When workers' skills become unbundled: Some empirical consequences for sorting and wages[a]

Oskar Nordström Skans[b], Philippe Choné[c] Francis Kramarz[d]

March 2022

Abstract

This empirical paper analyzes labor market sorting across establishments using Swedish register data on cognitive and non-cognitive abilities. We draw on the theoretical foundations of Choné and Kramarz (2021), in which workers are endowed with sets of multidimensional skills that need to be sold in "bundles" to employers that differ in their use of each of these skills. The theory also outlines how wage and sorting patterns should evolve when innovations "unbundle" the skills through the emergence of markets where each specific skill can be traded separately. Our empirical results show that labor is sorted across establishments on both comparative advantage and absolute ability. Furthermore, wage returns to each skill is higher in market segments where employers rely more heavily on workers who specialize in that particular skill. Changes over time are well in line with a process of unbundling; sorting on comparative advantage has increased and the market wages of generalists have risen relative to those of specialists.

Keywords: Cognitive skills, Non-Cognitive Skills, Firms, Technology

JEL Codes: J23, J24, J31

[a]This empirical paper was prepared as a companion paper to the theoretical paper by Choné and Kramarz (2021). We thank Maks Khomenko, Jonas Vlachos and seminar participants at Uppsala University and The Economics Of Creative Destruction (in Honor of Aghion and Howitt) for useful comments. Kramarz and Skans acknowledge funding from the ERC advanced grant FIRMNET (project ID 741467).

[b]Uppsala University, UCLS, IFAU and IZA

[c]CREST-ENSAE, IPParis

[d]CREST-ENSAE, IPParis, and Uppsala University

Contents

1 Introduction

This paper studies how workers are sorted across establishments based on workers' attributes and firms' characteristics. This classic question in labor economics has a long tradition (e.g. assignment models à la Sattinger (1975)). Recently, the literature has started to focus on the multidimensional nature of workers' skills, i.e. when workers come equipped with *bundles* of skills with no access to markets where skills can be traded separately. On the theory side, recent work of Choné and Kramarz (2021) and Edmond and Mongey (2020), has expanded the path-breaking work of Heckman and Scheinkman (1987) and Lindenlaub (2017). On the empirical side, several papers have approached related questions on sorting and skills. In particular, Fredriksson et al. (2018) explored how workers are sorted across heterogeneous jobs whereas Guvenen et al. (2020) study sorting across occupations.

This paper attempts to bring the empirical and theoretical strands closer together by providing reduced-form evidence using the model adopted by Choné and Kramarz (2021). In Choné and Kramarz (2021), following Heckman and Scheinkman (1987), *each worker is equipped with a set of skills that needs to be sold as a bundle to a single employer*. Firms have heterogeneous production technologies and, in equilibrium, use different types of workers. The model clearly outlines how the supply and demand for different types of skills determine the structure of sorting and of market wages across worker types. In addition, the model describes how sorting patterns and the wage structure change when innovations to technology and/or markets *unbundle* the skills so that they can be sold separately.

In this text, we examine the empirical counterpart of the above theoretical structure. We contrast key predictions of the Choné and Kramarz (2021) model (CK, hereafter) with Swedish register data on cognitive and non-cognitive skills, when bundled and when unbundled. Skills are measured for nearly all Swedish males who entered into adulthood during three decades starting in the early 1970s. We use these data to track the distribution of workers' sorting and the relationship to wages during the period 1996-2013.

Some of our empirical analyses have closely related predecessors, many of which used the same data sources. Fredriksson et al. (2018) used the same data to study how skill sorting at the job-level evolve with tenure among new matches, in particular, when inexperienced workers search for a suitable job, but the theoretical motivations are entirely different. Our analysis of the evolution of labor market sorting to that contained in Card et al. (2013), Song et al. (2019), or Skans et al. (2009) for Sweden. Even more related is, however, Håkanson et al. (2021) who use the same data as this paper to study how ability sorting has evolved over time. Again, our objective differs, hence some of their results are connected to ours but with a different interpretation. Hensvik and Skans (2020) describe the association between skill content trends in labor demand at the occupational-level (rather than at the job-level). Recent work by Böhm et al. (2020) complements our analysis by relating sorting to wages using the same data source. However, their analysis identifies the effects of interest from workers' movements in the tradition of Abowd et al. (1999), when we show that this has no theoretical foundation in our approach. Their analysis is, in this sense, similar in spirit to Fredriksson et al. (2018) as both identify their effects from non-competitive frictions across firms or jobs. More generally, our wage analysis focuses on the *market returns*, in the spirit of the Chicago school, by studying the impact of bundling constraints on wages, abstracting from search frictions and other market imperfections.

Our results clearly demonstrate that workers are non-randomly sorted across establishments. Indeed, establishments specialize both in the horizontal (mix of skill-types), as CK predicts, and vertical (quality of each skill-type) dimensions, as assortative matching predicts. However, the horizontal dimension seems to dominate, in particular at the top of the ability distribution. For instance, high-skilled workers with more cognitive than non-cognitive skills are much more likely to work with workers sharing these exact traits. In addition, as CK predicts, such workers are more likely to work with "middle skilled" workers whose skills are also "specialized" in this cognitive dimension. *But*, again in line with CK, they are *less* likely to work with high-skilled workers who hold *non-cognitive* skills. Importantly, this tendency of specialists to work with others specialized in the same skill has increased progressively across cohorts and over time, regardless of the age at which we evaluate the patterns.

Furthermore, market wages have properties that are well in line with CK's predictions. In particular, wage returns to each specific skill is higher in the segments that are dominated by firms that rely more heavily on workers of that type. Finally, in parallel with the increase in sorting through time, wages of generalists have grown more rapidly that those of specialists as predicted by CK's analysis of skills' unbundling (due to innovations such as outsourcing or platform markets à la Uber).

The paper is structured as follows. Section 2 presents the main elements of the theory contained in Choné and Kramarz (2021). Section 3 presents the data. Section 4 shows results on sorting. Section 5 present results on wages. Section 6 concludes.

2 The CK model of skill-bundling and unbundling

We start by presenting some useful elements of Choné and Kramarz (2021)'s theory. We outline the nature of the theoretical problem and the essence of its mathematical solution. This solution will constitute the basis of the empirical elements presented in the following Sections.

2.1 The setting: An economy with skill bundling

CK models the matching between heterogeneous workers, endowed with multidimensional skills, and firms, heterogeneous in their production functions.

Formally, a worker's skill endowment is a vector $x = (x_1, \ldots, x_j, \ldots, x_k)$, where each element x_j represents worker's endowment-level of skill type j. We may refer to $\lambda = |x|$ as the overall quality of a worker of type x. Similarly, we refer to $\tilde{x} = x/|x|$ as her skill profile. The skill profile represents a *horizontal* dimension of heterogeneity, or the comparative advantage of the worker. It is natural to think of some workers as generalists when they have a balanced skills-set, whereas others are specialized, when their endowment is large in some dimension but small in another. Similarly, heterogeneity in $\lambda = |x|$ represent *vertical* heterogeneity, i.e. some workers have larger skill absolute endowments but an identical skill profile. Throughout, we assume that the supply of skills is exogenously fixed, before the matching takes place.[d]

The multidimensional nature of workers' skills matters because firms are heterogeneous in their needs for these skills. CK models an economy where each firm's production process involves $k \geq 2$ tasks. Task j, $j = 1, \ldots, k$, is produced through a linear aggregation of the employees' endowments in skill j:

$$X_j = \int x_j \, \mathrm{d}N^d(x; \phi), \tag{1}$$

where $\mathrm{d}N^d(x; \phi)$ is the number of workers of type x hired by the firm with type ϕ.

All firms' production functions $F(X; \phi)$ are concave in the firm-level aggregate skill vector X. As mentioned above, firms differ both in their vertical and horizontal dimensions. In the vertical dimensions, firms are endowed with a total factor productivity (denoted by z). In the horizontal dimension, firms differ in their need to use different tasks (and thus,

[d]Skills are distributed according to a positive probability measure $\mathrm{d}H^w(x)$ on $\mathcal{X} = \mathbb{R}^k_+$.

skills) in the production process. Horizontal differences are captured by the parameter α. Firm-level heterogeneity thus takes the form $\phi = (\alpha, z)$, where $F(X; \alpha, z) = zF(X; \alpha, 1)$. It is natural to assume that firm and worker heterogeneity has the same dimension, i.e. $k - 1$.[e]

The output in equation (1) is an aggregation of workers' skills for each skill-type j used to produce an intermediary input, task j, that enters the firm's production function $F(X; \alpha, z)$. Hence, we use the terms tasks (both an input of the firm's production function and an output of skill aggregation) and skills (an input to produce tasks) interchangeably in what follows.

A matching between workers and firms is characterized by a coupling $\pi(x, \phi)$ of the measures H^w and H^f, i.e. a measure on $\mathcal{X} \times \Phi$ that admits H^w and H^f as marginals on $\mathcal{X}$ and Φ respectively. The surplus to be shared between firms and workers is the total output in the economy

$$\text{Total Output} = \int F\left(\int x \, \mathrm{d}\pi(x|\phi); \phi\right) \mathrm{d}H^f(\phi), \tag{2}$$

which differs from $\iint F(x; \phi) \, \mathrm{d}\pi(x; \phi)$, the grand sum of firm-specific tasks, because F is nonlinear in X. To fix ideas, CK often makes use of the CES production function with constant elasticity of substitution and decreasing returns to scale:

$$F(X; z, \alpha) = (z/\eta) \left[\sum_{j=1}^{k} \alpha_j X_j^{\sigma}\right]^{\eta/\sigma}, \tag{3}$$

with $\sum_{j=1}^{k} \alpha_j = 1$, $\eta < 1$, $\sigma \neq 0$, and $\sigma < 1$. The parameter α_j reflects the intensity of the firm's demand for skill-type j.

Competitive bundling equilibria: Under bundling, *the workers' sets of skills cannot be untied* since there are no separate markets for each skill. Firms and workers are restricted to trade in packages of skills $x = (x_1, \ldots, x_k)$. The worker skills are observed by the firm and are contractible. We rule out agency problems: a firm that pays $w(x)$ for x gets exactly x. Apart from the bundling friction, we abstract from all labor market frictions. CK shows that there is a market wage for a worker of type x, denoted by $w(x)$.

Given the wage schedule $w(x)$, the skill demand of a firm of type ϕ is a positive measure

[e]Firms' types are distributed according to a probability measure $\mathrm{d}H^f(\phi)$ on a set Φ. We normalize the numbers of firms and workers to one w.l.o.g.

$\mathrm{d}N^d(x;\phi)$ that maximizes its profit

$$\Pi(\phi; w) = \max_{\mathrm{d}N^d} F\left(\int x\,\mathrm{d}N^d(x;\phi)\right) - \int w(x)\,\mathrm{d}N^d(x). \tag{4}$$

The objective function of the firm depends only on its aggregate skill $X^d(\phi) = \int x\,\mathrm{d}N^d(x)$ and on the associated wage bill $\int w(x)\,\mathrm{d}N^d(x)$. The wage schedule in equilibrium must be such that the sum of all firms' skill demands $N^d(x;\phi)$ and supply of skills coincide for all worker-types. The existence of equilibria as well as the equilibrium properties are presented in CK. In particular, in equilibrium, the wage schedule is shown to be convex and homogenous of degree one.

At the firm ϕ's aggregate skill X^d demand, the productivity of each skill equals its marginal price:

$$F_j(X^d(\phi);\phi) = w_j(X^d(\phi)). \tag{5}$$

This first-order condition (5) generalizes the standard condition that wage equals marginal productivity at a competitive equilibrium. When the wage schedule is locally linear, i.e., price equals marginal productivity. Otherwise, the *implicit* price of skill i in the neighborhood of the aggregate skill X^d is the partial derivative $w_i = \partial w/\partial x_i$ evaluated at that point.

The case with two skills and tasks. When $k = 2$, we may parameterize skill profiles as $\tilde{X} = (\cos\theta, \sin\theta)$ and represent the aggregate demand $X^d = (\Lambda^d \cos\theta^d, \Lambda^d \sin\theta^d)$ in polar coordinates, where Λ^d is the total quality of workers employed at firm ϕ.

The strict convexity of the wage schedule implies that it is cheaper for firms to purchase the bundle $(x_1(\theta), x_2(\theta))$ from a generalist worker (a worker endowed with both skills in sufficient quantities) than to purchase $x_1(\theta)$ units of skill 1 and $x_2(\theta)$ units of skill 2 separately from specialist workers.

Furthermore, when the production function is CES with two skills:

$$F(X_1, X_2; \alpha_1, \alpha_2, z) = \frac{z}{\eta}\left(\alpha_1 X_1^{\sigma} + \alpha_2 X_2^{\sigma}\right)^{\eta/\sigma},$$

CK show that a sorting condition holds in equilibrium.

The sorting between workers and firms is represented by the increasing function ensuring that firms with a technology that is efficient in the use of skill j also employ more

j-specialists. Furthermore, CK shows that the total level of skills (quality) of the workers employed by firm, $\Lambda^d(\alpha_2, z)$, increases with firm's total factor productivity z, thus in this sense the model exhibits positive assortative matching (*PAM*). Overall, and in contrast to alternative models such as Lindenlaub (2017), the sorting pattern highlighted here pertains to both the horizontal and the vertical dimensions of workers' skills (skill profile and total quality) rather than to each of the two skills separately.

Essential for labor economists, an equilibrium wage exists. Furthermore, for any strictly convex wage schedule $w(x)$, for any homogenous production function $F(.;\phi)$ satisfying a classic single-crossing assumption, and for any workers' distribution H^w, there exist distributions of the firms' technological parameters ϕ for which w is the equilibrium wage.

When the equilibrium wage schedule includes linear parts (facets): Until now, we focused on the case when the equilibrium wage schedule we studied was *strictly* convex. But there is also a possibility that the equilibrium wage schedule includes linear parts. This can happen when the market demand for (local) generalists is sufficiently high that it starts to be profitable for firms to instead hire and combine, or "bunch", (local) specialists of different kinds instead of only hiring the generalists.

The equilibrium wage schedule typically has strictly convex parts together with linear parts. In the case where $k = 2$ it is useful to envision the space of skills as represented in a positive quadrant. We can consider workers at the extreme left and extreme right of the angle that defines a linear segment of the convex wage schedule. Consider workers in the middle of the angle; these are local "generalists" and the workers on each side are local specialists. Essentially adding the price for the skills of a specialist worker at the extreme left to that for the skills of a specialist worker at the extreme right will yield the price for the skills of the (appropriately selected) generalist workers in the middle if the wage schedule is linear. Put differently, the sum of the wages for two (local) specialists is equal to the sum of the wages for two (local) generalists.

Hence, in the case of bunching, the firm will obtain its optimal mix by hiring workers with different skill profiles rather than focusing on a unique skill profile as in the case of strictly convex parts of the equilibrium wage schedule. But even when this happens, there remains a perfect separation in the sense that each firm's aggregate skill mix θ always increases with α. Thus, we still have full sorting in terms of the *skill mix* of the workers in relation to the technology of the firm.

2.2 Unbundling

CK discusses in detail what happens when new technologies (Uber being a prominent example) or changes in market institutions (the Hartz reforms that facilitated the use of temp-agencies in Germany in the 2000s) enable unbundling of skills. Then, workers and firms become able/allowed to trade skills as separate commodities. In a first step, they discuss the case when this unbundling technology is costless for all market participants. In a second stage they assume that it entails some costs incurred by workers and/or firms.

Costless Unbundling: In this case, full efficiency prevails when competitive markets for individual skills do exist. This unconstrained efficiency therefore requires that the marginal productivities are constant across firms, i.e., for any $j = 1, \ldots, k$, there exists μ_j such that

$$F_j(X^*(\phi); \phi) = \mu_j$$

for all firms ϕ. Because there are k markets, one for each skill, there are k prices. On the supply side, the total supply of skills is unchanged. However, each worker can split her entire supply of skills between an employing firm and the market, making individual labor supply *endogenous* (in contrast to the bundling case where workers were "forced" by the technology to sell all their skills to a unique firm, which used them in full).

Assuming two tasks and a CES technology, CK characterizes those workers benefiting from full unbundling and those harmed in the process. More precisely, CK show that, when the production function is given by (3) with $k = 2$, and except in the case where the wage schedule is linear under bundling (i.e., there is full bunching), at least some generalist workers are strictly better off after unbundling. Furthermore, if skills are complements ($\sigma < \eta$), at least one type of specialist workers ($\theta = 0$ and/or $\theta = \pi/2$) is strictly worse off. The extent to which generalists benefit from unbundling and specialists are harmed by the process is an empirical question, partly addressed in the following empirical analysis.
CK characterizes this unbundling process further and show that after unbundling, specialized firms tend to specialize further, with their skill mixes being better aligned with their technologies. So specialization is an outcome, a result of the opening of markets rather than an assumption embedded in unbundling.

Costly Unbundling: So far, we have assumed that the unbundling of skills is a costless process. However, if unbundling comes from an innovation (such as Uber which creates a market for driving skills), workers are likely to have to pay a fee or, more generally, incur a cost to have their skills unbundled. This creates wedges between the market wages paid to workers and prices paid by firms. Two interpretations for these wedges are possible:

1. There is one market price p_i^f for skill i, but workers incur a cost c_i per unit of unbundled skill i;

2. The platform(s) purchase(s) skill i from workers at price p_i^w and resell(s) it to firms at price p_i^f, with a margin c_i.

CFurthermore, the range of implicit prices for each skill satisfies:

$$\max w_i^u - \min w_i^u \leq c_i, \tag{6}$$

where c_i is the cost incurred per unit of unbundled skill i. If a positive amount of skill i is traded on the market, then equality prevails in (6), with $p_i^f = \max w_i^u$ and $p_i^w = \min w_i^u$ being respectively the firm price and the worker price for that skill.
The presence of wedges between firm and worker prices implies that contracted workers – those who supply one of their skill through the market – and employed workers – those who supply their skills bundle to a firm – are paid different prices for the same skill used at the same firm.

3 Data and empirical strategies

3.1 Data overview

We use a broad data set covering Swedish male workers' multidimensional skills. The data originates from the Swedish military conscription tests taken by most males born between 1952 and 1981.[f] The tests were taken at age 18 and the data should therefore be understood as capturing pre-market abilities. There are two main components; *cognitive abilities*, henceforth denoted as C, which is measured through a written tests and *non-cognitive abilities*, henceforth denoted as N, which is measured during a structured interview with a specialized psychologist. As noted in the introduction, the data have been used to assess labor

[f]See Mood et al. (2012) for details on the data collection. Although the share of test takers is somewhat lower in the final years, we have no reason to believe that this will interfere with our analysis.

market sorting in previous work, most notably by Fredriksson et al. (2018) and Håkanson et al. (2021). Our definitions and set-up draws heavily on Fredriksson et al. (2018) in several dimensions.

Our used data on employment cover 1996 to 2013 and we include all workers (aged 20 to 64) with reported test results. An important component of the analysis is that the cross-worker heterogeneity in skill-types that is being measured at age 18 remains relevant for understanding worker heterogeneity later in life. Previous work (and our own results presented below) has shown that this is a plausible assumption, skill-types are related both to wages and to the type of work people perform throughout their careers, see e.g. Fredriksson et al. (2018), Håkanson et al. (2021), and Lindqvist and Vestman (2011).

We include all workers in their main job in November as long as we can identify their establishment.[g] When studying the link between skills and wages, we use wage data from the Structure of Earnings Statistics. These data come from a firm-level survey which heavily over-samples large firms. The data cover 30 percent of private sector employees and all public sector workers. We can verify that our main wage results are insensitive to the sampling by using average monthly earnings, which we observe for all. For the same set of workers, we observe occupations. For all our our analyses, we only include one job per worker and year.[h]

Our main target for the sorting analysis concerns how workers are sorted across *Establishments*. We include all establishments with between 6 and 600 workers with measured skills. But we also present results for *Jobs* defined as the intersection of the occupation (at the 3-digit level) and establishment of the worker as in Fredriksson et al. (2018). All results are stable across these two definitions.

3.1.1 Defining generalists and specialists

The skill data are measured on an ordinal discrete (integer) scale ranging from 1 to 9. Standard practice in the literature is to treat these data as if continuous and cardinal after standardizing them to mean zero and standard deviation one within each birth cohort, see

[g]An establishment is a physical place of work within one firm. About 10 percent of all workers do not have a fixed physical place of work and these are therefore not included.

[h]The preferences order to is to first use observations where the wage can be observed. Wages are sampled in October or November. If there is no (unique) such observation, we select the observation with the highest earnings.

e.g. Lindqvist and Vestman (2011). We proceed differently and, whenever we can, instead strive to build our empirical strategies to account for the fact that the data is reported on a discrete ordinal scale. We assume that the ordinal scales have monotonic relationships to the underlying productive abilities they represent.

We use as our main empirical tool a classification of workers as *Generalists* or *Specialists* depending on the relationship between the two reported scores. This corresponds to the concept of x_1/x_2 in the theory section. As we are unable to precisely compare the two scales, we allow the data to "wiggle" one step before referring to workers as specialists and therefore count workers with less than a one-step difference between the scores as generalists. We thus heuristically define workers as *Generalists* if $abs(C_i - N_i) < 2$ and consequently define workers as *C-Specialists* if $C_i > N_i + 1$ and *N-Specialists* if $N_i > C_i + 1$.[i]

These definitions force us to assume that there is some shared relationship between the two scales (i.e the measures C_i vs. N_i) for each given worker i. On the other hand, the computation does not rely on any cardinal interpretation of differences along each of the scales.

Building on this worker-level classification, we classify the skill-type environment each worker has in his establishment. The classification relies on (between 5 and 599) coworkers with measured skills and we need to assume that these coworkers' skills reflect the overall skill environment of the establishment. We define establishment types as follows: *Generalist establishments* have a share of Generalists that exceeds 50 percent.[j] Other establishments are classified as either *C-establishments* or *N-establishments* depending on the which type of specialists that dominate amongst the employees. This classification does, according to the theory, inform us about α, i.e. the type of production function used by the establishment. To ensure that we do not generate any mechanical relationship between the worker's own skills and the measure of establishment types, we only use the *coworkers* when classifying establishments. [k]

[i]In the appendix we show that this classification scheme results in fairly stable shares of worker types across time. This is reassuring as there are some minor changes in the test protocol which generates a few minor discrete changes in the share of test takers at each value as also shown in the appendix.

[j]Establishments with an exactly equal share of N-Specialists and C-Specialists are also considered "Generalist". This is obviously only relevant to the smallest establishments.

[k]This means that the same establishment, in principle, can be classified differently for different workers within the same establishment (because the excluded worker is different). This empirical curiosity does not have any impact on our conclusions.

For some of our analyses it is also useful to classify workers in terms of overall ability levels. Here we define workers as *low skilled* if the "sum" of (measured) cognitive and non-cognitive ability falls below 9 and *high-skilled* if the same sum is above 11 whereas the *mid skilled* are those where the sum is in-between. This classification is obviously more cardinal in nature as the base is an accumulation of high and low values on to the inherently ordinal scale. This caveat should obviously be kept in mind when interpreting the results but a mitigating factor may be that we only use this classification in contexts where we simultaneously account for the workers' specializations in the C/N dimension.

3.2 Descriptive statistics

Figure 1 below depicts the joint distributions of the skills as reported on their 1-9 scale. The lower panels show the joint distributions, and as is evident the skills are correlated (correlation in 0.37 in the used data) but also contain independent information.

Table 1 shows descriptive statistics for the used sample. The first column shows the full used data. As is evident, average test scores are marginally above 5 in both dimensions. Around half of the sample are classified as generalists (i.e. being on the diagonal of the joint distribution depicted in Figure 1) and about one quarter each are specialists in the cognitive or the non-cognitive dimension.

The following columns split the data in these three groups (generalists, C- vs. N-Specialists). The table shows, as expected, that the groups are equally distributed across years, ages and birth cohorts. Cognitive skills are "twice" as high (6.9 vs. 3.6) among cognitive specialists as among non-cognitive specialists, but as discussed above, these scales do not have a natural interpretation in terms of the productive content of these scores. The differences in terms of non-cognitive skills are also intuitive (6.3 vs. 4.1 for the two types of specialists). There is a tendency for C-specialists to over-represented in the group of "high skilled", but as is evident all ability levels are well represented among generalists and among both types of specialists.

Since most workers are classified as generalists, most establishment are also dominated by generalists. And this also makes it more common for the generalists to be working in an establishment dominated by the own group (in that sense, "matched").

The final column present statistics for the half of the overall sample where we can observe wages. As is shown, this sample is nearly identical to the sample where we can observe occupations. The most important aspect of this column is that the data are very similar to the first column (All) in all aspects (such as skill levels and composition), except for establishment size. The latter arises mechanically from an oversampling of large firms. Fortunately, we are able to verify the stability of our wage-results by estimating the same models using monthly earnings data (that we observe for all) instead.

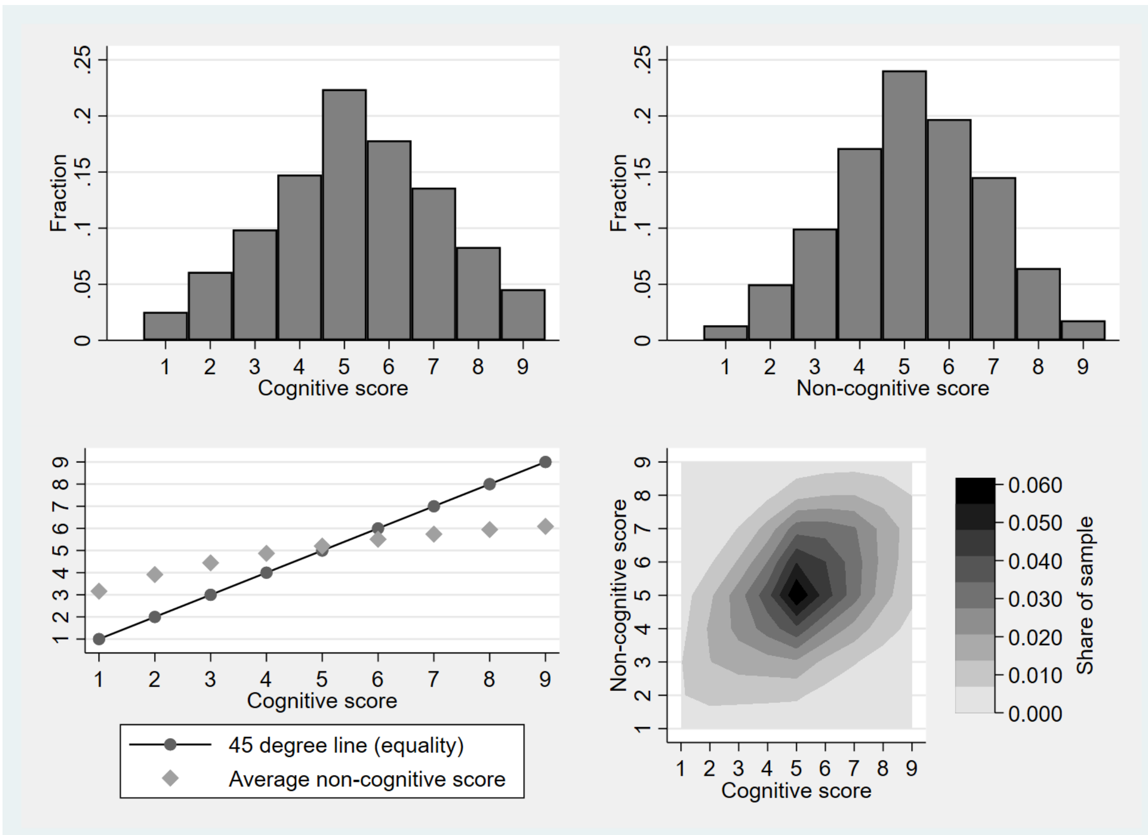

Figure 1: Measured ability scores

Note: The figure shows the test score results in our used data. See restrictions in the text. The bottom panels illustrate the joint distributions.

4 Sorting

We are interested in analysing how workers skills are related to the skill-requirements. In the spirit of Fredriksson et al. (2018), we will classify the establishments based on coworker skill set as explained above. We then regress the worker skill type on the type of workers they are working together with. As a starting point, we only use one year (2005) and defer the analysis for trends over time to later. Thus, we estimate models of the following form:

$$Y_{ij}^{\tau} = a + b^{C,\tau} * C_j^{-i} + b^{N,\tau} * N_j^{-i} + \epsilon_{ij} \tag{7}$$

where Y_{ij}^{τ} represent the type of worker i, employed at workplace j. Types will be indicators for being a specialist of type $\tau = C, N$, or a generalist. C_{jt}^{-i} and N_{jt}^{-i} measures the share of coworkers that C-specialists and N-specialists (the residual type is generalists). If workers are sorted into contexts where other workers are of a similar type (arguably, because this is what the firm-level technology asks for), we expect positive values on $b^{C,C}$, but negative values on $b^{C,N}$ and so forth.

4.1 Simulating assignment principles

We contrast the real sorting patterns with corresponding estimates that we derive from a simulated allocation of observations across the actual establishment size distribution. In practice, we first sort the establishment at random, preserving their size. We then sort the workers and assign them to establishments. We start by randomly sorting workers at random before matching them to the establishments. The second simulated assignment ranks workers on *absolute ability* as proxied by C+N before allocating them to establishments, thus placing all the highest skilled workers in the (randomly sized) first establishment and so forth. This assignment captures the idea that better workers are assigned to more productive firms, and it is therefore closely related to the concept of positive assortative matching ("PAM"). Third, we rank workers according to *relative abilities* as proxied by by C/N, thus placing all workers with the strongest C-specialisation in the (randomly sized) first establishment and so forth.[1] This generates four different allocations (Actual, Random, Absolute ranking and Relative ranking) all of which have the identical number of workers per ability type, and an identical establishment-size distribution.

[1] The collected data is discrete, but it is natural to think about the actual abilities as being continuous. We therefore first generate simulated raw continuous skills data that exactly aggregates up to our actual data in terms of number of workers with each combination of skills *and* which ensures that the correlations across skill types also is replicated within these types. We then allocate the workers according to these continuous scores, this process is inconsequential for the analyses presented here.

Table 1: Descriptive statistics of used data

	(1) All	(2) Generalist	(3) C-Specialist	(4) N-Specialist	(5) Wage obs
Year	2004.8	2004.8	2004.9	2004.7	2005.1
Cohort	1965.8	1966.0	1965.4	1965.8	1965.1
Age	39.0	38.8	39.5	39.0	40.0
Worker skills:					
Cognitive (C=1-9)	5.252	5.190	6.914	3.643	5.366
Non-cognitive (N=1-9)	5.179	5.206	4.090	6.267	5.239
C+N low (< 9)	0.252	0.237	0.207	0.339	0.233
C+N mid ($9 - 11$)	0.376	0.422	0.316	0.325	0.371
C+N high (> 11)	0.371	0.341	0.476	0.336	0.396
Establishment size	82.1	81.9	88.2	76.0	118.4
Generalist establishment	0.767	0.777	0.722	0.787	0.782
Cognitive establishment	0.136	0.125	0.209	0.087	0.141
Non-cognitive est.	0.097	0.098	0.069	0.126	0.077
Matched	0.504	0.777	0.209	0.126	0.507
Observed occupation	0.517	0.514	0.539	0.503	0.978
Observed wage	0.529	0.526	0.551	0.513	1.000
ln(Wage)	10.182	10.182	10.227	10.131	10.182
ln(Earnings)	10.102	10.104	10.138	10.059	10.157
N	12,627,401	6,964,632	2,744,810	2,917,959	6,682,011

Note: Descriptive statistics for the used data covering 1996-2013. Establishments are restricted to be size 6 (i.e. 5 coworkers) to 600. In columns (2) to (4) we split the sample and according to if the worker is a Generalist, defined as $abs(C - N) < 2$ or a Specialist in C or N. Column (5) only uses workers for whom we have information on wages. Generalist establishments have a majority of employees as generalists, or an exactly equal share of specialists of the two types. Non-generalist establishments are classified according to the dominating type of specialists among employees. These classifications only use *coworkers*, i.e. not the subject himself. "Matched" workers are C-Specialists in Cognitive establishments and so forth. Monthly earnings are recorded for all observations.

The results presented in Table 2 show that workers indeed are systematically sorted across establishments, although not as strongly or one-dimensional as suggested by the extreme absolute and relative sorting scenarios. Each type of worker is more prevalent if there are more coworkers of the same type. The table also illustrates the empirical approach "works" in the sense that the random allocation indeed do generate an independence between the own type and coworker types (i.e. all estimates are insignificant if workers are allocated at random). Strikingly, the actual allocation is such that there are less C-Specialists in establishments with many N-Specialists and conversely (remember, generalists are the omitted category). In terms of signs (although not magnitudes) this is exactly as implied by the relative sorting scenario emphasized by CK.

In section 4.3 below, we will analyse how these patterns change over time, and there we also present robustness tests that asserts that these general sorting patterns are robust to a number of variations of the empirical model and data.

4.2 Two-dimensional types

Next, we define a more detailed set of worker and establishment types by also accounting for ability levels. As discussed above, we define workers as low skilled if the sum of cognitive and non-cognitive ability falls below 9 and high-skilled if the sum is above 11 whereas the mid skilled are those in-between. By combining these categories with the indicators for generalists and C vs. N-specialists we get 9 types of workers. We then run regressions based on equation (7) where we let each of these 9 types be the outcomes in separate regressions. The explanatory variables are the coworker (leave-out) mean levels of these attributes. We start by estimating the impact of horizontal (specialists) and vertical (high/low) attributes separately, and then present estimates from fully interacted models.

Tables 3 shows the first set of estimates. The table highlights in **bold** the estimates that should be interpreted as indicating similarity between the subject and his coworkers. As is evident from column (1) panel A, *high-level N-Specialists* are found among other high ability workers and other N-specialists, all other estimates are negative. The pattern repeats itself for *high-level generalists* in Column (2) of the same panel and for *high-level C-specialists* in Column (3).

The following panels reiterate the same patterns for mid- and low-level workers and the patterns are qualitatively very similar, workers at all level are more likely to work with workers with a similar specialization, and workers with a similar ability level. Overall, the horizontal sorting does, however, appear to be stronger higher up in the ability ladder. The one estimate that deviates the overall patter is that there appears to be a positive association between N-specialists and mid-level generalists.

Table 4 goes one step deeper by characterizing coworkers in the same 9-dimensional way as the dependent variables. To reduce clutter, we only show regression estimates for the 6 regressions where the outcomes are for specialists and ignore the generalists at this stage. Unsurprisingly, given the estimates presented above, we find that workers are sorted into establishments where other workers are of the *exact* same type, in particular if they are of high ability.

In addition, column (1) also shows that there are fewer (-0.078) High-level N-specialists in establishments where there are many High-level C-specialists, but more where there are many mid-level N-specialists (0.055). Thus, in this segment, workers appear to sort more on the horizontal specialization than on the ability level. The same result holds for High-level C-specialists in column (2) where the impact of High-level N-specialists is negative (-0.040) but the impact of mid-level C-specialists is positive and very strong (0.164). As in the CK-theory, this implies that firms tend to hire workers who have a similar specialization, but who differ in ability levels.

Turning to the mid-level specialists in columns (3) and (4), we see similar patterns with positive estimates for coworkers at different levels but same specialization (0.015, 0.012, 0.063 and 0.061) but negative estimates for all types of coworkers with the same level but a different specialisation. In columns (5) and (6) we study the low level workers and here the estimates are somewhat less conclusive; we still find positive impacts of mid-level workers with the same specialisation, but we also find positive estimates for low-level workers with another specialisation. Thus, the results appear to suggest that specialization on relative skills is more prevalent at the higher end of the ability scale.

Overall, the key take-away is that the results confirm the picture that workers always are sorted into establishments where other workers are of similar types. This is consistent with the notion that employers have heterogeneous production functions that differ in how much productive use they can make out of N and C skills respectively.

Table 2: Leave-out mean regressions on worker types

	(1) Actual sorting	(2) Random sorting	(3) Sorting on C+N	(4) Sorting on C/N
Panel A:				
Dependent variable: Being N-specialist				
Coworker share of N-specialists	0.224	0.009	0.283	0.987
	(0.006)	(0.007)	(0.006)	(0.000)
Coworker share of C-specialists	-0.263	0.004	0.124	-0.005
	(0.004)	(0.005)	(0.005)	(0.000)
Constant	0.229	0.215	0.127	0.004
	(0.002)	(0.002)	(0.002)	(0.000)
Panel B:				
Dependent variable: Generalist				
Coworker share of N-specialists	-0.023	-0.010	-0.417	-0.980
	(0.008)	(0.008)	(0.008)	(0.000)
Coworker share of C-specialists	-0.155	-0.003	-0.423	-0.974
	(0.007)	(0.008)	(0.008)	(0.000)
Constant	0.593	0.555	0.740	0.990
	(0.003)	(0.003)	(0.003)	(0.000)
Panel C:				
Dependent variable: Being C-specialist				
Coworker share of N-specialists	-0.201	0.001	0.134	-0.008
	(0.004)	(0.005)	(0.005)	(0.000)
Coworker share of C-specialists	0.418	-0.001	0.299	0.978
	(0.007)	(0.007)	(0.006)	(0.000)
Constant	0.178	0.230	0.132	0.007
	(0.002)	(0.002)	(0.002)	(0.000)
Observations (all panels)	731,946	731,946	731,946	731,946

Note: Dependent variable is own type, estimates are for the share of coworkers of different types. Reference is the share of generalists. Data are for 2005. At least 6 workers and at most 600 workers with measured skills are employed in each establishment. Three last columns show regression on simulated allocations across the actual establishment size distribution, see text for details. Standard errors are clustered at the establishment level.

Table 3: Leave-out mean regressions on two-dimensional worker types

	(1) N-Specialists	(2) Generalists	(3) C-Specialists
Panel A (High total ability). Dep. var. types:	High N-Specialist	High Generalist	High C-Specialist
Estimates:			
Coworkers N-Specialists	**0.075*****	-0.055***	-0.105***
	(0.004)	(0.004)	(0.003)
Coworkers C-Specialists	-0.098***	-0.027***	**0.223*****
	(0.004)	(0.006)	(0.006)
(reference: Generalists)			
Coworkers High ability	**0.075*****	**0.329*****	**0.184*****
	(0.004)	(0.006)	(0.004)
(reference: Mid ability)			
Coworkers Low ability	-0.078***	-0.127***	-0.039***
	(0.003)	(0.004)	(0.003)
Constant	0.072***	0.117***	0.023***
	(0.002)	(0.002)	(0.002)
Observations	731,946	731,946	731,946
Panel B (Mid total ability). Dep. var. types:	Mid N-Specialist	Mid Generalist	Mid C-Specialist
Estimates:			
Coworkers N-Specialists	**0.083*****	0.029***	-0.049***
	(0.004)	(0.005)	(0.003)
(reference: Generalists)			
Coworkers C-Specialists	-0.063***	-0.079***	**0.096*****
	(0.003)	(0.005)	(0.004)
Coworkers High ability	-0.078***	-0.211***	-0.027***
	(0.002)	(0.005)	(0.003)
(reference: Mid ability)			
Coworkers Low ability	-0.039***	-0.129***	-0.030***
	(0.003)	(0.005)	(0.003)
Constant	0.106***	0.355***	0.079***
	(0.002)	(0.004)	(0.002)
Observations	731,946	731,946	731,946

Note: Results from 9 different regressions (table continues on next page) where the worker types are dependent variables. Types are defined from the combination of indicators for C/N-Specialists vs generalist combined with indicators for total ability being low, mid or high. Explanatory variables are coworker averages of the C/N-specialists (generalists as the reference) and Low/High ability (mid ability as the reference). Estimates in **bold** are for the same type. Data are for 2005. At least 6 workers and at most 600 workers with measured skills are employed in each establishment. Standard errors are clustered at the establishment level.

Table 3: Leave-out mean regressions on two-dimensional worker types (Table 3, cont'd)

	(1) N-Specialists	(2) Generalists	(3) C-Specialists
Panel C (Low total ability) Dep. var. types:	Low N-Specialist	Low Generalist	Low C-Specialist
Estimates:			
Coworkers N-Specialists	**0.042*****	-0.005	-0.016***
	(0.003)	(0.004)	(0.002)
Coworkers C-Specialists	-0.051***	-0.034***	**0.032*****
	(0.003)	(0.003)	(0.003)
(reference: Generalists)			
Coworkers High ability	-0.085***	-0.141***	-0.045***
(0.002)	(0.003)	(0.002)	
Coworkers Low ability	**0.126*****	**0.263*****	**0.053*****
	(0.003)	(0.004)	(0.002)
Constant	0.076***	0.126***	0.047***
	(0.002)	(0.002)	(0.001)
Observations	731,946	731,946	731,946
R-squared	0.030	0.053	0.006

Note: See note in previous table.

Table 4: Detailed leave-out mean regressions on two-dimensional worker types

	(1) High N-sp.	(2) High C-sp.	(3) Mid N-sp.	(4) Mid C-sp.	(5) Low N-sp.	(6) Low C-sp.
Coworkers High N-Sp	**0.228*****	-0.040***	*0.015****	-0.066***	*-0.074****	-0.032***
	(0.009)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Coworkers High Gen.	0.088***	0.172***	-0.084***	-0.023***	-0.101***	-0.039***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
Coworkers High C-Sp	-0.078***	**0.487*****	-0.139***	*0.063****	-0.108***	*-0.027****
	(0.004)	(0.009)	(0.003)	(0.004)	(0.003)	(0.003)
Coworkers Mid N-Sp	*0.055****	-0.065***	**0.097*****	-0.057***	*0.043****	-0.016***
	(0.005)	(0.004)	(0.007)	(0.004)	(0.005)	(0.004)
Mid Gen (ref.)						
Coworkers Mid C-Sp	-0.055***	*0.164****	-0.089***	**0.118*****	-0.080***	*0.052****
	(0.005)	(0.005)	(0.005)	(0.007)	(0.005)	(0.004)
Coworkers Low N-Sp	*-0.060****	-0.060***	*0.012***	-0.074***	**0.188*****	0.016***
	(0.004)	(0.003)	(0.005)	(0.004)	(0.007)	(0.004)
Coworkers Low Gen	-0.073***	-0.044***	-0.036***	-0.026***	0.123***	0.062***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)
Coworkers Low C-Sp	-0.050***	*0.008**	-0.067***	*0.061****	0.010*	**0.118*****
	(0.005)	(0.004)	(0.006)	(0.006)	(0.006)	(0.008)
Constant	0.065***	0.029***	0.107***	0.076***	0.082***	0.043***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	731,946	731,946	731,946	731,946	731,946	731,946

Note: Results from 6 different regressions where the worker types are dependent variables. Types are defined from the combination of indicators for C/N-Specialists vs generalist combined with indicators for total ability being low, mid or high. Explanatory variables are coworker averages of the same combined attributes with mid-level generalists as the reference. Estimates in **bold** are for the same exact type. Estimates in *italics* are for workers with different ability levels but the same (C,N) specialisation. Data are for 2005. At least 6 workers and at most 600 workers with measured skills are employed in each establishment. Standard errors are clustered at the establishment level.

4.3 Sorting over time

In this section we document how labor market sorting has changed over time. The purpose is to illustrate the extent to which the general time trends are consistent with a process of *unbundling* as outlined by CK. Because our data do not cover all cohorts, changes over time will also generate changes in the age-composition. To ensure that this does generate spurious patterns, we follow Håkanson et al. (2021) and zoom in on a specific age group that we can follow consistently over time (ages 40 to 45) for the baseline analysis.

We estimate a version of equation 7 where the covariates of interest are interacted with time trends covering our 1996-2013 data period. The model accounts for year dummies and, for robustness tests, other plant-level controls. The model can thus be written as:

$$Y_{ijt}^{\tau} = a + g^{C,\tau} * t * C_{jt}^{-i} + g^{N,\tau} * t * N_{jt}^{-i} + b^{C,\tau} * C_{jt}^{-i} + b^{N,\tau} * N_{jt}^{-i} + D_t + X_{ijt}\beta^{\tau} + \epsilon_{ijt}^{\tau}$$

where Y_{ijt}^{τ} represent the type of worker i, in year $t = Year - 2005$ employed at workplace j. Types will be indicators for being a specialist of type $\tau = C, N$, or a generalist. C_{jt}^{-i} and N_{jt}^{-i} measures the share of coworkers that C-specialists and N-specialists (the residual type is generalists). D_t are time dummies and X_{ijt} reflect additional controls. If concentration has increased we expect positive estimates for $g^{C,C}$ (i.e. a growing positive impact of coworker C on Y_{ijt}^{C}) and $g^{N,N}$, but negative estimates for $g^{N,C}$ and $g^{C,N}$.

The estimates are displayed in table 5. Panel A shows the estimates for the outcome Y_{ijt}^{C} and panel B for Y_{ijt}^{N}. Column (1) is the baseline specification without any controls except for time dummies. The estimates suggest that sorting has increased over time as C-specialists increasingly work with C-specialists and less with N-specialists. The converse is true for N-specialists. In column (2), we add controls for occupations. The sample here is considerably smaller as we do not observe occupations for all workers. The picture is, however, very similar. In column (3), we change the concept of coworkers and instead focus on other workers in the same *job* defined as occupation*establishment as in Fredriksson et al. (2018). Here the sample is reduced even further as we require that there are at least 5 other employees in the same job, but the estimated time-trends show a similar pattern as in the main specification. In Column (4), we return to the baseline model, but add controls for establishment size (8 groups) and for the share of low- and high-skilled workers in the establishment. The results remain robust. In Column (5), we remove low-tenured workers as in Fredriksson et al. (2018) without much change in results. Finally, in column (6), we widen the age span to also include workers aged 35 to 50 which makes the estimates more modest, although the qualitative results remain.

We display the patterns graphically in Figures 2 to 4. Figure 2 shows the exposure of specialists to other specialists over time. The graph shows that regardless of the age at which we evaluate the effect, each cohort is more exposed to the own type than the previous cohort. Within cohorts we do not see that age matters, however. This pattern suggests that the changing nature of sorting across time does not arise because workers adjust their sorting patterns as their career evolves.

We also include results for sorting across *jobs* to illustrate that time-patterns are nearly identical. The share of similar specialists at the establishment increased from 25 to 27.5 percent over the period, whereas the share of similar specialists in the job increased from 26 to 29 percent. Note that although these numbers may appear low in levels, it is because we define most workers to be generalists. To recap, 55 percent of workers are generalists, 23 percent are C-specialists and 22 percent are N-specialists.[m]

An interesting feature of the underlying process is that the trend increase in sorting is much clearer among specialists than among generalists. Thus, sorting appear to mostly increase at the extremes. To illustrate this point, we let Figure 3 show that it is becoming increasingly rare that specialists work with specialists of the *opposing* type (excluding generalists), whereas Figure 4 shows that the concentration of generalists has remained fairly constant over time.

[m]In the appendix, we show that these shares are stable across test cohorts.

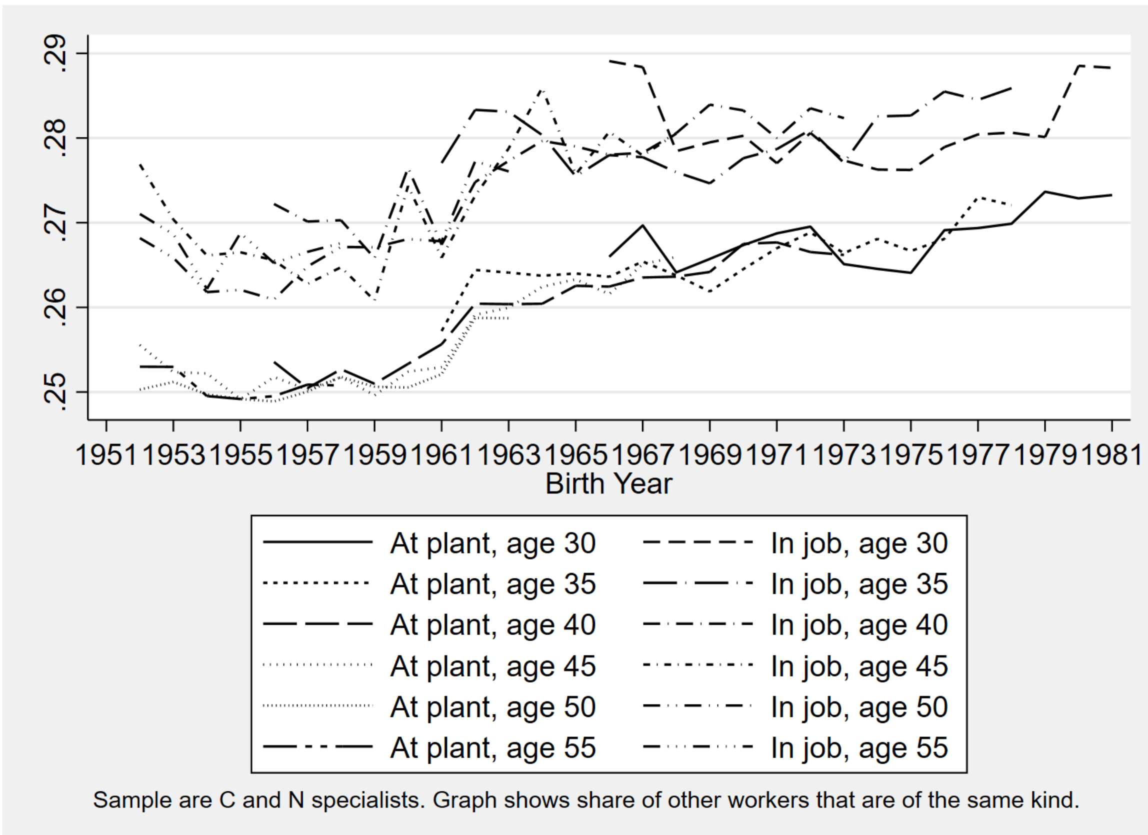

Figure 2: Concentration of similarly specialized coworkers over time (cohorts)

Note: The figure shows the share of specialist's coworkers that are of the same type, by cohort and age

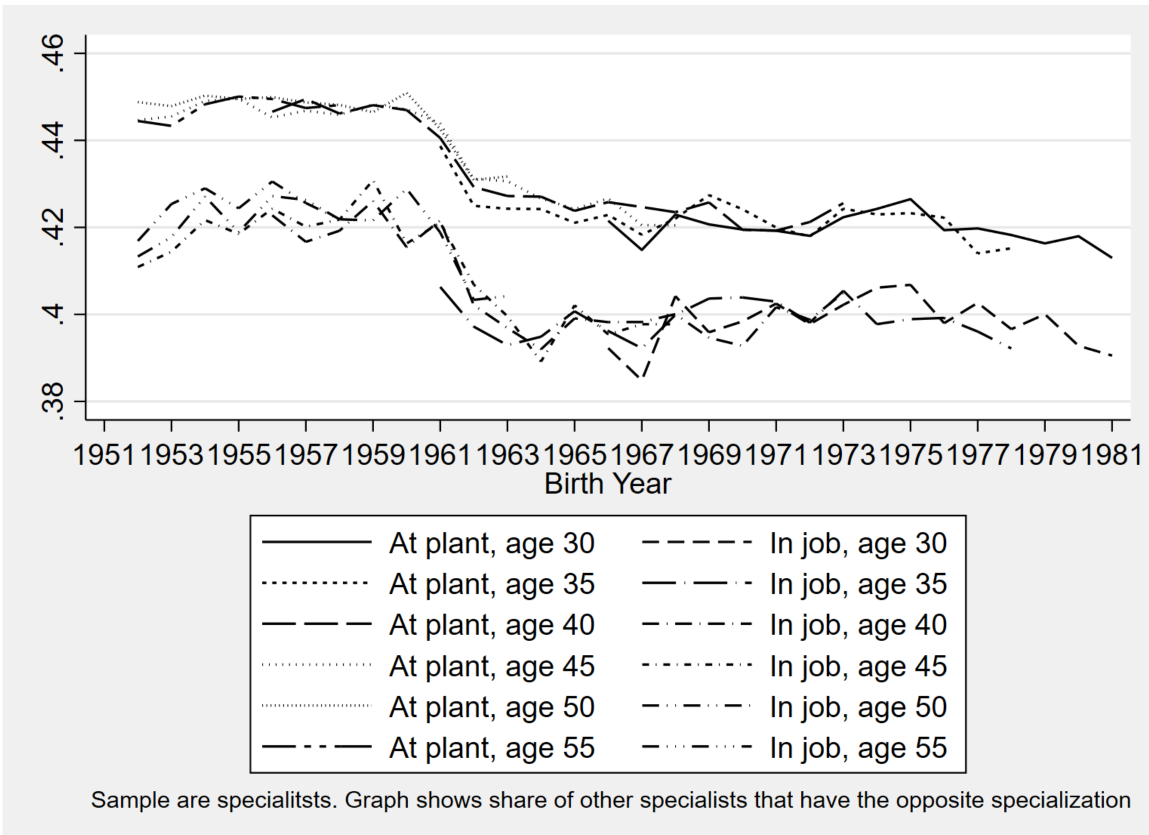

Figure 3: Exposure to coworkers of the *opposite* type among specialists

Note: The figure shows the share of specialists that are of the opposite specialisation, by cohort and age

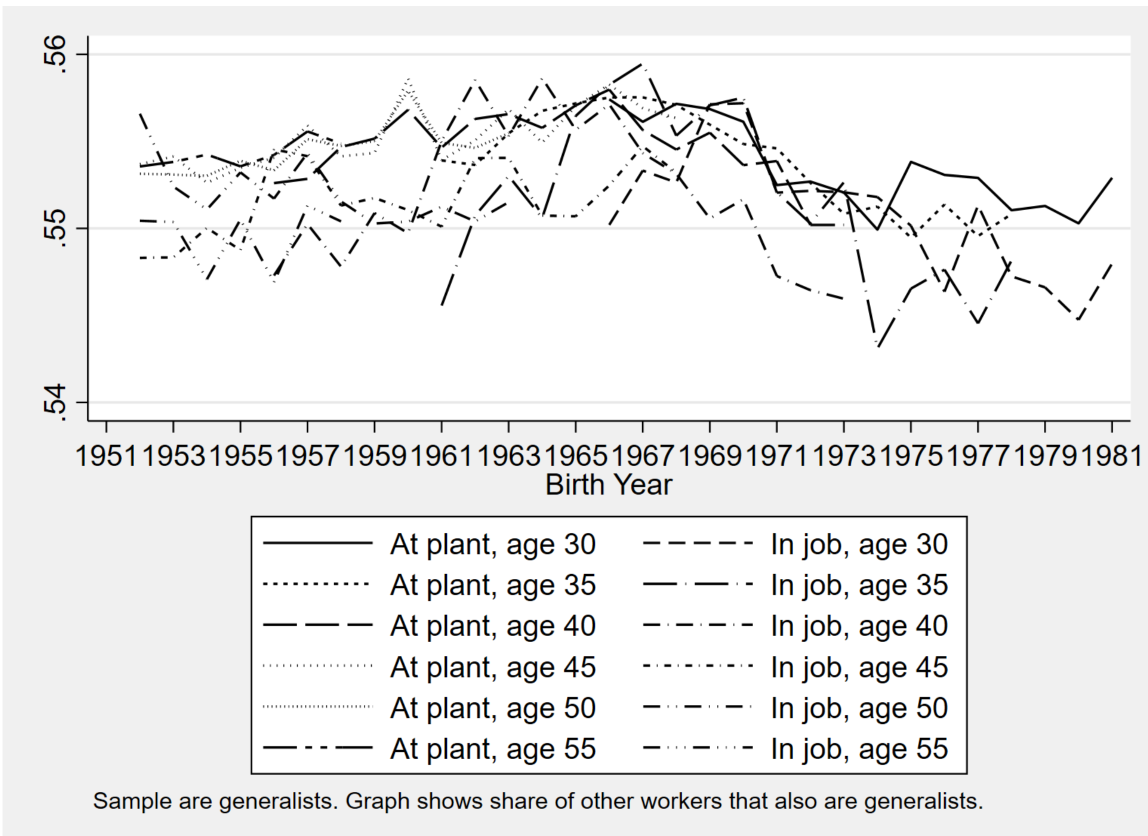

Figure 4: Concentration of generalists over time (cohorts)

Note: The figure shows the share of generalists' coworkers that also are generalists, by cohort and age

Table 5: : Specialist coworkers increasingly predict same-type specialists

Panel A Outcome: C-specialist	(1) Base	(2) Control for Occupation	(3) Coworkers in Job	(4) Additional Controls	(5) Only Tenured	(6) Broader Age Span
Time*C-spec.	0.008***	0.008***	0.006***	0.008***	0.009***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Time*N-spec.	-0.003***	-0.002	-0.003***	-0.003***	-0.003***	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
C-specialists	0.415***	0.269***	0.455***	0.345***	0.349***	0.361***
	(0.006)	(0.008)	(0.008)	(0.006)	(0.007)	(0.005)
N-specialists	-0.203***	-0.122***	-0.241***	-0.168***	-0.171***	-0.170***
	(0.004)	(0.007)	(0.006)	(0.004)	(0.005)	(0.003)
Low-skilled cow.				-0.025***	-0.023***	-0.022***
				(0.004)	(0.005)	(0.002)
High-skilled cow.				0.109***	0.121***	0.114***
				(0.005)	(0.005)	(0.003)
N	2,317,898	1,255,003	896,931	2,317,898	1,656,627	8,787,016
Panel B Outcome: N-specialist	(1) Base	(2) Control for Occupation	(3) Coworkers in Job	(4) Additional Controls	(5) Only Tenured	(6) Broader Age Span
Time*N-spec.	0.004***	0.002	0.003**	0.004***	0.004***	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Time*C-spec.	-0.003***	-0.004***	-0.002*	-0.003***	-0.003***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
N-specialists	0.227***	0.144***	0.264***	0.200***	0.203***	0.208***
	(0.005)	(0.008)	(0.008)	(0.006)	(0.007)	(0.004)
C-specialists	-0.251***	-0.147***	-0.261***	-0.198***	-0.198***	-0.207***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.005)	(0.003)
Low-skilled cow.				0.019***	0.019***	0.014***
				(0.004)	(0.005)	(0.003)
High-skilled cow.				-0.080***	-0.085***	-0.085***
				(0.004)	(0.004)	(0.003)
N	2,317,898	1,255,003	896,931	2,317,898	1,656,627	8,787,016

Note: Dependent variable is a an indicator for being a C-specialist in panel A (N-specialist in Panel B). Subjects are 40 to 45 years old. Explanatory variables are share of coworkers that are C/N-specialists interacted with time. Normalised so that main effects of coworkers reflect 2005. All specifications include year dummies. Col (2) also controls for occupation dummies at the 3-digit level (sample requires that occupations are observed). Column (3) measures coworkers in job (occupation*establishment) instead (sample requires at least 5 coworkers in job). Columns (4) to (6) controls for eight plant size dummies and the share high/low skilled among coworkers. Column (5) only include workers with at least 3 years of tenure. Column (6) widens the age span to 35 to 50. Standard errors clustered at the establishment level. Data cover 1996-2013.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Skills and wages

In this section we use our data to document how sorting is related wages. In particular, we are interested in the extent to which market returns to each skill is higher in settings where the technology is likely to be more intensively using that particular skills.[n] We define employer from co-worker skill-specializations as explained in the data section. In terms of theory, our data aspire to represent the set of firms where the α-parameter in the production function makes the firm want to specialize in either one or the other type of worker, i.e. C- and N-establishments. We then interact the establishment type with the specialization of the worker and estimate if the returns to being a C-intensive worker are higher if the employment pattern is such that the firm appears to be using a C-intensive technology (and conversely for N). As we are interested in the sorting of specialists, we exclude generalist establishments at this stage. The model controls for the overall impact of level of skills in each dimension through dummies for each score on the discrete 1-9 scale:

$$lnW_{ijt} = \gamma^{C}_{C(i)} + \gamma^{N}_{N(i)} + D^{N-plant}_{jt} + b^{N}_{j} * D^{N-in-N}_{ijt} + b^{C}_{j} * D^{C-in-C}_{ijt} + X_{ijt}\beta + \epsilon_{ijt} \qquad (8)$$

where lnW_{it} represents the wage of worker i in establishment j in year t and where the γ's are dummies for each value of C and N skills. The two key variables of interest are the are the interaction terms D^{N-in-N} (for N-specialists in N-establishments) and D^{C-in-C} which captures the additional returns to N-skills in N-intensive employers, and C-skills in C-intensive employers, respectively. The vector of control variables will always include time dummies, plant size dummies and an age polynomial.

The results are presented in Table 6. Throughout, the results suggest that the wages in the market sections where employers rely intensively on C-skills also pay higher returns to these skills. Similarly, the results suggest a premium for N-skills in market segments dominated by N-intensive firms. These patterns are robust to controls for occupations, analysing data at the job-level, controlling for very detailed skills, focusing on tenured workers, or zooming in on the center year of 2005. In panel B, we show that the message is identical is we instead use monthly earnings, which allows us to expand the data to include all observations instead of just the half where we observe wages. Panel C zooms in on establishments that are highly specialized (the most common worker type is either C or N-Specialists). All establishment-level results are robust.[o]

[n] Some evidence in this direction at the *job*-level is presented in Fredriksson et al. (2018), with a focus on new hires, but here we revisit the issue at the *establishment* level for the *stock* of employees.

[o] The one deviating estimate in the table is for N-specialists in the the job-level analysis.

Table 6: Returns to specific skills are higher when coworkers are specialist in those skills

	(1) Base	(2) Control for Occupation	(3) Coworkers in Job	(4) Interacted Skills	(5) Only Tenured	(6) Only 2005
Panel A: Wages						
C-sp. in C-est.	0.027***	0.009***	0.040***	0.027***	0.028***	0.020***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.007)
N-sp. in N-est.	0.016***	0.005*	0.023***	0.021***	0.017***	0.009
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.007)
C-establishment	0.087***	0.020***	0.126***	0.089***	0.092***	0.075***
	(0.004)	(0.003)	(0.005)	(0.004)	(0.005)	(0.006)
N	1,458,790	1,432,159	1,259,521	1,458,790	961,640	85,291
Panel B: Earnings						
C-sp. in C-est.	0.036***	0.009***	0.044***	0.036***	0.032***	0.033***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.007)
N-sp. in N-est.	0.023***	0.005*	0.026***	0.029***	0.025***	0.019***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.007)
C-establishment	0.081***	-0.002	0.108***	0.083***	0.082***	0.067***
	(0.003)	(0.003)	(0.005)	(0.003)	(0.004)	(0.006)
N	2,945,409	1,432,159	1,259,521	2,945,409	1,899,162	168,815
Panel C: Specialized						
C-sp. in C-est.	0.038***	0.014***	0.129***	0.038***	0.035***	0.038***
	(0.004)	(0.005)	(0.007)	(0.004)	(0.005)	(0.010)
N-sp. in N-est.	0.024***	0.009*	-0.067***	0.031***	0.028***	0.006
	(0.005)	(0.005)	(0.008)	(0.005)	(0.006)	(0.011)
C-establishment	0.103***	-0.010**	0.026***	0.104***	0.103***	0.084***
	(0.004)	(0.004)	(0.009)	(0.004)	(0.005)	(0.008)
N	1,297,390	556,605	616,218	1,297,390	824,400	73,423

Note: Dependent variable is log wages. Control variables are the dummies for C-skills (1 to 9) and N-skills (1 to 9), dummies for being a C- or an N- specialist, as well as year dummies, an age polynomial and eight plant size dummies. Displayed estimates are for C-specialists in C-establishments (and conversely for N-specialists). Sample excludes establishments where the majority of workers are generalists. Specialization of establishment is based on the specialization among coworkers. Column (2) adds controls for occupations. Column (3) performs the analysis at the job (occupation times establishment) level instead. Column (4) interacts the skills controls (C,N) into 81 groups Column (5) only include workers with at least 3 years of tenure. Column (6) zooms in on data for 2005. Panel A uses wages that only exist for a 50 percent sample. Panel B and C uses monthly earnings instead. Panel C only includes highly specialized establishments where C- or N-specialists are the most common type of worker. Samples in panels A and B overlap when conditioning on observed occupations (col 2 and 3). Standard errors clustered at the establishment level. Data cover 1996-2013.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 The growing wages of generalists

According to Choné and Kramarz (2021), a process of "unbundling" should lead to increased market wages of generalists relative to specialists. The reason is that the bundling constraint depresses the market wages of the generalists. In order to test if the evolution of the overall wage structure concur with this prediction, we estimate wage regressions where our variable of interest is the interaction between time and a dummy for being a generalist (defined as above). The model controls for overall wage growth through year dummies and include a fixed effect for each "detailed type" of worker defined as the interaction of raw cognitive and non-cognitive scores (thus, 81 types). Our identification thus comes from the relative wage changes among workers on the generalist skill-diagonal relative to other workers. The model can be written as:

$$lnW_{it} = \gamma_{CN(i)} + b^G * G_i * t + D_t + X_{ijt}\beta + \epsilon_{it} \tag{9}$$

where lnW_{it} represents the wage of worker i in year t, and where $\gamma_{CN(i)}$ is the fixed effect for the worker type. We estimate the model for 40 to 45 year old workers as above, and allow for a set of control variables X_{ijt} that will vary across specifications. We provide separate estimates for the sample of workers who are "well matched" (or, not bunched) in the sense that the type of the worker correspond to the type of the firm (e.g. C-specialists working in C-establishments, see data section for definitions).

The estimates are displayed in table 7. Panel A shows the estimates for the overall population and Panel B zooms in on the "well-matched" sample. Column (1) is the baseline specification without any controls except for time dummies and the type-specific fixed effects. The estimates suggest that wages of generalists have grown more than wages for workers in general. The magnitudes suggest a modest 1.2 percent additional wage increase across one decade.

The following columns establish that the qualitative conclusion is very robust. In column (2), we add controls for occupations interacted with the worker type. In Column (3), we keep the controls for occupations and also introduce a set of controls for time trends that interact each possible value of N and C with time (thus, 18 trends) as well as controls for establishment size (8 groups). To ensure that the results are not driven by ceiling effects at the top, we let column (4) show results for the baseline model but where we only include "mid skilled" workers that all have total skills (C+N) in the range 9 to 11. In column (5), we instead remove low-tenured workers and in column (6), we widen the age span to include all workers aged 35 to 50.

Panel B, use the same set of specifications but only include those workers who are employed in establishments where the majority of other workers are of the same broad type (Generalist, C-specialist, N-specialist). Estimates are unchanged in qualitative terms, but the magnitudes are much larger, suggesting that wages of well-matched generalists have grown by 2-3 percent more across a decade than wages of well-matched specialists. This amounts to around one-tenth of the average real wage growth during the period.[p]

Table 7: Generalists' relative wages grow over time

Panel A All workers	(1) Base	(2) Control for Occupation	(3) +additional controls	(4) Only Mid-skilled	(5) Only Tenured	(6) Broader Age Span
Generalist×time	0.0012***	0.0007***	0.0006***	0.0005**	0.0011***	0.0009***
	(0.0002)	(0.0001)	(0.0001)	(0.0003)	(0.0002)	(0.0001)
N	1,281,151	1,255,003	1,255,003	476,822	928,127	4,723,064
Panel B Well matched sample only						
Generalist×time	0.0031***	0.0020***	0.0014***	0.0019***	0.0024***	0.0019***
	(0.0006)	(0.0004)	(0.0004)	(0.0006)	(0.0007)	(0.0006)
N	654,687	641,005	641,005	266,173	476,688	2,415,481

Note: Dependent variable is log wages. Subjects are 40 to 45 years old. Estimates are for interaction between year and a generalist dummy. All specifications include year dummies and control for 81 fixed effects for interactions between measured C (1 to 9) and N (1 to 9). Col (2) (3) have more detailed fixed effects that also interacts with occupation dummies at the 3-digit level (sample requires that occupations are observed). Column (3) controls for eight plant size dummies and 18 additional time trends, each interacted with one of the possible 9 values of C and N. Column (4) only include workers with C+N below 9 and 11. Column (5) only include workers with at least 3 years of tenure. Column (6) widens the age span to 35 to 50. Standard errors clustered at the establishment level. Data cover 1996-2013.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[p] It is possible that we find larger estimates for well-matched workers because this sample is better at capturing the true worker types, and therefore also better reflects the market valuations. This could be the case, e.g., because of deviations between our measured worker-level skills and the true skills of the workers (because of the aggregation into the integer scale, or measurement errors). Workers that are misclassified should be more likely to turn up as poorly matched (and hence excluded in Panel B) if errors are uncorrelated across workers within establishments.

7 Conclusions

This empirical paper has illustrated patterns of worker sorting and the relationship between sorting and wages with Choné and Kramarz (2021) as the theoretical foundation. The results show that workers are sorted across establishments in both the vertical dimension (skill intensity) and the horizontal dimension (specialization). Horizontal sorting dominates at the top of the ability distribution. High-level specialists are *less likely* to work with the opposing type of specialists than under random sorting, but *more likely* to work with mid-level specialists of the same type.

Furthermore, the paper shows that sorting has increased over time. Every cohort of specialists is more likely to work with specialists of the same type, and less likely to work with specialists of the opposing type, than the previous cohort evaluated at the same age. In terms of wages, we show that the wage returns to specific skills is higher in the more specialized market segments. Furthermore, the results document a secular trend of growing relative wages for generalists relative to specialists. The two trends we document (increased sorting on relative skills and growing wages of generalists) are both fully in line with a process of "unbundling" as outlined by Choné and Kramarz (2021). If new markets open up that allow workers to sell their skills separately, generalist wages will be less under pressure from competing specialist workers, thus allowing their wages to increase.

Some (but far from all) of our results mimic conclusions drawn in earlier or parallel work using similar data, most notably Fredriksson et al. (2018), Håkanson et al. (2021) and Böhm et al. (2020). But we add to the literature by compiling the results in one unified empirical setting, by adding a set of important missing pieces, and by setting the results in context by relating them to what we believe to be a more comprehensive theoretical framework.

The presented results are distinctively reduced form in nature and the analysis is purely descriptive. A natural next step is to incorporate more detailed data on the firm side and use these data to estimate a structural model of worker-firm matching and to assess the model-performance in settings where we can observer clear cases of "unbundling". This is the direction of our ongoing work.

References

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

Böhm, M., K. Esmkhani, and G. Gallipoli (2020): "Firm Heterogeneity in Skill Returns," *CEPR Discussion paper*, 15480.

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

Choné, P. and F. Kramarz (2021): "Matching Workers' Skills and Firms' Technologies: From Bundling to Unbundling," *CREST Working Paper*, 202110.

Edmond, C. and S. Mongey (2020): "Unbundling labor," *Mimeo*.

Fredriksson, P., L. Hensvik, and O. Nordstrom Skans (2018): "Mismatch of talent: Evidence on match quality, entry wages, and job mobility," *American Economic Review*.

Guvenen, F., B. Kuruscu, S. Tanaka, and D. Wiczer (2020): "Multidimensional Skill Mismatch," *American Economic Journal: Macroeconomics*, 12, 210–244.

Håkanson, C., E. Lindqvist, and J. Vlachos (2021): "Firms and skills: the evolution of worker sorting," *Journal of Human Resources*, 56, 512–538.

Heckman, J. and J. Scheinkman (1987): "The importance of bundling in a Gorman-Lancaster model of earnings," *The Review of Economic Studies*, 54, 243–255.

Hensvik, L. and O. N. Skans (2020): "The Skill-Specific Impact of Past and Projected Occupational Decline," *IZA Discussion Paper*, 12931.

Lindenlaub, I. (2017): "Sorting multidimensional types: Theory and application," *The Review of Economic Studies*, 84, 718–789.

Lindqvist, E. and R. Vestman (2011): "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment," *American Economic Journal: Applied Economics*, 3, 101–28.

Mood, C., J. O. Jonsson, and E. Bihagen (2012): "Socioeconomic persistence across generations: The role of cognitive and non-cognitive processes," in *From Parents to Children: The Intergenerational Transmission of Advantage*, ed. by J. Ermisch, M. Jäntti, and T. M. Smeeding, Russell Sage Foundation, chap. 3.

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

Skans, O. N., P.-A. Edin, and B. Holmlund (2009): *Wage dispersion between and within plants: Sweden 1985-2000*, 217–260.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019): "Firming up Inequality," *The Quarterly journal of economics*, 134, 1–50.

A Appendix

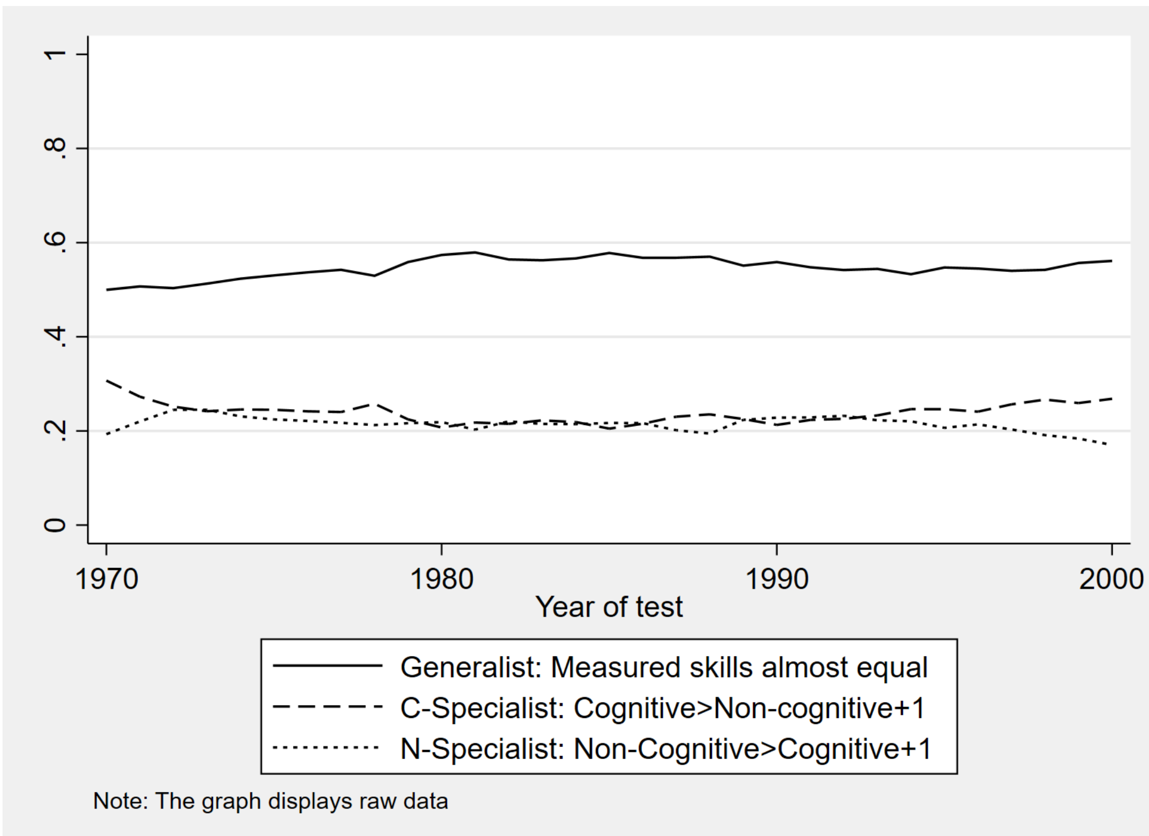

Figure 5: Year of test and ability types scores

Note: The figure shows the types as defined from test scores in the raw data as a function of test year.

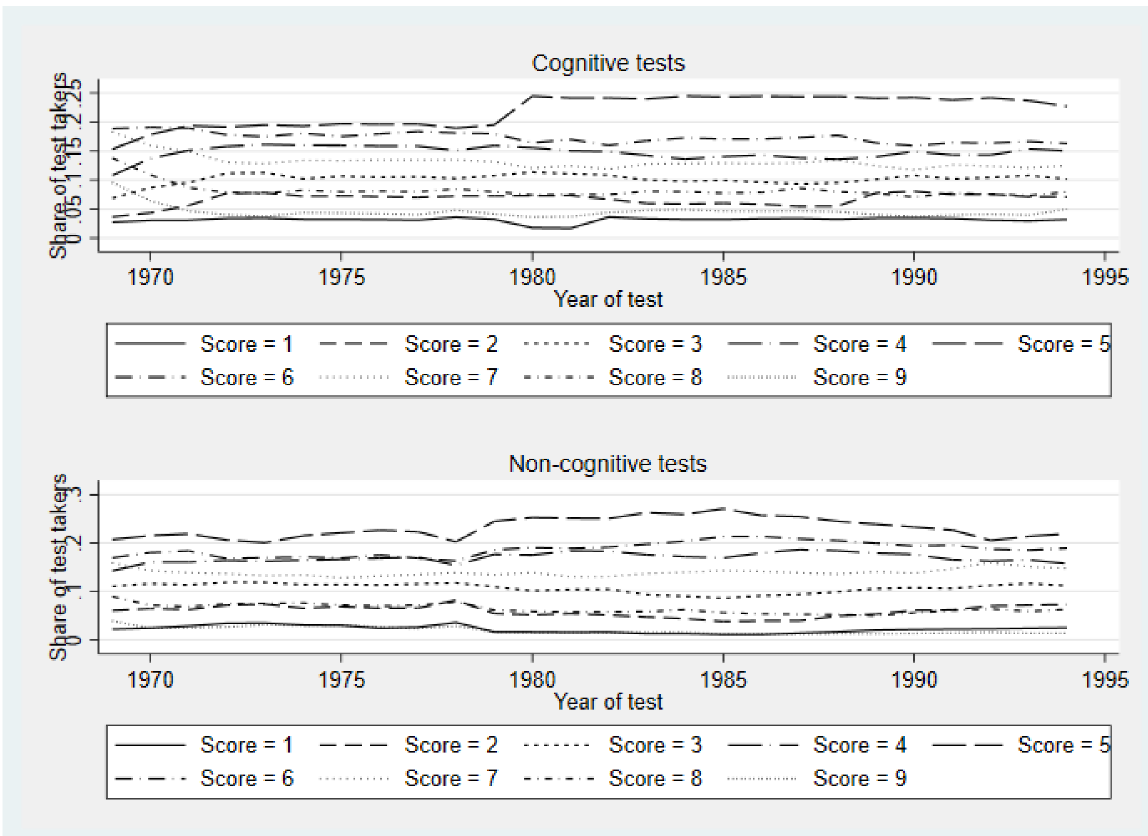

Figure 6: Year of test and measured ability scores

Note: The figure shows the test score results in the raw data as a function of test year. The increase in the number of test takers receiving a cognitive scores of 5 in 1980 corresponds to a change i test protocol.