the estimated $\hat{\gamma}_b$'s against $\frac{V[\eta]}{V\left[\hat{C}\right]}$, where $V[\eta]$ is estimated using (20). The relation between the estimated coworker-match effect and the error-variance share is striking: the lower the error-variance, the higher point-estimates become. Extrapolating the linear trend line implied in (22) suggests that the unbiased effect of coworker match (i.e., the value at which this trend line crosses the vertical axis) lies roughly between 0.4 and 0.55.[26] Interestingly, disregarding the heteroscedasticity in measurement errors at the educational-pair level yields very similar results. Appendix A.3 shows that when one assumes that $V[c_{ee'}] = \sigma_c^2$ for all pairs (e, e') extrapolated effects are all but indistinguishable from the ones in Figure 2a.

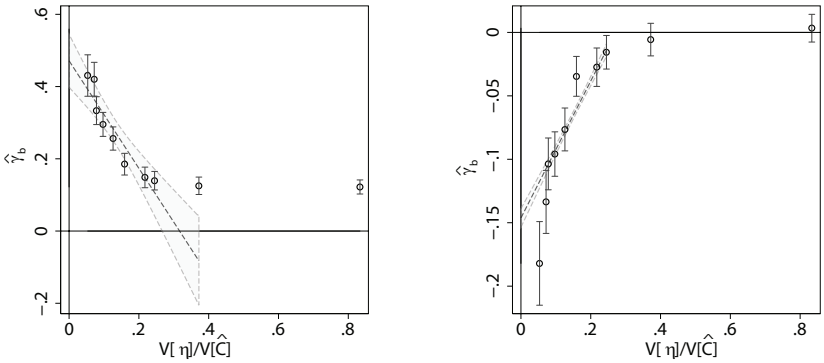
Figure 2b shows that, whereas $\hat{\gamma}_b$ rises as $\frac{V[\eta]}{V\left[\hat{C}\right]}$ becomes smaller, the estimated effect of $S_{e(w,t),p(w,t)}$ drops. This is indeed what one would expect if $\frac{V[\eta]}{V\left[\hat{C}\right]}$ quantified the mismeasurement of $C_{e(w,t),p(w,t)}$. In that case, the strong positive correlation between coworker match and substitutability would typically mean that the downward bias in the estimated effect of the former induces an *upward* bias in the estimated effect of the latter.

Errors-in-variables correction: 2SLS

A different way to correct for mismeasured coworker matches is to instrument this variable with an alternative proxy for how well a worker is matched to her work environment. A plausible candidate for such a proxy is the match between a worker's education and her occupation, μ_{eo}, as defined in (14). Assuming that measurement errors in coworker-match and education-occupation match are uncorrelated, using the latter as an instrument for the former should remove

[25] $V\left[\hat{C}\right]$ is calculated as the variance of the residual of a regression of $C_{e(w,t),p(w,t)}$ on X_{wt}, $Q_{p(w,t),t}$ and $S_{e(w,t),p(w,t)}$ in bin b.

[26] The trend line downweights the two outliers with extreme error-variances. It is plausible that measurement errors are overestimated in these bins, because the educational track dummies in (21) also absorb structural measurement-error components that are specific to educational tracks. As a consequence, the residual measurement-error variance may be smaller than what is plotted along the horizontal axis.



(a) Coworker match

(b) Substitutability

Figure 2: Estimated effect of coworker-match by error-variance bin

Split-sample estimates of effect of coworker match (Figure 2a) and substitutability (Figure 2b) by error-variance decile of $C_{e(w,t)p(w,t)}$ against the decile's average error-variance share, $\frac{V[\eta]}{V[\hat{C}]}$. Dashed line depicts regression of $\hat{\gamma}_b$ on $\frac{V[\eta]}{V[\hat{C}]}$, including a 95% confidence interval, downweighting outliers in a robust regression using biweights.

the attenuation bias caused by measurement error. Table 8 compares the results of this approach (column (2)) to the preferred OLS specification of Table 5 in column (1).

The 2SLS estimate using μ_{eo} to instrument for $C_{e(w,t),p(w,t)}$ exceeds its OLS counterpart by a substantial margin. Interestingly, once again the negative effect of substitutability strengthens considerably as well. In fact, the effects of coworker match and substitutability reported in column (2) are - given their standard errors - indistinguishable from the ones implied by Figure 2.

4.3 Causal effects

The error-corrected estimates suggest that the drop in effect estimates associated with the introduction of worker fixed effects may be attributable to an exacerbation of measurement error. In that case, although the fixed-effects models may correct for ability-based sorting, they would severely underestimate the true importance of good coworker matches. Consequently, models with worker fixed effects would provide a lower bound on the true effect of coworker match. However, there are several reasons for concern over such an interpretation. First, although worker fixed effects *can* eliminate biases related to ability-based sorting, they do not *necessarily* do so. For instance, worker fixed effects will not absorb time-varying aspects of ability, nor can they adequately correct for workers' intrinsic ability if this ability is only revealed with time (see Gibbons et al., 2005). A second concern is that workers may be assigned to jobs where they earn the highest wages as in the Roy model (Roy, 1951). In this case, observed wages represent a (nonrandomly) selected sample of a population of *potential* wage of-

Table 8: Errors-in-variables estimates

dep. var.: log(wage)	(1) OLS	(2) 2SLS
cow. match	0.274***	0.547***
	(0.0030)	(0.0071)
cow. subst.	-0.044***	-0.163***
	(0.0018)	(0.0033)
log(est. size)	0.044***	0.043***
	(0.0003)	(0.0003)
4th polyn. of age?	yes	yes
edu. level dum.?	yes	yes
fixed effects?	yr	yr
# obs.	2,144,965	1,640,144
# clust.	364,642	323,400
	First Stage	
match edu.-occ.		0.164***
		(0.0006)
t-stat.		281.7

***: p¡0.01; **: p¡0.05; *: p¡0.10. Standard errors, clustered at the worker level, in parentheses.

fers.[27] This section will attempt to estimate the causal effect of coworker match on wages by instrumental variables (IV) estimations. This approach should correct for various sources of endogeneity, as well as measurement error.

Focusing on workers who neither change establishments, nor upgrade their education $C_{e(w,t),p(w,t)}$ and $S_{e(w,t),p(w,t)}$ can be written as C_{ept} and S_{ept}. Now, consider wage equation (12) in first differences:

$$\Delta log_{10}\left(wage_{wt}\right) = \Delta X_{wt}\beta_x + \Delta Q_{pt}\beta_p + \Delta C_{ept}\beta_c + \Delta S_{ept}\beta_s + \Delta\epsilon_{wt} \tag{23}$$

where Δ denotes the first-differencing operator. The main concern is that ΔC_{ept} and ΔS_{ept} are correlated with $\Delta\epsilon_{wt}$ because of sample-selection biases, learning-about-ability or measurement error. To isolate exogenous variation in changes in C_{ept} and S_{ept}, I rely on shifts in the local availability of good matches for worker w. The underlying idea is that if there were an exogenous increase in the local supply of matching workers, this would lower the price of these workers, and therefore increase the likelihood of them being hired.

The supply shift I exploit is based on the number of graduates in a region. However, because students' educational choices may reflect future employment prospects, I do not use *observed* graduation rates as an instrument. Instead, I predict the number of local graduates in each educational track by combining information on the local educational structure between 1990 and 1995 with the national growth rate of the educational track in question.[28]

As an illustration, assume that, between 1990 and 1995, Gothenburg trained 20% of all Swedish automotive engineers. Furthermore, assume that a total of 800 students graduated in automotive engineering in Sweden but *outside* Gothenburg in 2001. If Gothenburg's educational share of automotive engineers in Sweden didn't change, Gothenburg should produce 200 graduates in 2001. I use this logic to predict the number of graduates (G_{emt}) for each educational track e and every municipality m in year t as:

$$\hat{G}_{emt} = \frac{q_{em}}{1 - q_{em}} \sum_{m \neq m'} G_{em't} \tag{24}$$

where $q_{em} = \frac{\sum_{t=1990}^{1995} G_{emt}}{\sum_{t=1990}^{1995} \sum_{m''} G_{em''t}}$ represents municipality m's historical share of graduates in educational track e. To calculate the number of local graduates that match education e well, I consider an education e' well-matching if $c_{ee'}$ exceeds a given threshold. Next, I divide the predicted number of well-matching graduates by the number of well-matching workers living in the region at the beginning of

[27] Similar issues are identified in studies on returns to overeducation: "... overeducation researchers are [...] aware of the fact that estimated returns to overeducation may reflect differences across people in terms of other unobserved components of their human capital stock or of their motivation. ... [But it] it is extremely difficult to obtain credible estimates of causal effects of being over/underschooled" Leuven et al. (2011, p. 306).

[28] The end of this period roughly coincides with a substantial overhaul of the Swedish higher education system. However, also absent this overhaul, one wouldn't expect the local shares of graduates in the early 1990s to affect year-on-year wage-changes occurring after 2001.

the period.[29] This number represents an exogenous shift in the local availability of good matches for education e. I construct this instrument for two different thresholds[30] and at two spatial scales, namely, at the municipality level and at the level of labor market areas.[31] Finally, to correct for changes in local market conditions, I control for region-industry-year fixed effects.

One complication of the specification in first-differences is that the wage data refer to a worker's annual wage, i.e., to total wages earned in a given year. Because coworkers may enter or leave an establishment at any time during the year, the change in coworker match between years t and $t+1$ may already affect the wages of year t. Therefore, wage changes are measured as annualized changes over a two-year period, from $t-1$ to $t+1$. Furthermore, I restrict the sample to workers who work continuously at the same establishment between years $t-2$ and $t+1$ to ensure that all wages reflect a full year's employment in one and the same establishment.

Due to their collinearity, it proves to be infeasible to instrument substitutability and coworker match simultaneously using this identification strategy.[32] Table 9 therefore only shows two-stage least squares (2SLS) estimates for the effect of coworker match.

Model (1) shows the 2SLS, reduced form and first-stage estimates for the preferred specification, which uses all four instruments and controls for industry-municipality-year fixed effects. The point estimate of 0.637 implies that a move from the 10th to the 90th percentile in coworker match translates into a 47% increase in wages. This is over twice as high as the original OLS effect (column (5) of Table 5). However, given the large standard errors, the OLS estimate is still well within a 95% confidence interval around this point estimate. Moreover, the Hansen J-statistic on over-identifying restrictions does not differ significantly from zero in any of the models with multiple instruments.

To explore the robustness of these findings, I re-estimate model (1) with only year (column (2)), industry-year (column (3)) and region-year fixed effects (column (4)). Although point estimates in these models exceed the ones in model (1), differences are well within the margin of error.[33] Furthermore, although the Kleibergen-Paap F-statistic of 14.1 indicates that the instruments are reasonably strong in model (1), the large, imprecisely estimated effects still raises concerns over weak identification. To explore this, Model (5) uses a Limited Information Maximum Likelihood (LIML) estimator and Model (6) estimates a just-identified model with only the strongest instrument. Both models produce

[29] That is, I construct the instrument using information on the municipality of residence, not of work.

[30] The first (education-specific) threshold is chosen such that, for each education, the expected likelihood that a random pair of workers will be close matches is 0.1%, whereas the second sets this expected likelihood at 1%.

[31] I exclude the worker's own municipality to increase the difference in what the two instruments measure.

[32] When instrumenting both variables simultaneously, the Kleibergen-Paap F-statistic drops to 1.874 when using four instruments and 4.661 when using only one.

[33] Moreover, the lower Kleibergen-Paap statistic means that these increases may reflect weak-instrument problems.

similar results to model (1), suggesting that the high point-estimates are not caused by weak instruments.

4.4 Coworker complementarity

Table 10 collects the main findings of this section. OLS estimates of coworker-match effects are substantially lower than IV estimates, for which I have proposed three explanations. First, in spite of the statistically insignificant Hansen J-statistic, the identification strategy in section 4.3 may not address endogeneity concerns perfectly. For instance, shifts in the local supply of well-matching workers may directly impact wages by improving a worker's outside options and, therewith, her bargaining position in an establishment. In that case, the exclusion restriction will not hold.[34]

To explore this explanation further, I rerun the IV analyses of Table 9 on a sample of *static* establishments, i.e., establishments that do not change their workforce.[35] In such establishments, the instrument can only affect workers' wages directly, but not indirectly through a change in coworker match. The sum of the instrument's direct and indirect effects on wages is given by the reduced-form effect. In the full sample, these reduced-form effects are highly significant. For instance, in model (6) of Table 9 the preferred instrument's reduced-form effect on wages has a t-value of 2.83. In the static-establishment sample in contrast, the corresponding t-value is -0.09. In fact, in none of the models do instruments have a statistically significant reduced-form effect on wages, and consequently.[36] Consequently, there is no evidence that instruments affect wages directly.[37]

A second explanation for why OLS estimates exceed IV estimates is sample-selection effects. When jobs are assigned according to where workers earn the highest wages, the nonrandom sampling of observed (i.e., accepted) wage offers will lead to downward biases in OLS estimates. In that case, the higher IV estimates arise because IV estimators correct for this bias. The third explanation for deflated OLS estimates is measurement error. Also now, higher IV estimates result from problems in the OLS specifications, not in the identification strategy. This third explanation finds support in the error-correction approaches: various ways to correct for measurement errors all yield point estimates that are surprisingly similar to each other *and* to the IV estimates in Table 9. Moreover, this interpretation is also in line with the fact that not just the positive effect of coworker match, but also the negative effect of substitutability is amplified

[34] The importance of this mechanism depends on an establishment's relative size to the local labor market: De Fontenay and Gans (2003) show that, when the number of outsiders in a firm becomes large relative to the number of insiders, intrafirm bargaining power of workers dissipates and wages approach the competitive-market equilibrium.

[35] To be precise, I focus on workers in establishments that neither hire nor fire anyone between the current and the subsequent year.

[36] As a consequence, the accompanying IV estimates are all insignificant.

[37] All reduced-form estimates corresponding to the models in Table 9 for workers in static establishments, as well as F-tests for the joint significance of the instruments, are shown in Table A.2 in Appendix A.4.

Table 9: Instrumental variable regression, first differences

		(1)		(2)	(3)	(4)	(5)	(6)
model:	2sls	fs	rf	2sls	2sls	2sls	liml	pref
cow. match	0.637**			1.099***	0.948***	0.956***	0.677**	0.788***
	(0.265)			(0.286)	(0.274)	(0.283)	(0.283)	(0.301)
log(est. size)	0.027***	-0.017***	0.016***	0.040***	0.035***	0.033***	0.027***	0.029***
	(0.005)	(0.001)	(0.001)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
4th polyn. of age?	yes	yes	yes	yes	yes	yes	yes	yes
fixed effects?	yr×m×i	yr×m×i	yr×m×i	yr	yr×m	yr×i	yr×m×i	yr×m×i
IV - matching grad. (0.1%, m)		0.0013***	0.0013*					
		(0.0004)	(0.0007)					
IV - matching grad. (1%, m)		0.0014***	0.0006					
		(0.0005)	(0.0009)					
IV - matching grad. (0.1%, r)		0.0004	0.0029*					
		(0.0007)	(0.0015)					
IV - matching grad. (1%, r)		0.0022**	-0.0015					
		(0.0009)	(0.0018)					
Kleibergen-Paap	14.1	14.1	14.1	11.7	11.7	11.9	14.1	44.4
P-val. Hansen J	0.220			0.457	0.363	0.694	0.224	
# obs.	727,982	727,982	727,982	771,684	771,684	771,470	727,982	727,982
# clust.	115,802	115,802	115,802	123,820	123,820	123,791	115,802	115,802
Instruments	all			all	all	all	all	(1%, r)

***: p¡0.01; **: p¡0.05; *: p¡0.10. First-differenced regression as specified in equation (23) using workers who do not change establishments or educations between the years $t-2$ and $t+1$. The dependent variable is the logarithm of the annualized wage growth between years $t-1$ and $t+1$. Model (1) shows results from a 2SLS regression (together with the first stage - fs - and reduced form - rf - estimates, using four instruments: the number of highly matching (threshold of 0.1%) and of well-matching (threshold of 1%) graduates in the establishment's municipality and analogous variables at the labor market area. Models (1), (5) and (6) control for year-municipality-industry (yr×m×i) fixed effects. Models (2), (3) and (4) show analyses with only year, year-municipality and year-industry fixed effects. Model (5) uses the LIML instead of the 2SLS estimator. Model (6) uses only the preferred instrument (threshold of 0.1% at the labor market area level). Standard errors, clustered at the year-municipality-educational track level, in parentheses.

in all estimations that correct for measurement error. Overall therefore, the weight of the evidence seems to lie with this third explanation.

Regardless of the exact size of coworker-match effects, all models in Table 10 consistently show that better coworker matches are associated with higher wages. This points towards a particular interpretation of the coworker-match variable: it seems to capture the *complementarity* among coworkers.[38] After all, if higher wages reflect higher productivity, these positive effects would mean that coworker match acts as a measure of Milgrom and Roberts's (1995) Edgeworth complementarity.[39]

In fact, if employers tend to hire teams of workers who complement each other, it is not surprising that coworker match - which quantifies how often educational tracks are combined in establishments - behaves as a measure of complementarity. However, employers will typically hire a mix of complements and substitutes. After all, whenever workloads exceed the capacity of a single worker, this worker's skill set needs to be duplicated. This explains why coworker match is positively correlated with substitutability, i.e., why workers often work together with close substitutes.

The fact that the effect of coworker match strongly rises when conditioning on substitutability suggests that, to accurately capture coworker complementarity, one must isolate the component of coworker match *orthogonal* to substitutability. To do so, I calculate the residual of a regression of coworker match on substitutability:

$$C_{e(w,t)p(w,t)t} = \alpha^c_{lt} + \beta^c_{lt} S_{e(w,t)p(w,t)t} + m_{wpt} \tag{25}$$

where m_{wpt} is an I.I.D. error term. To remain agnostic about the functional relation between coworker match and substitutability, the coefficient β^c_{lt} is allowed to vary freely over time and with worker w's level of education. Similarly, the intercept, α^c_{lt}, interacts the worker's level of education with year dummies. In what follows, I will interpret the estimated "excess coworker-match" from equation (25), $\hat{m}_{wpt}$, as a worker's *complementarity* to her coworkers and explore how this complementarity evolves over workers' careers and how it affects a number of well-known wage premiums.

5 Career paths

5.1 Evolution of coworker complementarity

I start by studying how coworker complementarities evolve as workers gain work experience. As a benchmark, I compare this evolution to a worker's wage evo-

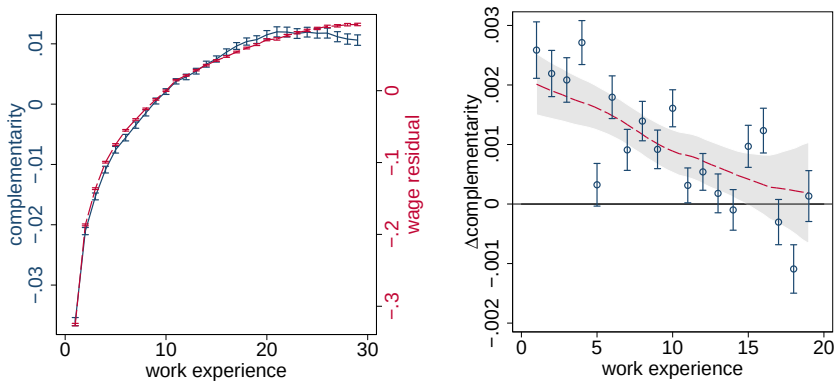
[38] This interpretation is in line with how Dibiaggio et al. (2014) interpret their measure of complementarity of technological fields. These authors use the fact that different technologies are used in the same patents as an indication of their complementarity.

[39] Two production factors are Edgeworth complements if the expansion of one increases the productivity of the other. As explained in section 2, the positive effect of coworker match would signal the existence of complementarities among coworkers, even if workers aren't paid their marginal productivity.

Table 10: Wage regressions: summary

		OLS		causal	error-correction	
dep. var.:	(1)	(2)	(3)	(4)	(5)	(6)
log(wage)	OLS	FE	FE	IV	extrapolated	error-IV
cow. match	0.274***	0.165***	0.057***	0.637**	0.471***	0.547***
	(0.003)	(0.005)	(0.007)	(0.265)	(0.035)	(0.007)
cow. subst.	-0.044***	-0.046***	-0.016***		-0.146***	-0.163***
	(0.002)	(0.003)	(0.004)		(0.002)	(0.003)
fixed effects?	yr	yr, w.	yr, w.×est	yr×m×i	yr	yr

***: p¡0.01; **: p¡0.05; *: p¡0.10. Columns (1) to (3): OLS, worker fixed effects and worker-establishment fixed effects models; column (4): 2SLS outcomes with supply-shift instruments; column (5): error-correction, extrapolated zero-error-variance estimates of Figure 2; column (6): error-correction, 2SLS model instrumenting coworker match by education-occupation match.



(a) Complementarity and wage residuals

(b) Change in complementarity at job switch

Figure 3: Coworker complementarity and work experience

lution, as described by the residual of the following regression:

$$log_{10}\left(wage_{wpt}\right) = \alpha^{c}_{lt} + \omega_{wpt} \tag{26}$$

where α^{c}_{lt} is an interaction of year dummies with educational-level dummies. Figure 3a plots the estimated residual, $\hat{\omega}_{wpt}$, as well as complementarity, $\hat{m}_{wpt}$, over the first 30 years of a worker's career.[40]

As workers progress in their careers, coworker complementarities rise along a concave curve. The figure shows how strikingly long-lasting an imprint education choices leave on workers' work environments: educational complementarities among coworkers keep rising for up to 20 years after workers have started their careers (and typically, have finished their education).

The rise in complementarity with work experience results from the interplay of two phenomena. First, many workers who change employers find more complementary teams. This is shown in Figure 3b, which plots the changes in coworker complementarity for workers who change jobs. Second, as section 5.2 will show, low coworker-complementarities increase the likelihood *that* workers change jobs.

Intuitively, as complementarities rise, it should become harder to find even more complementary teams, while, at the same time, the higher wages associated with better matches will increase the opportunity costs of search. Consequently, one would expect the rate at which workers manage to improve their coworker complementarity to drop. This expectation is confirmed by, both, the concavity of Figure 3a, and the negative slope of Figure 3b. In fact, in line with research

[40] Work experience refers to the actual number of years worked. For workers who were already employed in 1990, the first year with reliable establishment identifiers, I augment this with their potential work experience in 1990.

on educational mismatch,[41] most improvements in coworker complementarity are realized within the first 5 to 10 years after entering the labor market.

A final interesting finding is that complementarity and wage-residual curves have remarkably similar shapes. Figure A.7 in Appendix A.5 shows that this also holds true when plotting separate curves for each educational level. However, note that the increase in wage over a worker's career is an order of magnitude larger than what, given the estimates in section 4, the increase in complementarity can account for. Therefore, the close relation between the two quantities is unlikely to be fully causal. Instead, both curves may reflect that improvements in a worker's job match diminish over time.

5.2 Job switching

If job-to-job mobility is a means to finding better-matching work environments, one would expect that bad matches result in higher establishment-switching rates. The relation between establishment tenure and coworker match is described in Table 11. Column (1) shows results of a linear probability model (LPM) where the dependent variable is a dummy for whether or not a worker switches establishments. Models (2) to (5) are LPMs with a dummy that assesses whether a worker has at least 2, 3, 4, or 5 years of tenure at an establishment.[42]

Increases in coworker match decrease the likelihood that a worker switches establishments and increase the likelihood of reaching at least 2, 3, 4 or 5 years of tenure. The effects are sizable: moving from the 10th to the 90th percentile decreases switching rates by 2.5 percentage points (pp), against a base switching-probability of 13.7%. In line with this, the likelihood of long tenure increases by between 9.2 and 12.0 pp over base probabilities that range from 72.6% (at least 2 years of tenure) to 38.2% (at least 5 years of tenure). Substitutability, in contrast, is associated with higher switching rates and a lower likelihood of reaching long tenures. When moving from the 10th to the 90th percentile, switching rates increase by 1.3 pp and the likelihood of achieving long tenures decreases by between 3.5 and 5.8 pp. Interestingly, the degree to which coworker fit variables increase the likelihood of long tenure are concentrated in columns (2) and (3), but do not strengthen (nor weaken) thereafter. This suggests that, most mismatches are resolved in the first two to three years in a new job.

6 Wage premiums

The analyses in section 4 have shown that a high coworker match is associated with a significant wage premium. In this section, I explore how this wage

[41] For instance, in a study on the Dutch labor market, Groot and Maassen van den Brink (2000) report that workers decrease their overeducation mostly by early-career job switching. Moreover, older workers have been found to exhibit lower quantitative (Leuven et al., 2011), as well as qualitative educational mismatches (Sattinger, 2012).

[42] Regressors in models (2) to (5) are averaged across a worker's entire observed tenure at the establishment.

Table 11: Job-switching regressions

dep. var.:	(1) est. switch	(2) ten.≥2 yrs	(3) ten.≥3 yrs	(4) ten.≥4 yrs	(5) ten.≥5 yrs
cow. match	-0.096***	0.349***	0.442***	0.453***	0.430***
	(0.004)	(0.024)	(0.023)	(0.023)	(0.022)
cow. subst.	0.032***	-0.090***	-0.131***	-0.147***	-0.144***
	(0.003)	(0.009)	(0.010)	(0.010)	(0.009)
log(est. size)	-0.038***	0.076***	0.101***	0.105***	0.102***
	(0.000)	(0.006)	(0.006)	(0.006)	(0.006)
4th polyn. of age?	yes	yes	yes	yes	yes
edu. level dum.?	yes	yes	yes	yes	yes
FE?	yr	yr	yr	yr	yr
R^2	0.024	0.029	0.045	0.052	0.055
# obs.	1,697,715	185,637	185,637	185,637	185,637
# clust.	307,166	68,894	68,894	68,894	68,894
base probability	0.137	0.726	0.564	0.458	0.382

***: p¡0.01; **: p¡0.05; *: p¡0.10. Observations in column (1) are at the worker-establishment-year level, in columns (2) to (5) at the worker-establishment level. The dependent variable in column (1) is an establishment-switching dummy, in columns (2) to (5) a dummy indicating that the worker stays for at least 2, 3, 4 or 5 years with the establishment. The row "base probability" provides estimated probabilities of switching or reaching long tenure for workers with average regressor values. Standard errors, clustered at the worker level (column (1)) or at the establishment level (columns (2) to (5)), in parentheses.

premium relates to other well-known wage premiums, namely, the returns to schooling, the urban wage premium and the large-plant premium.

6.1 Returns to schooling

Returns to schooling in Sweden are low in international comparison. In Appendix A.6, I show that, whereas, in Sweden, college degrees yield a wage-premium of about 45% over a secondary school degree, the comparable premium in the U.S., of college-educated workers over their peers who completed 11th grade, is almost three times as high (127%). To a large extent, this is a result of Sweden's compressed wage structure.[43] Regardless of their overall level, however, returns to schooling may depend on whether or not a worker works with complementary coworkers. After all, if education allows workers to specialize, the higher a worker's own level of education, the more she will depend on coworkers to complement these skills. To investigate this, I divide all worker-year observations into five equally-sized bins by their coworker complementarity,

[43]Table A.4 in appendix A.6 shows that, when regressing, instead of wages, a worker's percentile rank in the overall wage distribution on educational attainment, the premium to college education corresponds to a 21.5 percentiles rise in the overall wage distribution in Sweden, compared to a rise of 35.9 percentiles in the U.S.. See also Hanushek et al. (2015), who show that absolute returns to skills among 22 countries are lowest in Sweden.

$\hat{m}_{wpt}$.[44] Next, interact educational level dummies with these complementarity quintiles in the following equation:

$$log_{10}\left(wage\right) = X_{w,t}\beta_x + Q_{p(w,t),t}\beta_p + \sum_{b=1}^{5}\sum_{l=1}^{6}\beta_{b\times l}B_bL_l + \epsilon_{wt} \tag{27}$$

B_b is a dummy for complementarity quintile b and L_l a dummy for educational level l. $X_{w,t}$ and $Q_{p(w,t),t}$ follow the preferred specification and control for establishment size, year dummies and a 4th order polynomial of age.

Figure 4a depicts the estimated $\beta_{b\times l}$ parameters. Observed returns to education differ substantially by how complementarity-rich a worker's work-environment is. Moreover, as education rises, returns to education become increasingly dependent on coworker complementarities.[45] At the extreme end, the return to college education runs from 18% in the lowest to 88% in the highest complementarity quintile. That is, whereas college-educated workers in the bottom quintile earn wages that are indistinguishable from workers who only took secondary education, their counterparts in the top quintile earn 70% more. Figure 4b repeats the analyses of Figure 4a, controlling for worker fixed effects.[46] Although this reduces returns to education substantially, the dependence of these returns on complementarity remains.

Table 12 shows the converse analyses of Figure 4, namely, how returns to complementarity differ by level of education.[47] Column (1) repeats the preferred specification of Table 5 for workers with over secondary schooling. Column (2) shows estimates when the sample is expanded to include workers with primary or secondary schooling. Columns (3) to (6) show separate results by educational level. On average, an 80 percentiles increase in coworker match is associated with 18.1% higher wages and the same increase in substitutability translates into a 4.8% wage reduction. However, highly educated workers are much more dependent on their coworkers' skill mix. For college-educated workers, an 80 percentiles increase in complementarity translates into a 47.7% higher wage. To put this into perspective, this premium is comparable to the premium college-educated workers earn over workers who only completed primary school.[48] Moreover, college-educated workers also suffer the most pronounced negative effects from substitutability, with an 80 percentiles increase in substitutability being associated with an 18.2% wage-reduction.

[44] Results (reported in Appendix A.5) are qualitatively similar when dividing workers into coworker-match bins *and* substitutability bins.

[45] The exception is workers at the post-graduate (Ph. D.) level. Although there is still a clear complementarity gradient in the returns to post-graduate education, this dependence is weaker than for college-educated workers. However, the sample for post-graduates is rather small and includes many workers in research careers, who might value intellectual achievement over wages.

[46] The variance that allows estimating $\beta_{b\times l}$ parameters comes mostly from workers who switch complementarity quintiles.

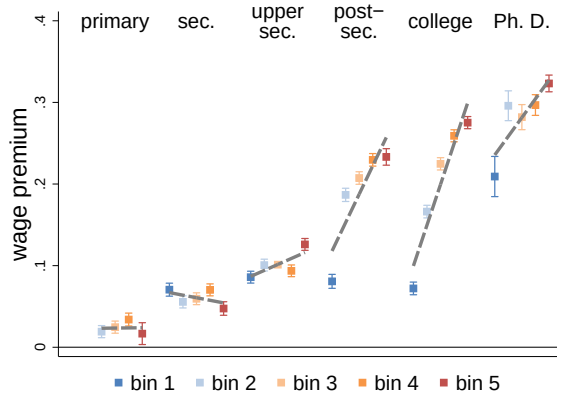
[47] Fixed-effects models by educational level are reported in the Appendix A.5.

[48] Based on model (2), Table 12, this college premium is 52.3% (not shown).

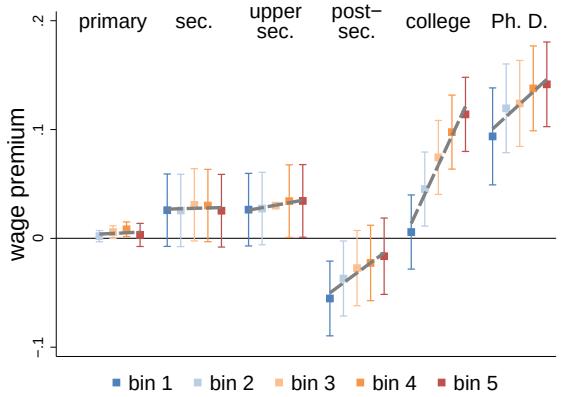
Table 12: Wage regressions by educational level

dep. var.: log(wage)	(1) all	(2) all	(3) upper sec.	(4) post-sec.	(5) college	(6) Ph. D.
cow. match	0.274***	0.257***	0.094***	0.468***	0.642***	0.328***
	(0.0030)	(0.0029)	(0.0033)	(0.0152)	(0.0078)	(0.0383)
cow. subst.	-0.044***	-0.058***	-0.003	-0.036***	-0.182***	-0.127***
	(0.0018)	(0.0017)	(0.0020)	(0.0087)	(0.0052)	(0.0331)
log(est. size)	0.044***	0.044***	0.042***	0.026***	0.050***	0.047***
	(0.0003)	(0.0003)	(0.0003)	(0.0010)	(0.0006)	(0.0032)
4th polyn. age?	yes	yes	yes	yes	yes	yes
edu level dum?	yes	yes				
fixed effects?	yr	yr	yr	yr	yr	yr
R^2	0.300	0.296	0.220	0.286	0.322	0.255
# obs.	2,144,965	2,576,964	1,522,026	184,033	426,827	12,079
# clust.	364,642	440,578	260,310	29,934	80,157	2,280

***: p¡0.01; **: p¡0.05; *: p¡0.10. OLS regressions of log_{10} ($wage$). The model in column (1) uses the sample of workers with at least upper secondary schooling, column (2) is based on the full sample and columns (3) to (6) report results by educational level. Standard errors, clustered at the worker level, are provided in parentheses.



(a) OLS model



(b) Worker fixed effects model

Figure 4: Returns to schooling by complementarity quintile

Educational wage premiums over omitted category (primary school, bottom quintile) by complementarity quintile in log_{10} points. 95% confidence intervals use standard errors clustered at the worker level. Colors code complementarity quintiles (bottom to top: dark blue to dark red). Figure 4a: OLS, Figure 4b: model with worker fixed effects.

6.2 Urban wage premium

Another well-established wage premium is the urban wage premium (UWP), i.e., the positive elasticity of wages with respect to city size. For instance, Combes et al. (2010) find that, for every doubling of a region's population, wages go up by about 5%. The UWP is often attributed to better learning opportunities in larger cities. Here, I explore an alternative explanation: large cities help workers find complementary coworkers.[49]

[49] Appendix A.7 shows that, although the relation between coworker complementarities and a region's size depends on a worker's educational level, for workers with high levels of

Table 13: Urban wage premium, conditional on coworker fit

controls	upper sec. OLS	post-sec. OLS	tertiary OLS	PhD OLS	all OLS
no controls	0.024***	0.037***	0.050***	0.023***	0.047***
	(0.0026)	(0.0044)	(0.0046)	(0.0060)	(0.0035)
$C_{ept} + S_{ept}$ cntrls	0.027***	0.040***	0.035***	0.011**	0.051***
	(0.0029)	(0.0060)	(0.0047)	(0.0056)	(0.0045)
$C^{\#}_{ept} + S^{\#}_{ept}$ cntrls	0.025***	0.029***	0.033***	0.006	0.043***
	(0.0029)	(0.0044)	(0.0048)	(0.0052)	(0.0034)

***: p¡0.01; **: p¡0.05; *: p¡0.10. Wage elasticity with respect to the size of a labor market area (total employment in the area). Standard errors, clustered by labor market area, in parentheses.

If the UWP indeed reflects better matching opportunities in large cities, accounting for coworker fit should reduce the estimated wage elasticity with respect to region size. Table 13 shows that this is indeed the case for workers with high levels of education. The table summarizes how the urban wage premium changes when controlling for coworker fit.[50] The first row shows urban wage premiums that were estimated without any control variables. The second row adds average coworker-match and substitutability as defined in equations (10) and (11) as control variables. The third row uses as control variables the count of well-matching coworkers and coworkers who are close substitutes to the focal worker, as defined in (8) and (9).

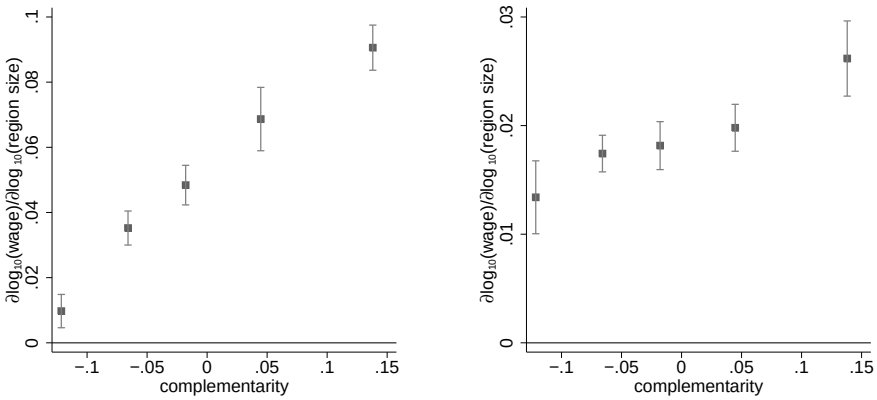
Controlling for average coworker match and substitutability does not change the point estimates of the UWP for workers with low levels of education. However, it does substantially change the observed UWP of workers with college or Ph. D. degrees, by, respectively 30% and 50%. Controlling for the count-based coworker-fit variables further reduces the UWP by 21% for workers with post-secondary degrees, by 34% for workers with a college degree and by 74% for workers with a Ph. D. degree. However, the UWP for workers with upper secondary school degrees is left unchanged.

Although coworker complementarities can only to some extent account for the UWP, they strongly moderate this premium. Figure 5a plots UWP-estimates for workers in different complementarity quintiles. The UWP is much more pronounced for workers who work in highly complementary teams than for those who don't, rising about five-fold between the bottom and the top complementarity quintile. Figure 5b repeats this analysis while controlling for worker fixed

education, coworker complementarities tend to go up with region-size.

[50]The reported results are from regressions of $log_{10}(wage)$ on $log_{10}(region\,size)$, where region size is the number of employees in one of Sweden's 110 labor market areas. These regressions without any worker-level control-variables capture the raw (unconditional) UWP. Adding worker-level control-variables reduces the UWP overall, but does not substantially alter the relative reductions in UWP when adding coworker-fit controls.

effects. Although the average UWP drops, suggesting that the UWP may in part reflect spatial sorting of workers, the interaction pattern with coworker complementarity remains clearly visible.



(a) OLS

(b) worker fixed effects

Figure 5: Urban wage premium by complementarity quintile

Urban wage premium by complementarity quintile, estimated in a regression of the logarithm of wages on the logarithm of labor market size (i.e., total employment in the region), controlling for year dummies and clustering standard errors by labor market area. Figure 5a: OLS; Figure 5b: model with worker fixed effects.

These findings resemble prior findings on heterogeneity in UWP. For instance, Wheeler (2001) finds that the UWP rises monotonically with education and Bacolod et al. (2009) show that the UWP increases with workers' cognitive skills. However, as shown in Figure A.9 of Appendix A.5, the interaction effect in Figure 5a cannot be attributed to education per se, but is also manifest in the subsamples of workers with post-secondary, college and Ph. D. degrees.

6.3 Large-plant premium

The final premium I explore is the premium associated with working in large establishments. Several mechanisms have been put forward for how this large-plant premium (LPP) is generated (e.g., Troske, 1999). For instance, the LPP has been attributed to rent sharing (Weiss, 1966; Mellow, 1982; Akerlof and Yellen, 1990) or efficiency wages to prevent shirking in large establishments (Bulow and Summers, 1986). Others have proposed various types of complementarities to explain the LPP: complementarities between skills and physical capital (Griliches, 1970), between entrepreneurs and workers (Oi, 1983) and assortative matching among workers (Kremer, 1993). In light of the present paper, a related, yet slightly different, interpretation of the LPP arises, namely, that large establishments have more complementary workforces. The force behind this explanation is that larger establishments allow for a greater division of labor, which leads to greater interdependencies among workers with different, yet complementary skills.

Table 14: Large-plant premium

controls	upper sec. OLS	post-sec. OLS	college OLS	Ph. D. OLS	all OLS
no controls	0.040***	0.028***	0.050***	0.040***	0.041***
	(0.0003)	(0.0010)	(0.0007)	(0.0031)	(0.0003)
$C_{ept} + S_{ept}$ cntrls	0.042***	0.026***	0.050***	0.047***	0.044***
	(0.0003)	(0.0010)	(0.0006)	(0.0032)	(0.0003)
$C^{\#}_{ept} + S^{\#}_{ept}$ cntrls	0.035***	-0.000	-0.011***	0.001	0.019***
	(0.0005)	(0.0019)	(0.0011)	(0.0060)	(0.0005)

***: p¡0.01; **: p¡0.05; *: p¡0.10. Wage elasticity with respect to establishment size. First row: analyses with year dummies, a 4th polynomial of age, and, in the last column, educational level dummies as controls, second row furthermore controls for average coworker match and substitutability as defined in (10) and (11), third row controls instead for the number of well-matching coworkers and the number of coworkers that are close substitutes to the focal worker as defined in (8) and (9). Standard errors, clustered by establishment, in parentheses.

If this were indeed the case, the LPP should disappear once coworker complementarities are controlled for. Contradicting this, Table 14 shows that, controlling for the weighted-average coworker-match and substitutability as defined in equations (10) and (11) doesn't affect the wage elasticity with respect to establishment size, regardless of a worker's level of education. However, given that weighted averages were chosen to *ensure* that the coworker-fit variables would be close to uncorrelated with establishment size, this is not surprising. When controlling instead for the number of well-matching coworkers and the number of close substitutes among one's coworkers (i.e., using the indices of equations (8) and (9)), the LPP is reduced by 53% in the full sample. Moreover, the LPP disappears completely for workers with more than upper secondary education. Given that, on average, depending on the educational level, only between 6.4% and 15.6% of coworkers are well-matching and close substitutes make up between just 6.8% and 11.4% of coworkers, it is remarkable that controlling for the presence of these relatively small sets of coworkers can account for the entire LPP.

Moreover, Figure 6, which plots returns to complementarity by establishment-size quintile,[51] shows that the benefits of complementarity rise with establishment size. This finding is in line with the notion that larger establishments apply a finer division of labor, which increases the specialization of workers and, therewith, their dependence on coworkers.

[51] This figure is based on a regression of $log_{10}(wage)$ on workers' age, educational level and coworker complementarity, $\hat{m}_{wpt}$, interacted with dummies for establishment-size quintiles.

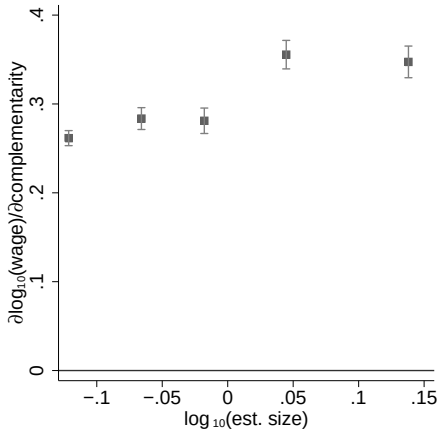


Figure 6: Complementarity premium by establishment-size quintile

7 Conclusion

Division of labor allows workers to specialize, but also makes them dependent on one another. That is, specialization often implies *co-specialization*: coworkers need to acquire *different, yet complementary* expertise. I have quantified these interdependencies in terms of the match and substitutability among coworkers, using Swedish administrative data that describe workers' educational attainment in terms of 491 different educational tracks. Coworker match is measured by how often these tracks co-occur in establishments' workforces, whereas substitutability is measured as the degree to which different educational tracks give access to the same occupations.

The effects of coworker match on wages are positive and substantial. Causal estimates imply that working with well-matching coworkers yields returns of a similar magnitude as having a college degree. Moreover, better coworker matches are associated with lower job-switching rates. In contrast, being easily substituted by coworkers diminishes wages and is associated with elevated job-switching rates. Given the positive wage effects, I have argued that the component of a worker's coworker match that is orthogonal to coworker substitutability can be thought of as a measure of how complementary a worker is to her coworkers. This coworker complementarity rises over the course of a worker's career in a way that closely tracks the Mincer curve. Furthermore, I have shown that well-established wage premiums are to some extent contingent on working with complementary coworkers. For instance, college-educated workers who have few complementary coworkers earn about the same as workers who only completed secondary school. Similarly, the urban wage premium is about nine times larger for workers in the top quintile of the complementarity distribution compared to those in the bottom quintile. Finally, for workers with

post-secondary degrees or higher, the large-plant premium, i.e., the relatively high wages paid by large establishments, can be wholly attributed to the fact that these establishments employ larger numbers of complementary coworkers.

These findings highlight a salient fact of modern societies: high levels of specialization make skilled workers reliant on coworkers who specialize in areas that are complementary to their own field of expertise. This interdependence of coworkers has consequences for how we should think about returns to schooling at a societal level, for the implied coordination challenges in upgrading education systems, and for the role urban labor markets play as places where workers match to coworkers, not just to employers.

References

Acemoglu, D. and J. Angrist (2001). How large are human-capital externalities? evidence from compulsory-schooling laws. In *NBER Macroeconomics Annual 2000, Volume 15*, pp. 9–74. MIT Press.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics*, 255–283.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics 30*, 98–108.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics 121*(1), 343–375.

Arcidiacono, P., J. Kinsler, and J. Price (2013). Productivity spillovers in team production: Evidence from professional basketball. *Iowa state University, Ames, Iowa*.

Bacolod, M., B. S. Blum, and W. C. Strange (2009). Skills in the city. *Journal of Urban Economics 65*(2), 136–153.

Becker, G. S. and K. M. Murphy (1992). The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics 107*(4), 1137–1160.

Bettencourt, L. M. A., H. Samaniego, and H. Youn (2014). Professional diversity and the productivity of cities. *Sci Rep 4*, 5393.

Bulow, J. I. and L. H. Summers (1986). A theory of dual labor markets with application to industrial policy, discrimination, and keynesian unemployment. *Journal of Labor Economics 4*(3, Part 1), 376–414.

Cahuc, P., F. Marque, and E. Wasmer (2008). A theory of wages and labor demand with intra-firm bargaining and matching frictions. *International Economic Review 49*(3), 943–972.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics 128*(3), 967–1015.

Chillemi, O. and B. Gui (1997). Team human capital and worker mobility. *Journal of Labor Economics 15*(4), 567–585.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration Economics*, pp. 15–66. University of Chicago Press.

David, H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *The American Economic Review 103*(5), 1553–1597.

De Fontenay, C. C. and J. S. Gans (2003). Organizational design and technology choice under intrafirm bargaining: Comment. *The American Economic Review 93*(1), 448–455.

Dibiaggio, L., M. Nasiriyar, and L. Nesta (2014). Substitutability and complementarity of technological knowledge and the inventive performance of semiconductor companies. *Research Policy 43*(9), 1582–1593.

Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2005). Comparative advantage, learning, and sectoral wage determination. *Journal of Labor Economics 23*(4), 681–724.

Glaeser, E. L., D. C. Mare, et al. (2001). Cities and skills. *Journal of Labor Economics 19*(2), 316–42.

Griliches, Z. (1970). Notes on the role of education in production functions and growth accounting. In *Education, income, and human capital*, pp. 71–127. NBER.

Griliches, Z. and J. A. Hausman (1986). Errors in variables in panel data. *Journal of econometrics 31*(1), 93–118.

Groot, W. and H. Maassen van den Brink (2000). Skill mismatches in the dutch labor market. *International Journal of manpower 21*(8), 584–595.

Hanushek, E. A., G. Schwerdt, S. Wiederhold, and L. Woessmann (2015). Returns to skills around the world: Evidence from piaac. *European Economic Review 73*, 103–130.

Helsley, R. W. and W. C. Strange (1990). Matching and agglomeration economies in a system of cities. *Regional Science and urban economics 20*(2), 189–212.

Jäger, S. (2016). How substitutable are workers? evidence from worker deaths. Technical report, mimeo.

Kirkebøen, L., E. Leuven, and M. Mogstad (2014). Field of study, earnings, and self-selection. Technical report, National Bureau of Economic Research.

Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics*, 551–575.

Leuven, E., H. Oosterbeek, et al. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education 4*, 283–326.

Mailath, G. J. and A. Postlewaite (1990). Workers versus firms: Bargaining over a firm's value. *The Review of Economic Studies 57*(3), 369–380.

Mas, A. and E. Moretti (2009). Peers at work. *The American Economic Review 99*(1), 112–145.

Mellow, W. (1982). Employer size and wages. *The Review of Economics and Statistics*, 495–501.

Milgrom, P. and J. Roberts (1995). Complementarities and fit strategy, structure, and organizational change in manufacturing. *Journal of accounting and economics 19*(2), 179–208.

Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *The American Economic Review 94*(3), 656–690.

Oi, W. (1983). The fixed employment costs of specialized labor. In *The measurement of labor cost*, pp. 63–122. University of Chicago Press.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers 3*(2), 135–146.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015). *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]*. Minneapolis: University of Minnesota.

Sattinger, M. (2012). Assignment models and quantitative mismatches. In *London: Expert Workshop" Skill Mismatch and Firm Dynamics: Integrating Skills with the World of Work*.

SCB (2011). Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier (lisa) 1990-2009. arbetsmarknads- och utbildningsstatistik 2011:4. Statistiska centralbyrån/Statistics Sweden.

Stole, L. A. and J. Zwiebel (1996a). Intra-firm bargaining under non-binding contracts. *The Review of Economic Studies 63*(3), 375–410.

Stole, L. A. and J. Zwiebel (1996b). Organizational design and technology choice under intrafirm bargaining. *The American Economic Review*, 195–222.

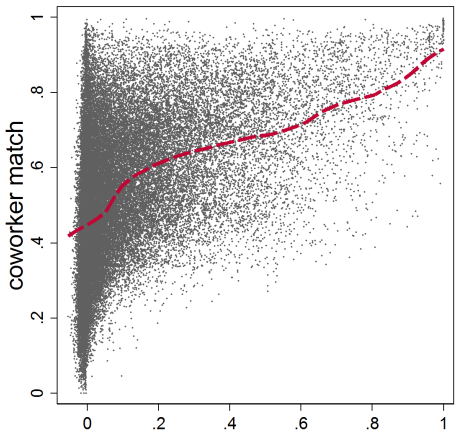


Figure A.1: Coworker match against substitutability scatterplot
Scatterplot of educational coworker match against educational substitutability. Observations are pairs of educational tracks. The dashed line shows the prediction of a kernel regression.

Troske, K. R. (1999). Evidence on the employer size-wage premium from worker-establishment matched data. *Review of Economics and Statistics 81*(1), 15–26.

Weiss, L. W. (1966). Concentration and labor earnings. *The American Economic Review 56*(1/2), 96–117.

Wheeler, C. H. (2001). Search, sorting, and urban agglomeration. *Journal of Labor Economics 19*(4), 879–899.

A Appendix for online publication: Additional results and derivations

A.1 Substitutability and coworker match of educational pairs

As explained in section 3.2 and shown in Tables 1 and 2, educations that are close substitutes are also often overrepresented in coworker counts. Figure A.1 shows a scatterplot of coworker match against substitutability. The two quantities are strongly and positively correlated. Interestingly, however, the scatterplot is by-and-large triangular: although educations that are substitutes, also often co-occur in establishments' workforces, the reverse does not hold. Apparently, establishments hire teams that cover a wide range of skills at different levels, but for two workers to be able to do the same job they must be trained at a similar level.

To further illustrate the relation between coworker match and substitutability, Table A.1 shows the educational tracks with the highest coworker match

Table A.1: Top 10 educational pairs: coworker match, controlling for substitutability

rank	edu. (1)	edu. (2)	match	subst.
1	5: Professional officers	5: Programme for air transport	0.969	0.016
2	5: Tactical military	5: Programme for air transport	0.966	0.019
3	5: Professional officers	5: Police service	0.953	0.010
4	5: Tactical military	5: Police service	0.950	0.009
5	5: Programme for air transport	5: Personal services, other	0.955	0.023
6	5: Professional officers	5: Programme for water transport	0.953	0.029
7	5: Tactical military	5: Programme for water transport	0.948	0.031
8	5: Police service	5: Programme for air transport	0.927	0.001
9	5: Police service	5: Programme for water transport	0.928	0.010
10	5: Dental surgery	6: Other medicine	0.926	0.007

Idem Table 1, but now the top 10 highest coworker match values, controlling for substitutability.

values, controlling for how substitutable these tracks are. To be precise, it shows the residual of the non-parametric regression of $c_{ee'}$ on $s_{ee'}$ shown in Figure A.1. The table shows that the coworker match and coworker substitutability of a pair of educations can differ substantially from one another.

Figure A.2 explores to what extent coworker match and substitutability are driven by similarity in the *content* or similarity in the *level* of educational tracks. Similarity in content is measured by the overlapping leading digits of the content code, whereas similarity in levels is expressed as the absolute difference between levels (which run from 1 for primary education to 6 for Ph. D. degrees). The figure shows that, although substitutability and coworker match both rise as educational levels and educational content become more similar. However, to be substitutes, educational tracks require more or less the same educational level, whereas workers can work together (i.e., have a high coworker match), even if they hold degrees of different levels. In fact, compared to substitutability, for coworker relations, similarity in content is relatively more important than similarity in levels. For instance, teams may combine workers with theory-oriented training and more applied vocational training, as long as the workers are trained in the same field. Indeed, in an ANOVA of substitutability and coworker match on similarity in educational levels and educational contents, similarity in educational contents accounts for two-thirds of the explained variation in substitutability, but three-quarters of the explained variation in coworker match.

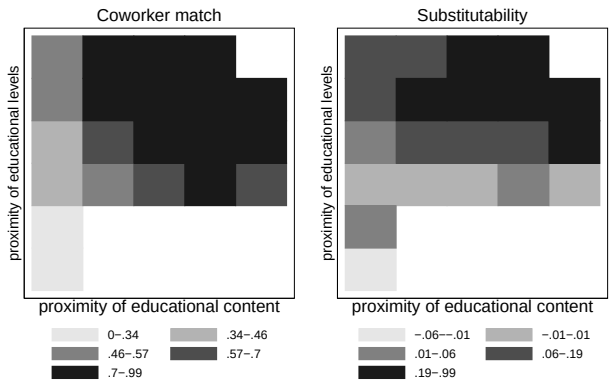


Figure A.2: Content and levels of educational tracks

The plots show the average coworker match (left panel) and substitutability (right panel) for pairs of educational tracks at different degrees of proximity in educational levels and contents. Proximity in educational levels is calculated as 5 minus the absolute difference in the levels of the educational tracks within a pair. Proximity in educational content is calculated as the number of leading digits two educational tracks have in common. Different shadings reflect increasing quintiles in coworker match or substitutability.

A.2 Coworker match and substitutability by industry

Do economic activities differ in terms of coworker fit? Figure A.3 explores this by plotting the average coworker match for workers in a given industry[52] against their coworker substitutability. That is, it plots the following quantities $\bar{C}_s$ and $\bar{S}_s$:

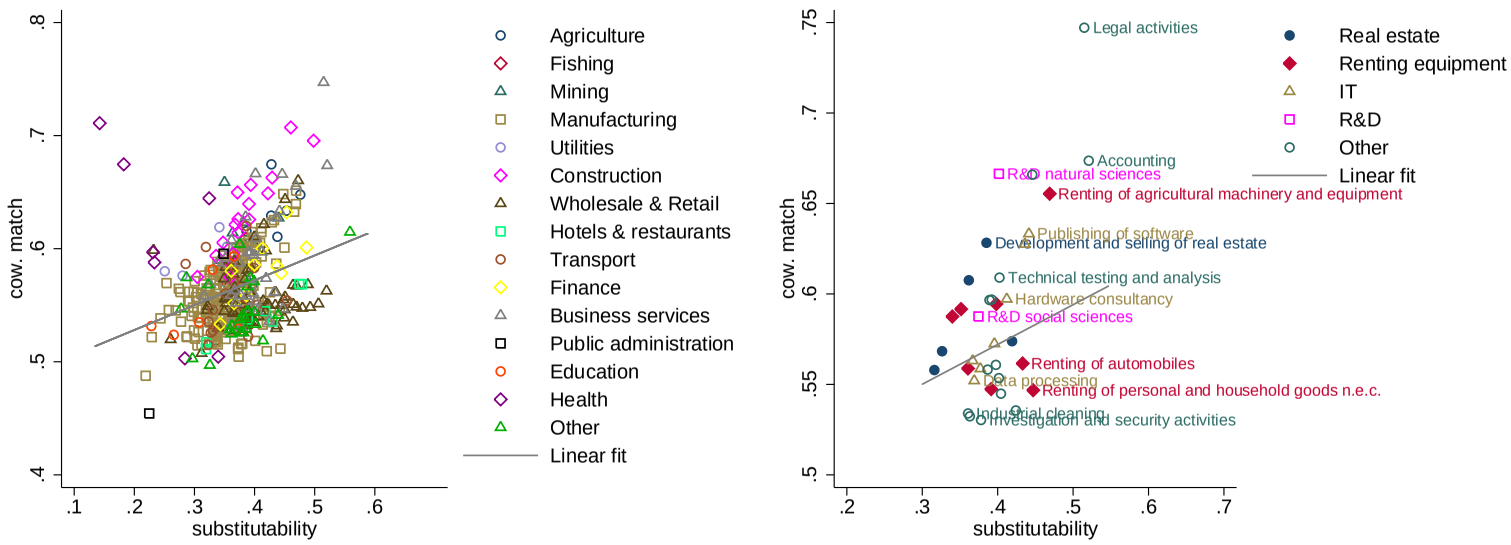
$$\bar{C}_s = \frac{1}{|W_s|} \sum_{(w,t)\in W_s} C_{e(w,t)p(w,t)} \tag{28}$$

$$\bar{S}_s = \frac{1}{|W_s|} \sum_{(w,t)\in W_s} S_{e(w,t)p(w,t)} \tag{29}$$

where W_s represents the set of worker-year observations in establishments of industry s. The panel on the left displays all industries in the economy, whereas the panel on the right focuses on industries in business services.

A high average coworker match implies that an industry's establishments tend to combine workers who generally often work together. A high average substitutability means that an industry's establishments employ workforces of workers who are very similar to one another. Figure A.3 once again highlights that who works with whom is not independent of who can substitute whom. Many industries that display high average coworker-match values also hire groups of workers who can substitute one another. However, there are marked differences between industries. For instance, the workforces of establishments in health care exhibit both high coworker-match levels and low levels

[52]Industries are classified at the 4-digit level of the Swedish SNI 2002 classification, which corresponds to the European NACE Revision 1.1 classification system.



(a) All industries

(b) Business services

Figure A.3: Coworker match and substitutability by industry

The left panel shows the average of $C_{e(w,t)p(w,t)}$ against the average of $S_{e(w,t)p(w,t)}$ in the estimation sample as defined in (10) and (11) across all workers of a given industry. The right panel shows the same graph, but restricted to industries in business services. The gray line represents an (unweighted) linear fit, where in both panels, this fit is based on the full set of industries. In the right panel, labels are shown for a number of industries at high or low levels of substitutability or coworker match and for industries that strongly exceed the linear fit.

of substitutability. In contrast, industries in retail and wholesale combine low coworker matches with high degrees of substitutability, whereas business services are among the most coworker-match-rich industries. Within this sector (rightmost panel) many industries lie above the regression line of average coworker match on average substitutability. In particular, employers in human-capital-intensive business services, like legal services, R&D, software publishing, hardware consultancy and technical testing, all tend to hire teams of well-matching workers. The lowest coworker-match levels (as well as relatively high levels of substitutability) are found in the workforces of cleaning, security and rental agencies.

Teams of workers who are closely matched but not close substitutes are typically associated with the highly skilled and specialized labor forces of high-technology industries, whereas homogeneous teams (i.e., teams whose workers can easily substitute one another) are often found in less skill-intensive activities, such as retail, hotels, restaurants, and cleaning services. However, there are important exceptions. For instance, many industries in construction employ workers who form close matches, but who are not each other's substitutes. Although the - often vocational - skills in this sector, are typically taught at lower levels, construction workers can have very different types of expertise (e.g, carpenters, masons, painters, electricians, construction engineers, and so on).

A.3 Error variance

Measurement error at the educational pair level

To arrive at an estimate of the variance in coworker match, consider the following Binomial model for the probability of observing $n_{ee'}$ co-occurrences between education e and e':

$$Pr\left[N_{ee'} = n_{ee'} \left| N_{..} = n_{..}, \Pi_{ee'} = \pi_{ee'}\right.\right] = \begin{pmatrix} n_{..} \\ n_{ee'} \end{pmatrix} \pi_{ee'}^{n_{ee'}} \left(1 - \pi_{ee'}\right)^{n_{..} - n_{ee'}} \tag{A.1}$$

Using Bayes' law, we get:

$$Pr\left[\Pi_{ee'} = \pi_{ee'} \left| N_{..} = n_{..}, N_{ee'} = n_{ee'}\right.\right] = \frac{Pr\left[N_{ee'} = n_{ee'} \left| N_{..} = n_{..}, \Pi_{ee'} = \pi_{ee'}\right.\right] Pr\left[\Pi_{ee'} = \pi_{ee'} \left| N_{..} = n_{..}\right.\right]}{\int_0^1 Pr\left[N_{ee'} = n_{ee'} \left| N_{..} = n_{..}, \Pi_{ee'} = \pi_{ee'}\right.\right] Pr\left[\Pi_{ee'} = q_{ee'} \left| N_{..} = n_{..}\right.\right] dq_{ee'}}$$

Choosing the $BETA\left[\alpha_{ee'}, \beta_{ee'}\right]$ distribution, the Binomial distribution's conjugate prior, as a prior for $\Pi_{ee'}$, this becomes:

$$Pr\left[\Pi_{ee'} = \pi_{ee'} \left| N_{..} = n_{..}, N_{ee'} = n_{ee'}\right.\right] = \frac{Pr\left[N_{ee'} = n_{ee'} \left| N_{..} = n_{..}, \Pi_{ee'} = \pi_{ee'}\right.\right] \pi_{ee'}^{\alpha_{ee'}-1} \left(1 - \pi_{ee'}\right)^{\beta_{ee'}-1} \frac{\Gamma(\alpha_{ee'}+\beta_{ee'})}{\Gamma(\alpha_{ee'})\Gamma(\beta_{ee'})}}{\int_0^1 Pr\left[N_{ee'} = n_{ee'} \left| N_{..} = n_{..}, \Pi_{ee'} = q_{ee'}\right.\right] q_{ee'}^{\alpha_{ee'}-1} \left(1 - q_{ee'}\right)^{\beta_{ee'}-1} \frac{\Gamma(\alpha_{ee'}+\beta_{ee'})}{\Gamma(\alpha_{ee'})\Gamma(\beta_{ee'})} dq_{ee'}} \tag{A.2}$$

where Γ represents the gamma function. Using the expression for the binomial probability density function (pdf), (A.2) simplifies to:

$$Pr\left[\Pi_{ee'} = \pi_{ee'} \left| N_{..} = n_{..}, N_{ee'} = n_{ee'}\right.\right] = \frac{\pi_{ee'}^{n_{ee'}+\alpha_{ee'}-1}\left(1-\pi_{ee'}\right)^{n_{..}-n_{ee'}+\beta_{ee'}-1}}{\int_0^1 q_{ee'}^{n_{ee'}+\alpha_{ee'}-1}\left(1-q_{ee'}\right)^{n_{..}-n_{ee'}+\beta_{ee'}-1} dq_{ee'}} \tag{A.3}$$

Because the denominator of (A.2) has to integrate to one (it is an integral of a pdf over its domain), we get:

$$\int_0^1 t^{n_{ee'}+\alpha_{ee'}-1}\left(1-t\right)^{n_{..}-n_{ee'}+\beta_{ee'}-1} dt = \frac{\Gamma\left(n_{ee'}+\alpha_{ee'}\right)\Gamma\left(n_{..}-n_{ee'}+\beta_{ee'}\right)}{\Gamma\left(n_{..}+\alpha_{ee'}+\beta_{ee'}\right)} \tag{A.4}$$

Substituting (A.4) into (A.3), yields:

$$Pr\left[\Pi_{ee'} = \pi_{ee'} \left| N_{..} = n_{..}, N_{ee'} = n_{ee'}\right.\right] =$$
$$\pi_{ee'}^{n_{ee'}+\alpha_{ee'}-1}\left(1-\pi_{ee'}\right)^{n_{..}-n_{ee'}+\beta_{ee'}-1}\frac{\Gamma\left(n_{..}+\alpha_{ee'}+\beta_{ee'}\right)}{\Gamma\left(n_{ee'}+\alpha_{ee'}\right)\Gamma\left(n_{..}-n_{ee'}+\beta_{ee'}\right)} \tag{A.5}$$

which describes a $BETA\left[n_{ee'}+\alpha_{ee'}, n_{..}-n_{ee'}+\beta_{ee'}\right]$ distribution. In other words, the posterior distribution of $\Pi_{ee'}$ is:

$$\Pi_{ee'} \left| N_{..} = n_{..}, N_{ee'} = n_{ee'}\right. \sim BETA\left[n_{ee'}+\alpha_{ee'}, n_{..}-n_{ee'}+\beta_{ee'}\right] \tag{A.6}$$

This leaves the task of choosing parameters $\alpha_{ee'}$ and $\beta_{ee'}$ such that they reflect a plausible prior for the mean and variance of $\Pi_{ee'}$. To arrive at such priors, assume that the total number of co-occurrences in which educations e and e' participate are given. In other words, think of co-occurrences as arising from a process in which each time an education e is present in an establishment, it draws a random second education from the pool of educational presences. Moreover, because the total number of co-occurrences in the economy $N_{..} \ggg N_{ee'}$ is large for any (e, e'), I take $N_{..}$ as fixed. Consequently, co-occurrences follow a Hypergeometric distribution, with the following prior means and variances for $\Pi_{ee'}$:

$$E\left[\Pi_{ee'}\right] = E\left[\frac{N_{ee'}}{N_{..}}\right] = \frac{1}{N_{..}}E\left[N_{ee'}\right] \backsimeq \frac{1}{N_{..}}N_{e.}\frac{N_{.e'}}{N_{..}} \tag{A.7}$$

$$V\left[\Pi_{ee'}\right] = \frac{1}{N_{..}^2}V\left[N_{ee'}\right] \backsimeq \frac{1}{N_{..}^2}\frac{N_{e.}N_{.e'}\left(N_{..}-N_{e.}\right)\left(N_{..}-N_{.e'}\right)}{N_{..}^2\left(N_{..}-1\right)} \tag{A.8}$$

where $\backsimeq$ indicates an equality by assumption of the Hypergeometric data generating process. The $BETA\left[\alpha_{ee'}, \beta_{ee'}\right]$ distribution implies:

$$E\left[\pi_{ee'}\right] = \mu_{ee'} = \frac{\alpha_{ee'}}{\alpha_{ee'}+\beta_{ee'}} \tag{A.9}$$

$$V\left[\pi_{ee'}\right] = \sigma_{ee'}^2 = \frac{\alpha_{ee'}\beta_{ee'}}{\left(\alpha_{ee'}+\beta_{ee'}\right)^2\left(\alpha_{ee'}+\beta_{ee'}+1\right)} \tag{A.10}$$

Solving for $\alpha_{ee'}$ and $\beta_{ee'}$, yields:

$$\alpha_{ee'} = \frac{\mu_{ee'}^2}{\sigma_{ee'}^2}\left(1-\mu_{ee'}\right) - \mu_{ee'} \tag{A.11}$$

$$\beta_{ee'} = \mu_{ee'}\left(\frac{\left(1-\mu_{ee'}\right)^2}{\sigma_{ee'}^2} + 1\right) - 1 \tag{A.12}$$

Equations (A.5), (A.7), (A.8), (A.11) and (A.12) now define a posterior expectation, $\hat{\Pi}_{ee'}^{post}$, of $\Pi_{ee'}$ for each educational pair.

Measurement error at the worker-establishment level

Equation (19) can be used to estimate the error-variance in worker-establishment coworker-matches for each observation in the data. Dropping subscripts w and t for notational clarity, collecting all regressors other than $C_{e(w,t),p(w,t)}$ in vector Z and letting y denote $log_{10}\left(wage_{wt}\right)$, (21) can be written as:

$$y = \gamma_b C + Z\Theta_b + \epsilon \tag{A.13}$$

where C is a mismeasured version of the underlying quantity $\tilde{C}$:

$$C = \tilde{C} + \eta$$

where η is a measurement error, which I will assume to be uncorrelated with the true coworker match, $\tilde{C}$, the regressors in Z and the disturbance term ϵ. Furthermore, I assume that the real relation between y and C is given by:

$$y = \gamma\tilde{C} + Z\Theta + \tilde{\epsilon} \tag{A.14}$$

That is, the effect of coworker match is constant across error bins.[53] The estimate of γ_b in (A.13) will be biased and the size of the bias depends on $V\left[\eta\right]$ in bin b. Given the Frisch-Waugh-Lovell theorem, the OLS estimate of γ_b can be written as:

$$\hat{\gamma}_b = \frac{Cov\left[\hat{C}, y\right]}{V\left[\hat{C}\right]} \tag{A.15}$$

where $\hat{C}$ is the residual of a regression of C on Z:

$$C = Z\Theta + \hat{C} \tag{A.16}$$

Similarly, $\hat{\tilde{C}}$ represents the residual of a regression of $\tilde{C}$ on Z:

$$\tilde{C} = Z\Theta + \hat{\tilde{C}} \tag{A.17}$$

[53] Although this will not be imposed in the empirical analyses, for convenience, I also assume that other parameters are constant across bins, i.e., $\Theta_b = \Theta$ for all b.

Because η is uncorrelated with the columns of Z, (A.16) and (A.17) must have the same parameters, Θ. Consequently:

$$\hat{C} = \hat{\tilde{C}} + \eta \tag{A.18}$$

Using (A.18) and (A.14), the numerator of (A.15) can be written as:

$$Cov\left[\hat{C}, y\right] = Cov\left[\hat{\tilde{C}} + \eta, \gamma\tilde{C} + Z\Theta + \epsilon\right]$$

$$\begin{aligned} Cov\left[\hat{C}, y\right] = \gamma Cov\left[\hat{\tilde{C}}, \tilde{C}\right] + \sum_k \Theta_k Cov\left[\hat{\tilde{C}}, Z_k\right] + Cov\left[\hat{\tilde{C}}, \epsilon\right] \\ + \gamma Cov\left[\eta, \tilde{C}\right] + \sum_k \Theta_k Cov\left[\eta, Z_k\right] + Cov\left[\eta, \epsilon\right] \end{aligned} \tag{A.19}$$

where k indexes columns of matrix Z. Except for the first term, all terms in (A.19) are equal to zero.[54] We therefore get:

$$Cov\left[\hat{C}, y\right] = \gamma Cov\left[\hat{\tilde{C}}, \tilde{C}\right]$$

which, using (A.17) can be written as:

$$Cov\left[\hat{C}, y\right] = \gamma Cov\left[\hat{\tilde{C}}, \hat{\tilde{C}} + Z\Theta\right]$$

Given that $\hat{\tilde{C}}$ is the residual of a regression on Z, this becomes:

$$Cov\left[\hat{C}, y\right] = \gamma V\left[\hat{\tilde{C}}\right] \tag{A.20}$$

Using (A.20) in (A.15), the following expression for $\hat{\gamma}_b$ results:

$$\hat{\gamma}_b = \gamma \frac{V\left[\hat{\tilde{C}}\right]}{V\left[\hat{C}\right]}$$

Using $V\left[\hat{C}\right] = V\left[\hat{\tilde{C}}\right] + V\left[\eta\right]$, this can be written as:

$$\hat{\gamma}_b = \gamma\left(1 - \frac{V\left[\eta\right]}{V\left[\hat{C}\right]}\right), \tag{A.21}$$

which is (22) in the main text.

[54] The second term equals zero, because $\hat{\tilde{C}}$ is the residual of a regression on Z and therefore orthogonal to each column of Z. The third term equals zero because $\hat{\tilde{C}} = \hat{C} + \eta$ and $\hat{C} \perp \epsilon$ because $\hat{C}$ is a linear combination of the regressors in (A.13) and $\eta \perp \epsilon$ by assumption. The fourth to sixth terms equal zero by the assumption that measurement errors are uncorrelated to regressors, the true, underlying coworker match, and to the residual in (A.13).

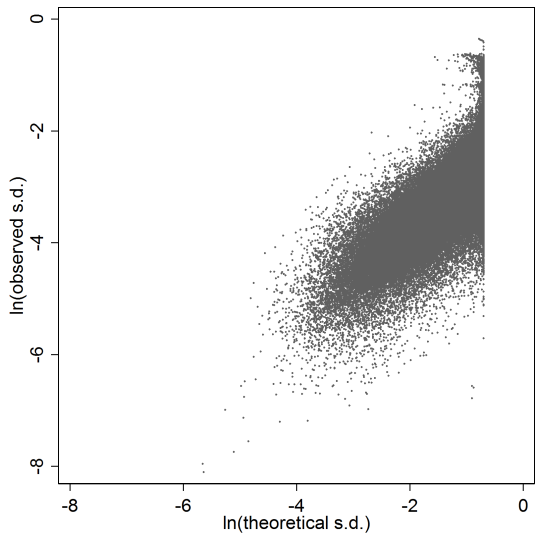


Figure A.4: Coworker-match standard-deviations (observed vs theoretical)
Observations are educational pairs. The vertical axis shows the standard deviation of coworker match estimates across years in the period 2001-2010. The horizontal axis shows the theoretical standard deviation calculated using (19).

Empirical magnitude of measurement error (educational pairs)

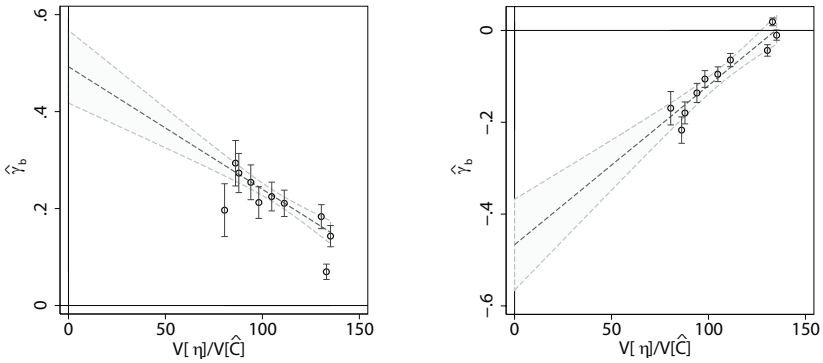
Figure A.4 shows a scatterplot of (log-transformed) empirical (timeseries) standard deviations against theoretical standard deviations of coworker matches between educational pairs. The estimated slope of 1.06 is very close to 1 with an R^2 of 0.60. This shows that the Bayesian model of section A.3 performs surprisingly well. However, the observed standard deviations are, on average, about a fifth of the theoretical ones. A possible explanation for this is that, because teams are relatively stable, the empirical timeseries variance of coworker match underestimates the measurement error. However, given that the extrapolation to error-free parameter estimates is unaffected by a uniform scaling of the error variance, this is inconsequential.

Homoscedastic measurement errors

In section 4.2, I used the fact that error variances can be estimated for each educational pair. As a robustness check, I show here the results when the effect-extrapolation to arrive at error-free estimates assumes homoscedastic measurement errors. In the homoscedastic case, $V\left[c_{e(w,t)e'}\right] = \sigma_\eta^2$ for all $e\left(w,t\right)$ and e', equation (20) simplifies to:

$$V\left[C_{e(w,t),p(w,t)}\right] = \sigma_\eta^2 \sum_{e'} \left(\frac{E_{e'p(w,t)t} - 1\left(e' = e\left(w,t\right)\right)}{\sum_{e''} E_{e''p(w,t)t} - 1} \right)^2 \tag{A.22}$$

In other words, the difference in error-variance across worker-establishment



(a) Coworker match

(b) Substitutability

Figure A.5: Estimated effect of coworker-match by error-variance bin (homoscedasticity)

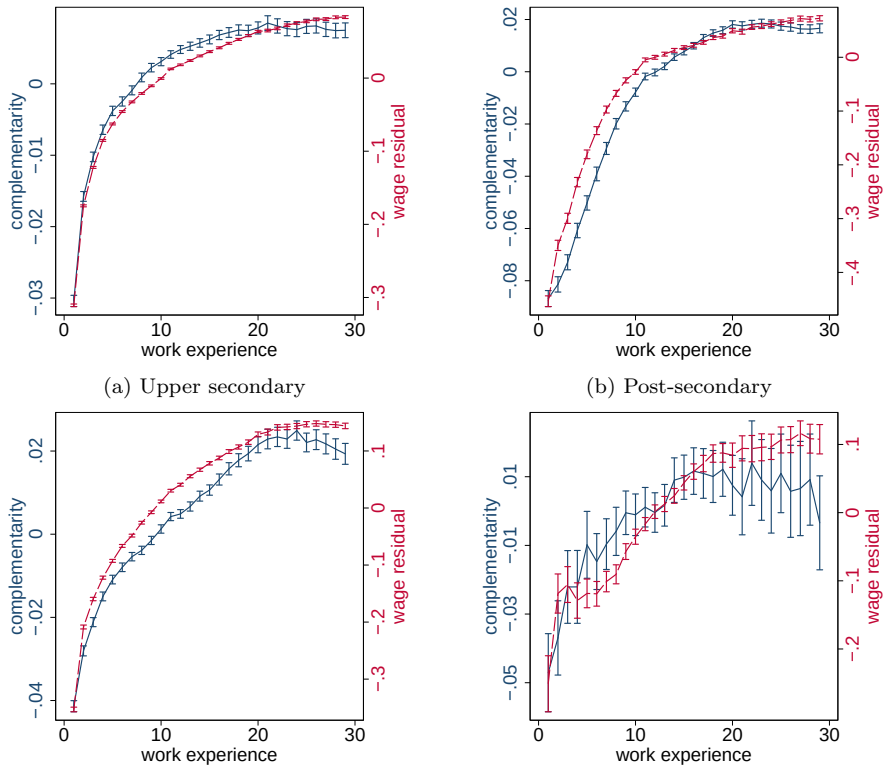
$\hat{\gamma}_b$ (depicted on the vertical axis) is the estimated effect of coworker match (Figure A.5a) and substitutability (Figure A.5b) in a regression of $log_{10}(wage)$ on coworker match, substitutability, the logarithm of establishment size, year-dummies, educational-track dummies, a fourth order polynomial of age and the shares of workers with primary, secondary, upper secondary, post-secondary, college and Ph. D. degrees. This regression is repeated for each decile of error-variances for $C_{e(w,t)p(w,t)}$ as calculated in (A.22). The horizontal axis depicts the average $\frac{V[\eta]}{V[\hat{C}]}$ in bin b, where error-variance bins are based on equation (A.22). The dashed line depicts the regression line of $\hat{\gamma}_b$ on $\frac{V[\eta]}{V[\hat{C}]}$, including a 95% confidence interval. To downweight outliers, this trend line is constructed with robust regression using biweights.

observations is driven by the sum of squared educational employment shares in the establishment. Although this approach ignores heterogeneity in precision when estimating $c_{ee'}$, it still allows some (albeit imperfect) sorting of observations by their coworker-match error-variance. Figure A.5 shows the results when bins are created using (A.22) instead of (20).[55] The implied unbiased coworker-match effects are all but indistinguishable from those based on heteroscedastic measurement errors in Figure 2.

A.4 Reduced-form results in static establishments

The identification strategy in section 4.3 requires that there are no direct effects of a shift in the predicted number of local graduates on changes in wages. For the exclusion restriction to hold, the effect of this instrument on wages must run completely through the endogeneous variable, i.e., through a change in coworker match in a worker's own establishment. To investigate the plausibility of this assumption, I focus on a sample of workers who work in *static establishments*, i.e., establishments in which no workers leave or enter in the year for which

[55] Note that (A.22) yields error-variances up to a scaling factor σ_η^2. As a consequence, $\frac{V[\eta]}{V[\hat{C}]}$ need not lie between 0 and 1.



(a) Upper secondary

(b) Post-secondary

(c) College

(d) Ph. D.

Figure A.6: Change in complementarity with work experience by level of education

Idem Figure 3a. Panels show relation between complementarity (solid blue) and wage residuals (dashed red) by workers' educational level.

the change in coworker match is measured. Table A.2 presents reduced-form estimates when using the sample of workers in static establishments for all models presented in Table 9. In none of these models do instruments have a statistically significant direct effect on wage changes. Consequently, this exercise does not yield evidence of a violation of the exclusion restriction in IV estimates.

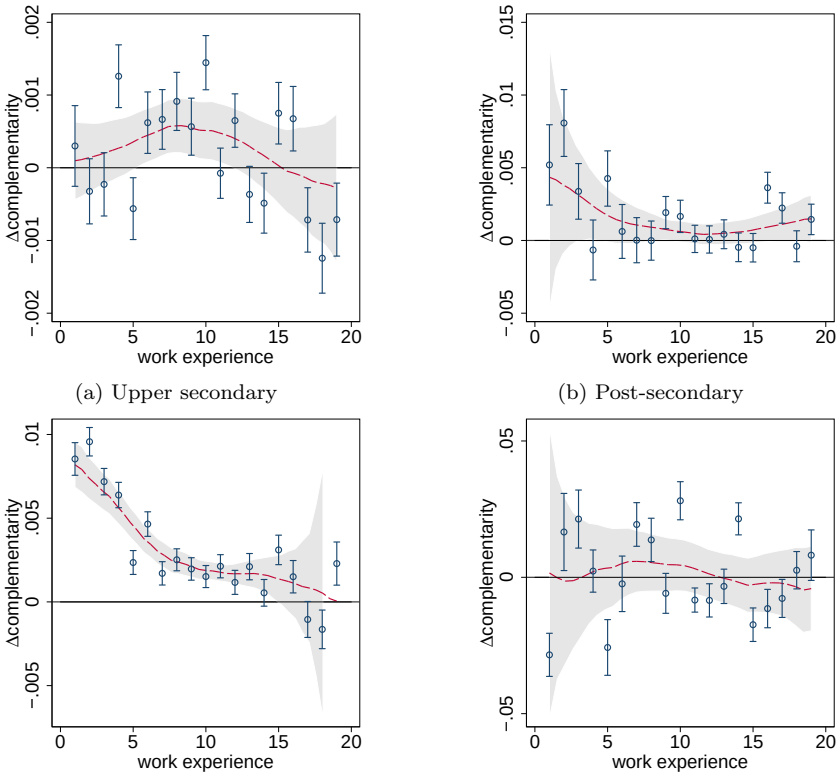
A.5 Additional education-specific analyses

In this section, I report additional results for selected analyses in the main text. Figures A.6 and A.7 repeat the analyses on the evolution of coworker complementarity over a worker's career described in section 5.1 (Figures 3a and 3b), in subsamples by workers' educational level. These graphs show that the similarity in the shapes of the complementarity and wage curves reported for the full sample in the main text is also apparent in these subsamples.

Table A.2: Reduced-form estimates, static-establishment sample

model:	(1) rf	(2) rf	(3) rf	(4) rf	(5) rf
IV - matching grad. (0.1%, m)	-0.0012	-0.0053	-0.0078**	-0.0045	
	(0.0069)	(0.0036)	(0.0039)	(0.0034)	
IV - matching grad. (1%, m)	0.0005	0.0064**	0.0062*	0.0053	
	(0.0062)	(0.0030)	(0.0033)	(0.0034)	
IV - matching grad. (0.1%, r)	0.0041	-0.0010	-0.0020	-0.0028	
	(0.0141)	(0.0096)	(0.0100)	(0.0096)	
IV - matching grad. (1%, r)	-0.0212	0.0073	0.0056	0.0027	
	(0.0187)	(0.0078)	(0.0087)	(0.0085)	
IV - matching grad. (1%, r inc.)					-0.0006
					(0.0064)
4th polyn. of age?	yes	yes	yes	yes	yes
fixed effects?	yr×m×i	yr	yr×m	yr×i	yr×m×i
F-stat.	0.43	1.32	1.39	0.87	0.01
p-value	0.784	0.260	0.234	0.481	0.932
# obs.	19,688	35,685	35,611	35,017	19,688
# clust.	12,303	22,225	22,151	21,867	12,303

***: p¡0.01; **: p¡0.05; *: p¡0.10. Reduced-form estimates in a sample of workers in static establishments. The dependent variable is the logarithm of the annualized wage growth between years $t-1$ and $t+1$. Model (1) shows reduced-form estimates corresponding to models (1) and (5) in Table 9, models (2), (3), (4) and (5) correspond to models (2), (3), (4) and (6) in Table 9. Row "F-stat." contains the value of the F-statistic against the null-hypothesis that the joint-significance of all instruments is zero, with the corresponding p-value below (row "p-value"). Standard errors, clustered at the year-municipality-educational track level, in parentheses.



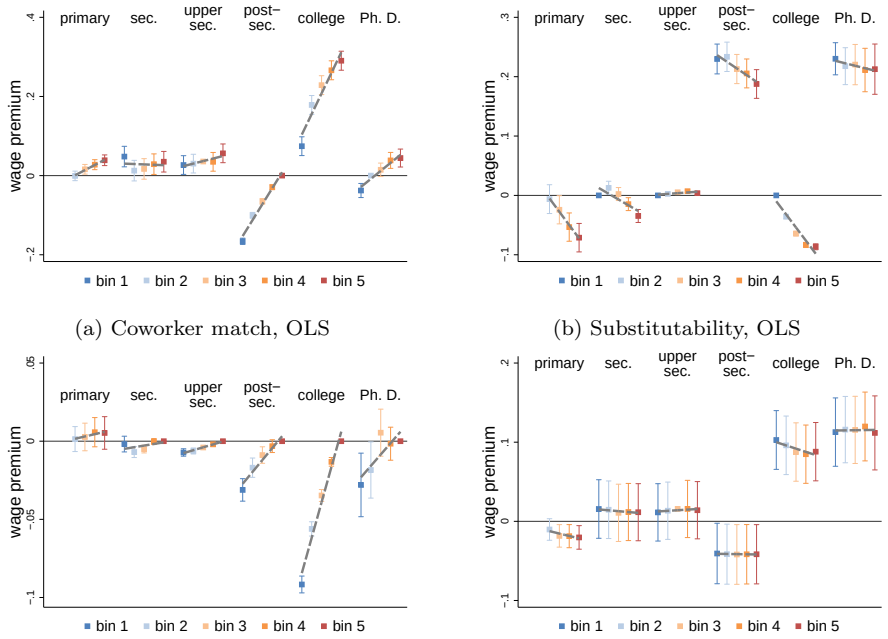
(a) Upper secondary

(b) Post-secondary

(c) College

(d) Ph. D.

Figure A.7: Change in complementarity at job switch by level of education
Idem Figure 3b. Panels depict data for workers with different levels of education.



(a) Coworker match, OLS

(b) Substitutability, OLS

(c) Coworker match, worker fixed effects

(d) Substitutability, worker fixed effects

Figure A.8: Education premium by coworker match and substitutability quintiles

Table A.3 reports results from wage regressions with worker and worker-establishment fixed effects by education level as a complement to the OLS results in Table 12. In line with the results in the main text, the effects of coworker fit (especially the ones of coworker match) tend to strengthen with rising levels of education.

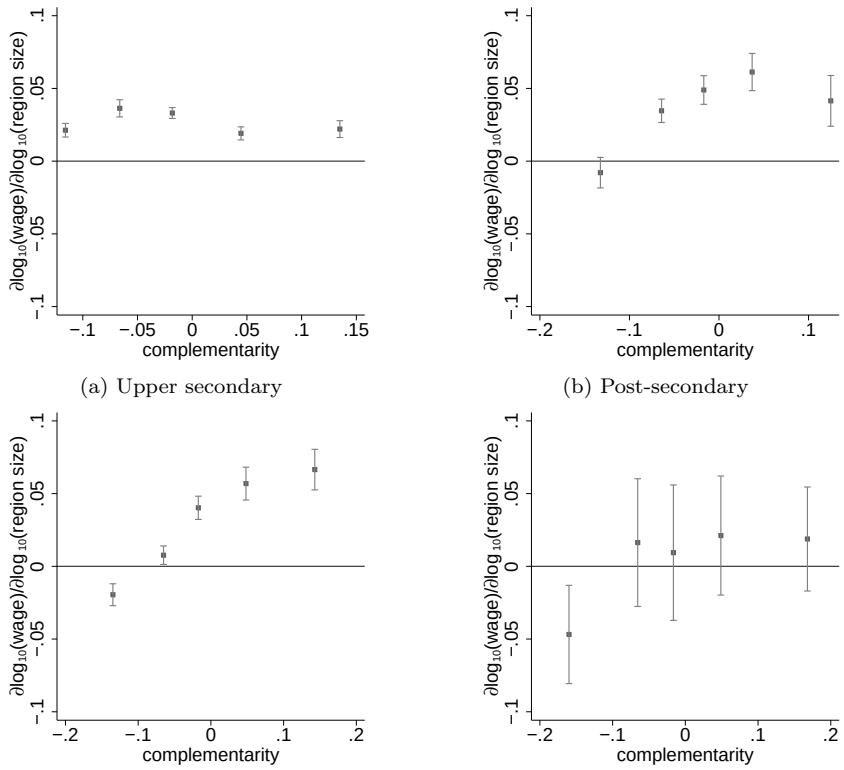
Figures A.8a/A.8b and A.8c/A.8d show how returns to schooling depend on coworker complementarity as in Figure 4a and 4b. However, instead of interacting educational levels with coworker complementarity ($\hat{m}_{wpt}$) quintiles, these levels are interacted with both coworker-match and substitutability quintiles. In these figures, the overall benefits of a higher level of education get distributed between the interactions with coworker match and substitutability. The separation between the two elements of coworker fit shows that coworker match tends to increase the returns to a given level of education, whereas substitutability tends to decrease these returns. These patterns are particularly robust for coworker match, whereas, for substitutability, they are only visible in OLS regressions.

Figures A.9 and A.10 repeat the analyses on the UWP of Figure 5 by educational level. The general pattern of rising urban wage premiums with increasing levels of coworker complementarity shown in Figure 5 is also observed within

Table A.3: Wage regressions by educational level

dep. var.: log(wage)	(1) upper sec.	(2) upper sec.	(3) post-sec.	(4) post-sec.	(5) college	(6) college	(7) Ph. D.	(8) Ph. D.
cow. match	0.046***	0.016*	0.099***	-0.013	0.276***	0.104***	0.101*	0.015
	(0.0058)	(0.0074)	(0.0191)	(0.0247)	(0.0126)	(0.0210)	(0.0502)	(0.0647)
cow. subst.	-0.014***	-0.005	-0.010	-0.005	-0.094***	-0.068***	-0.060	0.035
	(0.0031)	(0.0040)	(0.0100)	(0.0128)	(0.0078)	(0.0123)	(0.0410)	(0.0563)
log(est. size)	0.032***	0.049***	0.017***	0.038***	0.017***	0.048***	0.018***	0.063***
	(0.0005)	(0.0020)	(0.0013)	(0.0040)	(0.0009)	(0.0035)	(0.0047)	(0.0154)
4th polyn. age?	yes	yes	yes	yes	yes	yes	yes	yes
fixed effects?	yr, w.	yr, w.×est.	yr, w.	yr, w.×est.	yr, w.	yr, w.×est.	yr, w.	yr, w.×est.
R^2	0.779	0.865	0.849	0.906	0.820	0.891	0.800	0.852
# obs.	1,522,026	1,522,026	184,033	184,033	426,827	426,827	12,079	12,079
# clust.	260,310	122,131	29,934	26,307	80,157	46,902	2,280	1,738

***: p¡0.01; **: p¡0.05; *: p¡0.10. Regressions of log_{10} ($wage$) with worker fixed effects (uneven columns) and worker-establishment fixed effects (even columns). Standard errors, clustered at the worker level (uneven) or establishment level (even), in parentheses.



(a) Upper secondary

(b) Post-secondary

(c) College

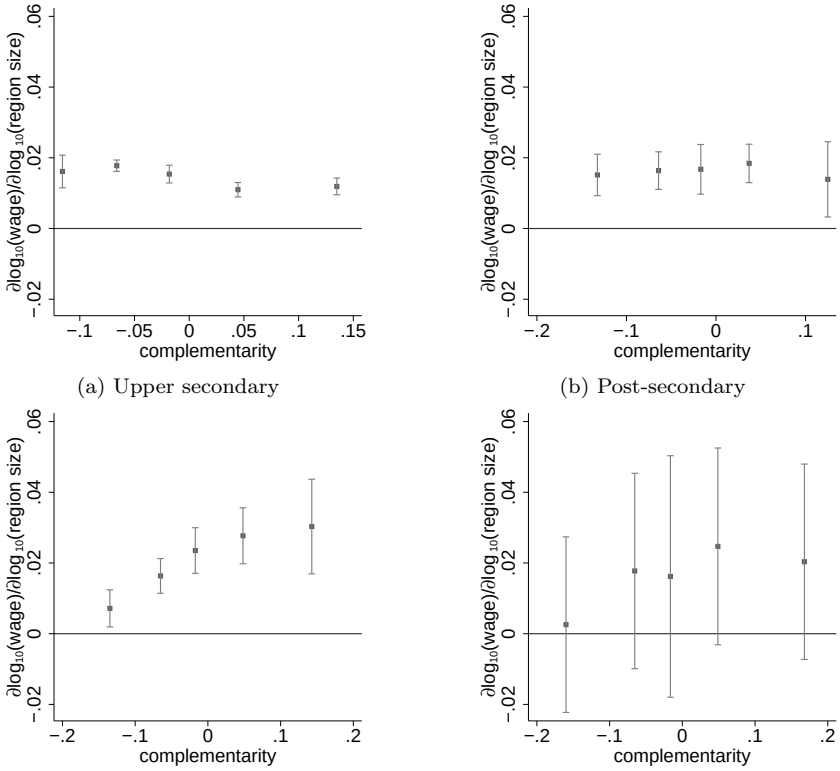
(d) Ph. D.

Figure A.9: Urban wage premium by complementarity quintile and educational level (OLS)
Idem Figure 5a with separate analyses for each educational level.

the subsamples of workers with post-secondary and college educations, although not of workers with upper secondary degrees. Point estimates for workers with Ph. D. degrees exhibit a pattern similar to the one in the overall sample, but due to the small sample-size, these estimates are very imprecise. Adding worker fixed effects in Figure A.10 increases the width of confidence intervals, but for college-educated workers, the rise in UWP across complementarity quintiles remains visible.

A.6 Returns to schooling in Sweden and the U.S.

Here, I compare returns to schooling in Sweden and the U.S.. Table A.4 presents the outcomes. All analyses contain a nth4 order polynomial of a worker's age and year dummies as control variables. Column (1) regresses $log_{10}(wage)$ on a worker's educational level in Sweden. Column (2) repeats this analysis using a worker's percentile rank in the overall wage distribution as a dependent vari-



(a) Upper secondary

(b) Post-secondary

(c) College

(d) Ph. D.

Figure A.10: Urban wage premium by complementarity quintile and educational level (worker FE)

Idem Figure 5b with separate analyses for each educational level.

able. Columns (3) and (4) show analogous regressions, using U.S. census data for the years 2001 to 2010 retrieved from IPUMS-USA (Ruggles et al., 2015) and cleaned following the procedures outlined in David and Dorn (2013). U.S. workers are divided into educational categories that are meant to mimic the Swedish categories as closely as possible. In these columns, the omitted category is workers with at most middle school, "sec." refers to workers who completed 11th grade, "upper sec." to workers who completed grade 12 or have a high school degree, "post-sec." to workers with associate degrees or some college education, "college" to workers with a bachelor's, master's or professional degree and "Ph. D." to workers with a doctoral degree. The sample is constructed using similar criteria as those used for the Swedish data: apart from using the same age restrictions, also self-employed, female and government employed workers, as well as workers in employment agencies are excluded. Moreover, workers below the poverty threshold in real 2010 USD are excluded, as well as workers in the top and bottom 0.5 wage-percentile. Wages are annual wages and the regressions are weighted by workers' sample weights. Column (5) repeats the preferred specification of Table 5, but now uses a worker's percentile rank in the wage distribution as a dependent variable to show the effect of coworker match and substitutability on a worker's wage rank. Note that this regression uses only workers with at least upper secondary education.

Although the absolute wage premium in the U.S. is about three times the wage premium in Sweden, in relative terms, the difference is less pronounced. College-educated workers in Sweden are on average 21.5 percentiles higher up in the wage distribution from workers with only secondary education, whereas in the U.S., college-educated workers are 35.9 percentiles above high-school educated workers (11th grade) in their economy's wage distribution. From this perspective, the returns to coworker match reported in column (5) are substantial. The point estimate implies that moving from the 10th to the 90th percentile of coworker match is associated with a 10.3 percentiles rise in the wage distribution, whereas the drop in wage-rank associated with a similar increase in substitutability is 2.9 percentiles.

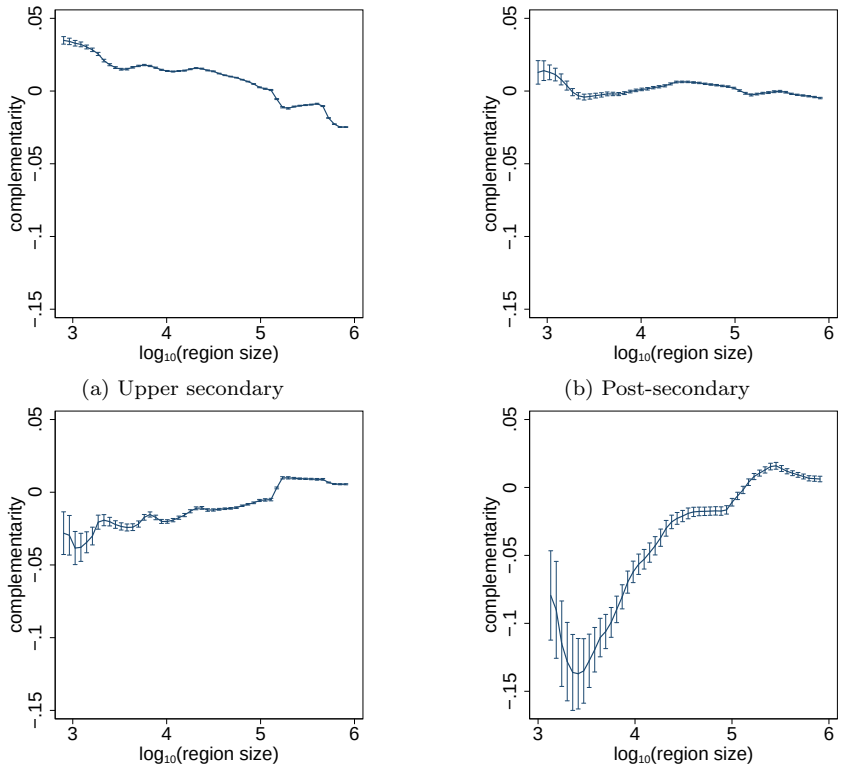
A.7 Coworker complementarities and the size of a region

To explore whether larger regions help workers find more complementary coworkers, Figure A.11 shows the relation between coworker complementarity, $\hat{m}_{wpt}$, and the logarithm of a labor market area's working population and Figure A.12 shows how coworker complementarities change when workers move from one region to another. Both figures are based on the average employment size of a region throughout the sample such that estimates are not affected by differential regional growth rates. Both graphs show that, at least for workers with high levels of education, coworker complementarities tend to rise with a region's size.

Table A.4: Returns to education

	Sweden		U.S.		Sweden
	(1)	(2)	(3)	(4)	(5)
dep. var.:	$\log_{10}$(wage)	wage pct.	$\log_{10}$(wage)	wage pct.	wage pct.
sec.	0.037***	0.049***	0.060***	0.063***	
	(0.0013)	(0.0020)	(0.0012)	(0.0013)	
upper sec.	0.077***	0.106***	0.142***	0.151***	-0.138***
	(0.0012)	(0.0019)	(0.0010)	(0.0010)	(0.0013)
post-sec.	0.175***	0.247***	0.209***	0.228***	
	(0.0015)	(0.0022)	(0.0010)	(0.0010)	
college	0.199***	0.264***	0.415***	0.422***	0.001
	(0.0013)	(0.0020)	(0.0010)	(0.0010)	(0.0014)
Ph. D.	0.301***	0.370***	0.506***	0.495***	0.063***
	(0.0034)	(0.0034)	(0.0022)	(0.0019)	(0.0030)
$\log_{10}$(est. size)					0.068***
					(0.0004)
cow. match					0.390***
					(0.0043)
cow. subst.					-0.060***
					(0.0026)
4th polyn. age?	yes	yes	yes	yes	yes
fixed effects?	yr	yr	yr	yr	yr
R^2	0.255	0.227	0.304	0.309	0.287
# obs.	2,576,964	2,576,964	3,237,003	3,237,003	2,144,965
# clust.	440,578	440,578			364,642

***: p¡0.01; **: p¡0.05; *: p¡0.10. Standard errors (in parentheses) are clustered at the individual level when using Swedish data and robust for the U.S. data sample.



(a) Upper secondary

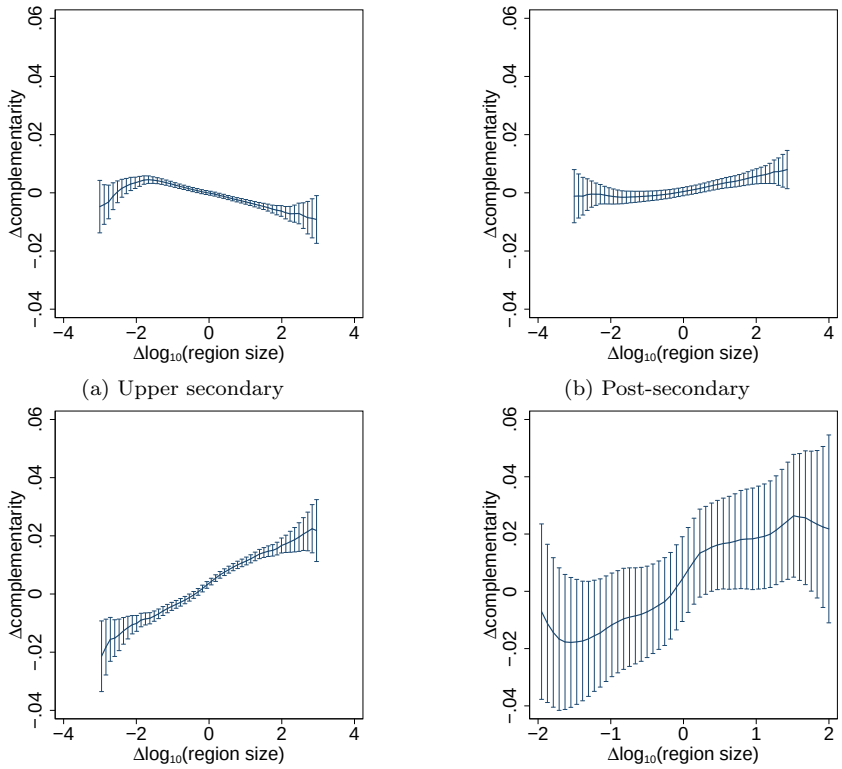
(b) Post-secondary

(c) College

(d) Ph. D.

Figure A.11: Coworker complementarity and region size

Local polynomial smooths of the relation between complementarity, $\hat{m}_{wpt}$, as defined in (25) and the base-10 logarithm of the average total employment in a labor market area in the period 2001 to 2010.



(a) Upper secondary (b) Post-secondary

(c) College (d) Ph. D.

Figure A.12: Change in coworker complementarity and region size (region switchers)

Local polynomial smooths of the relation between the change in complementarity, $\hat{m}_{wpt}$, as defined in (25), and the change in base-10 logarithm of the average employment size of a labor market area in the period 2001 to 2010 for workers who change jobs between labor market areas.