Measuring On-the-job Learning Rates in Multidimensional Skills

Mariia Bondar, Nikolaos Errikos Melissinos*

February 10, 2022

Abstract

In this paper, we introduce a framework with multidimensional skills, in which we estimate how fast skills accumulate due to on-the-job experience. We model an individual's wage as a weighted sum of her productivities in different skills. We call this skill-specific productivity expertise. Since expertise is not directly observable, we proxy this variable with skill-specific experience, which depends on the years of labor market experience across different occupations and the importance of the corresponding skill in those occupations. We compute skill-specific experience using the data on occupational skill requirements from O*NET. We then estimate the wage equation using skill-specific experience to evaluate the speed of expertise accumulation (learning rate) in different skills. We find that expertise in different skills grows with skill-specific experience and that different skills exhibit different learning rates. (*JEL* E24, J24)

*Goethe University, Frankfurt am Main. We thank Alexander Bick, Sena Coskun, Nicola Fuchs-Schündeln and Khanh Hoang, as well as the participants of FQMG Brown Bag seminars, SAFE Brown Bag seminars, 2021 Frankurt-Mannheim Macro Workshop and 16th BiGSEM Doctoral Workshop on Economics and Management for helpful comments and discussions.

1 Introduction

Recent research on workers' productivity concludes that skills have a multidimensional character, and that summarizing skill measures by a single number leads to loss of information. This fact forces us to include various skill categories in our analyses of the labor market. At the same time, it opens up a lot of questions. How many and which skills are relevant? What is the return to different skills? What dynamics do different skills follow? Which skills matter for different occupations? What is the relationship between skills and formal education? How do people differ with respect to skills?

In this paper, we introduce a framework with multidimensional skills, in which we are able to estimate how fast skills accumulate due to on-the-job experience. Our analysis is based on the skill-weights approach developed by Lazear (2009): all skills are general, but the uses of these skills are occupation-specific. Following Lazear (2009), we assume that an individual's wage is equal to a weighted sum of her productivities in different skills. We call this skill-specific productivity *expertise*. Since expertise is not directly observable, we proxy this variable with an empirical measure that we call *skill-specific experience*. Skill-specific experience depends on the years of labor market experience across different occupations and the importance of the corresponding skill in those occupations. First, we use the data on occupational skill requirements from O*NET Program database and the data on occupation histories of young individuals from the National Longitudinal Survey of Youth 1997 (NLSY97) to compute skill-specific experience. We then estimate the wage equation using skill-specific experience as an empirical counterpart of unobserved expertise to evaluate the speed of expertise accumulation (*learning rate*) in different skills.

Our work is related to a vast and fast-growing literature on human capital accumulation. Starting with a seminal paper by Becker (1962), the human capital research has shifted focus from firm- and industry-specific human capital to occupation- and task- (or skill-) specific human capital (for an overview of the history of thought in human capital research see Sanders and Taber (2012)). We contribute to the last strand of this literature, which concentrates on skill-specific human capital, specifically, on the returns to multidimensional skills (Ingram and Neumann (2006), Poletaev and Robinson (2008)), and the process of multidimensional skill accumulation (Yamaguchi (2012), Sanders (2016)).

Our approach is most similar to Gathmann and Schönberg (2010), Lise and Postel-Vinay (2020) and Guvenen, Kuruscu, Tanaka and Wiczer (2020). Similarly to us, Gathmann and Schönberg (2010) model expertise in different tasks (which they call "task productivity") as a weighted sum of task tenure.[1] Task tenure is the closest concept to our skill-specific experience. The authors use the task data from the German Qualification and Career Survey combined with the German Employee Panel, and find that task tenure is an important predictor of wage growth.

Lise and Postel-Vinay (2020) use data on occupational skill requirements from O*NET and data on occupation histories from NLSY79. They estimate a structural search-theoretic model of workers' careers using indirect inference. They find that expertise in different skills exhibits different characteristics. In addition, they find that skill mismatch is costly and that the costs vary depending on the skill. Finally, they model mismatch as the nonnegative difference between the occupational skill requirement and the corresponding expertise. In their setting mismatch fades the more a person stays at a specific occupation, because with time expertise adjusts to the occupation the worker is exercising.

Guvenen *et al.* (2020) use the same data as Lise and Postel-Vinay (2020). In their setup, individuals choose an occupation in which skill requirements are equal to their own perceived skill-specific abilities. The authors estimate a series of regressions, and the regressors that they use are a weighted average of skill-specific variables. This means that the regressors themselves, and hence their estimates, are not skill-specific. Like Lise and Postel-Vinay (2020) they also focus on the concept of skill mismatch. In their case, skill mismatch is defined as the absolute value of the difference between individual skill-specific learning rates and the corresponding occupational skill requirements. Unlike Lise and Postel-Vinay (2020), the skill mismatch stays constant as long as a worker stays in the same occupation. Guvenen *et al.* (2020) also find that mismatch is predictive of wages and the effect lasts for subsequent occupations.

In this paper, we use O*NET as our source of data for occupational skill requirements and NLSY97, which is a newer sample compared to NLSY79, for data on occupation histories. Like the other authors, we also propose a learning process for multidimensional skills. Unlike the other authors, our process is not based on skill mismatch but rather on skill-specific experience, which corresponds to time spent practicing a skill on the job. The closest NLSY-based measure to our skill-specific experience is what Guvenen *et al.* (2020) call cumulative past mismatch. However, unlike skill mismatch, skill-specific expe-

[1] When comparing to other authors we use terminology according to the definitions that we give in this paper.

rience depends only on the occupation history of the worker and not on the discrepancy between her own ability on the one side and occupational skill requirements on the other side. This means that using skill-specific experience allows us to not rely on test scores, as we can use skill-specific experience directly as a proxy of expertise. These differences lead to a relatively different interpretation of our estimates, as we are able to estimate parameters corresponding to skill-specific learning rates. Like the previous authors we also find significant results even after controlling for covariates, and we find heterogeneity between the learning rates of different skills.

O*NET database of occupational skill requirements contains more than a hundred different skill categories. Our paper contributes to the literature that uses factorization methods on O*NET data to derive basic skills out of a large space of occupation-related characteristics (see, for instance, Ingram and Neumann (2006), Yamaguchi (2012), Poletaev and Robinson (2008), Lise and Postel-Vinay (2020), Guvenen *et al.* (2020)). We use Nonnegative Matrix Factorization, a method novel to this literature, to reduce the O*NET data to four basic skills, and interpret them as social, physical, technical, and cognitive skills based on the correlation analysis.

The rest of this paper is organized as follows. Section 2 describes the model and introduces the concepts of expertise and skill-specific experience. It also outlines the estimation strategy and analyzes the sources of potential biases. We describe the data used in the estimation in Section 3. Section 4 explains our dimensionality reduction procedure. Section 5 presents the main results, while Section 6 shows the results of the heterogeneity analysis. The robustness checks and additional results are described in Section 7. Finally, Section 8 concludes.

2 Theoretical foundations

Our model is based on a skill-weights approach to firm-specific human capital, suggested by Lazear (2009). The author complemented the seminal paper on human capital by Becker (1962) by suggesting a way to explicitly model specific human capital. According to him, there are no firm-specific skills. All skills are general; all firms are choosing skills from a common pool, but attach different weights to the selected skills. In other words, weights are firm-specific, not skills.

By Lazear's theory of firm-specific capital, for instance, cognitive skills (A) and social skills (B) are general. However, a bank and a retail store may assign different weights to each of these skills in the applicant's portfolio. Lazear (2009) summarizes his approach with a following expression for the output of firm i:

$$Y_i = \lambda_i A + (1 - \lambda_i)B \tag{1}$$

where λ_i is a weight of skill A in the output of firm i. The worker's wage is then determined by Nash bargaining as a share of total output Y_i.

In this paper, we apply the skill-weights idea of Lazear to a setting where skill weights λ are occupation-specific, not firm-specific.[2] We then model log wage as a weighted sum of an individual's *expertise* in all skills:

$$w_{io} = \sum_{s=1}^{S} \lambda_{so} e_{is} \tag{2}$$

The wage of individual i in her current occupation o is equal to a weighted sum, where e_{is} represents the expertise of individual i in skill s, and λ_{so} represents the weight of skill s in occupation o. As a quantity, expertise captures the level of productivity in the respective skill. Skill weight λ_{so} corresponds to the importance of skill s for occupation o.[3] The more a skill is required by some specific occupation, the more important the corresponding expertise becomes to determine total productivity. For example, both an engineer and a high-school teacher might need mathematics, but the extent to which this skill is required can be different across the two occupations. Following Guvenen *et al.* (2020), we assume that there is perfect competition among occupations, such that an individual's output perfectly equals her remuneration.

[2] Kambourov and Manovskii (2009) show that occupation-specific human capital is much more important in determining wages than firm- or industry-specific human capital. The authors conclude that human capital is occupation-specific.

[3] In contrast to Lazear (2009), in our model occupation-specific weights λ do not sum up to 1. Section 4 elaborates on this point.

2.1 Estimating the wage equation

If expertise e_{is} were observable, we could use skill weights λ_{so} to estimate a regression based on equation 2.[4] In particular, we could run the following regression:

$$w_{io} = a_1 \lambda_{1,o} e_{i,1} + a_2 \lambda_{2,o} e_{i,2} + \ldots + a_S \lambda_{S,o} e_{i,S} + u_{io} \quad (3)$$

where $\lambda_{s,o} e_{i,s}$ for $s = 1, \ldots, S$ are the regressors and u_{io} is some process with zero mean. To the extent that (a) wages are exclusively driven by occupation-specific productivity and (b) the measures for the weights λ_{so} are also perfect, the estimates of all the parameters a_s should be equal to 1 and the regression should exhibit a perfect fit.[5] Crucially, if the weights are positively correlated with expertise, this in itself does not lead to biased estimates. Furthermore, even if the weights are imperfectly measured such that only $\tilde{\lambda}_{so} = \lambda_{so} + u_{\lambda,so}$ is observed, one can show that the estimates of a_s will be biased towards 0. This implies that the regression will produce an estimate of the lower bound of the effect. Due to symmetry, the same argument can be used when the weights are well-measured but there is measurement error in the data on expertise. Hence, as long as one has a proxy of expertise for each different skill, they can produce an estimate of the lower bound of the true effect.[6]

Guvenen *et al.* (2020) use a similar argument to justify the use of test scores as measures of true skill-specific ability, which is not directly observable. Instead, we will use a new variable to proxy for expertise – skill-specific experience.

Skill-specific experience as a proxy of expertise

Skill-specific experience is the cumulative amount of time spent developing a skill during a work career. In general, labor market experience is an important driver of skill accumulation, simply because people spend most of their adult life working. We use skill-specific experience as a proxy for unobserved expertise, as we think that workers improve their expertise in skills on the job through learning-by-doing while performing their everyday

[4] In this paper we use data from O*NET to construct the skill weights, as will be explained later.

[5] If the weights are measured in different scales, then the estimates of the regression will naturally be scaled according to the corresponding scale factors.

[6] Here we assume that the regressors $\lambda_{s,o} e_{i,s}$ are not correlated with each other. Otherwise the measurement error of one regressor can contaminate the estimates of the others.

tasks.[7] Clearly, some occupations develop specific skills more than others, so it is reasonable to distinguish between different kinds of skill-specific experience. Intuitively, a person that has exercised an occupation with high requirements for cognitive skills will have a high experience specific to cognitive skills. At the same time, she may have very little experience specific to technical skills, if her cognitive-intensive occupation did not involve a lot of technical tasks. We denote skill-specific experience with $\tilde{y}_{is}$ and we use the following definition:

$$\tilde{y}_{is} = \sum_{o=1}^{O} \lambda_{so} y_{io} \tag{4}$$

where the summation index runs through O possible occupations, and y_{io} is the amount of time that individual i spent exercising occupation o throughout her career. As is clear from equation (4), we center skill-specific experience around the occupations in a worker's career and not around specific jobs. Furthermore, the skill weights λ in equation (4) are the same skill weights that we use in equation (2). Thus, we assume that the skill weights λ indicate how important a skill is for the productivity of the worker in occupation o, and at the same time determine the amount of skill-specific experience that the worker is gaining while exercising that occupation. We believe that this assumption is reasonable and it simplifies our analysis.

While both Guvenen *et al.* (2020) and Lise and Postel-Vinay (2020) use test scores to proxy for individual expertise in different skills, we follow an alternative route by using skill-specific experience itself. We describe the relationship between expertise and skill-specific experience as follows:

$$e_{is} = \alpha_s + \beta_s \tilde{y}_{is} + \epsilon_{is} \tag{5}$$

The error term ϵ_{is} has zero mean for each individual i and skill s, and α_s and β_s are the parameters of the linear relationship for each skill s.[8] Combining equations (2) and (5),

[7] According to Cossa, Heckman and Lochner (1999), there are two main specifications of skill formation process in the literature. The first specification (Ben-Porath (1967), Kuruscu (2006), Flinn, Gemici and Laufer (2017), Engbom (2021)) assumes that workers allocate their time between working for pay and training. Human capital accumulation distracts workers from productive activities, resulting in foregone earnings. Such a model corresponds to the notion of vocational or on-the-job training. The second specification (Shaw (1989), Imai and Keane (2004)) is that of learning-by-doing: time devoted to work produces skills. Human capital accumulation happens simultaneously with productive activities, so there is no trade-off between working and training. Our model is in line with the second specification. For a discussion of the differences between the two specifications see Killingsworth (1982) and Blandin (2018).

[8] Section 7 extends this linear specification to a second order polynomial.

we arrive at the following regression equation:

$$w_{io} = \sum_{s=1}^{S} \lambda_{so}(\alpha_s + \beta_s \tilde{y}_{is} + \epsilon_{is}) + u_{io} = \sum_{s=1}^{S} \alpha_s \lambda_{so} + \sum_{s=1}^{S} \beta_s \lambda_{so} \tilde{y}_{is} + \sum_{s=1}^{S} \lambda_{so} \epsilon_{is} + u_{io} \qquad (6)$$

Equation (6) is the basis of the regressions that we will execute in this paper. Skill-weights λ_{so} will be extracted from the O*NET data and used to compute skill-specific experience $\tilde{y}_{is}$. The regression will then produce parameter estimates for α_s and β_s, which can be interpreted in two different ways. On the one hand, if skill weights λ are well measured, statistically significant estimates of α_s and β_s indicate that skill-specific experience $\tilde{y}_{is}$ is a valid proxy of expertise (especially if statistical significance survives the inclusion of covariates). On the other hand, the estimated coefficients β_s have an interpretation as average *learning rates* for skills.[9] Expertise accumulates through relevant work experience but not at the same pace for all skills. In some skills skill-specific experience may help a lot (high β_s), while in others it may advance a worker's expertise very little (low β_s). To the best of our knowledge, this is the first paper to estimate skill-specific learning rates.

Both interpretations of β_s rely on two crucial assumptions. The first one is that the wage perfectly reflects productivity. The second one is that the weights λ perfectly reflect the *market price of expertise* in skills.[10] If expertise in cognitive skills becomes less valuable in the labor market, the decrease in its price should be captured by a decreasing weight λ for cognitive skills, and not by a negative $\beta_{cognitive}$. If this assumption does not hold and if, for example, the weights λ fail to incorporate the fall in the market price of cognitive skills, this could push $\beta_{cognitive}$ to negative territory.

Reliability of regression estimates

Our framework makes very few assumptions about the structure of the labor market, which makes our analysis more flexible and more general. The drawback, however, is that we do not take a specific stand on the various processes that are driving the labor market. For this reason it is important to analyze the possible sources of bias for our regression estimates.

[9] Interpreting the α's is not as straightforward. These parameters correspond to the constant in equation (5), and they should represent the average level of expertise in absence of skill-specific experience. Conceptually, α_s may be very sensitive to sample selection and the inclusion of other control variables, e.g. years of education.

[10] This is arguably problematic, because in practice the O*NET database does not directly record this quantity, but rather captures demand for skill-specific expertise.

From equation (5), the bias of the estimates α_s and β_s for each s can be judged by whether the following terms are equal to zero or not:

$$E[\lambda_{so}(\lambda_{so}\epsilon_{is} + u_{io})] \quad \text{and} \quad E[\lambda_{so}\tilde{y}_{is}(\lambda_{so}\epsilon_{is} + u_{io})] \tag{7}$$

Lemma 1: *The estimates of α_s and β_s from a linear regression based on equation (6) are unbiased if ϵ_{is} and u_{io} are independent of $\tilde{y}_{is}$ and λ_{so}.*

According to Lemma 1, the estimates will be unbiased even if there is correlation between λ_{so} and $\tilde{y}_{is}$. This is an important result, because these two variables are bound to be highly correlated. This is clear from the definition of $\tilde{y}_{is}$, which is the sum of past skill weights λ_{so}.

We now extend the analysis to the case when individuals have different learning rates. Then, equation (5) takes the following form:

$$e_{is} = \alpha_s + (\beta_s + b_{is})\tilde{y}_{is} + \epsilon_{is} \tag{8}$$

and the resulting regression is:

$$w_{io} = \sum_{s=1}^{S} \alpha_s\lambda_{so} + \sum_{s=1}^{S} \beta_s\lambda_{so}\tilde{y}_{is} + \sum_{s=1}^{S} b_{is}\tilde{y}_{is}\lambda_{so} + \sum_{s=1}^{S} \lambda_{so}\epsilon_{is} + u_{io} \tag{9}$$

Here β_s plays the same role as before, but b_{is} is an individual-specific learning rate.

Lemma 2: *The estimates of α_s and β_s from a linear regression based on equation (9) are unbiased if b_{is}, ϵ_{is} and u_{io} are independent of $\tilde{y}_{is}$ and λ_{so}.*

The estimates of α_s and β_s can still be unbiased even if there are individual learning rates in the setup. As before, this result also survives when the skill weights are correlated with skill-specific experience.

Finally, we discuss the situation when the estimates are biased in the following lemma:

Lemma 3: *The estimates of α_s and β_s from a linear regression based on equation (9) are biased if ϵ_{is} or b_{is}:*

1. *are positively correlated with λ_{so} or*

2. *are positively correlated with $\tilde{y}_{is}$. In this latter case the bias on α_s will be negative and the bias on β_s will be positive.*[11]

The following examples provide the intuition behind correlation between b_{is} on the one hand and λ_{so} or $\tilde{y}_{is}$ on the other. A person who is a good learner of cognitive skills may have a higher probability of exercising an occupation that strongly requires cognitive skills ($corr(b_{i,cognitive}, \lambda_{cognitive,o}) > 0$). As a result, the estimate for $\beta_{congitive}$ may be biased. In practice, the positive correlation between individual-specific learning rate and occupation selection is known as *selection on match quality*, and is a fundamental empirical problem of the human capital literature (Sanders and Taber (2012)). Alternatively, someone who is a good learner of cognitive skills may have a higher probability of having worked in the past at jobs that require high cognitive skills ($corr(b_{i,cognitive}, \tilde{y}_{i,cognitive}) > 0$). In this case, the estimate for $\beta_{cognitive}$ will be unambiguously upward biased (while the estimate for $\alpha_{cognitive}$ will be unambiguously downward biased). Although the above scenarios cannot be ruled out, there are two arguments that support the case for unbiased estimates. First, a lot of skills could be largely homogeneous among individuals with respect to learning rates, especially given control variables. Second, Lemmas 1 and 2 show that a correlation between skill weights λ and skill-specific experience $\tilde{y}$ does not lead to biased estimates. Thus, if occupation choice is mainly driven by occupation history of the worker, no bias arises. To evaluate the potential bias of our estimates in practice, we conduct several robustness exercises in Section 7.1.

Empirical specification

Based on equation (6), we estimate a following regression model:

$$\log W_{iot} = \sum_{s=1}^{S} \alpha_s \lambda_{sot} + \sum_{s=1}^{S} \beta_s \lambda_{sot} \tilde{y}_{ist} + \delta X_{it} + \nu_i + \omega_t + \eta_{ind} + \theta_o + \epsilon_{iot} \qquad (10)$$

The dependent variable $\log W_{iot}$ is the log of hourly wage of individual i working in occupation o in year t. The right-hand side of equation (10) consists of initial level of expertise α_s, weighted skill-specific experience $\tilde{y}_{ist}$, a vector of control variables X_{it}, and a set of individual, year, industry, and occupation fixed-effects. We use a balanced panel and apply a fixed-effect estimator with standard errors clustered at the level of individual.

[11] For the results in the lemmas to hold, we need to assume that the regressors are uncorrelated with each other. This would be true if the λ's for the different skills were orthogonal.

The vector of control variables includes age and square of age, marital status, years of education, and tenure in current job. α_s and β_s are $2 \times S$ parameters to estimate.

Equation (10) allows us to recover the values of parameters α_s and β_s and to quantify the average pace of expertise accumulation. We can then plug these values back into equation (5) to see how expertise increases with skill-specific experience. Furthermore, the estimated parameter β_s in combination with skill weights λ_{sot} can be thought of as the returns to skill-specific experience in occupation o.

3 Data

3.1 O*NET

We do not estimate skill weights λ_{sot}. Instead, the skill-weights are extracted from the occupational dataset O*NET. O*NET measures skills, abilities and knowledge necessary to perform tasks within narrow occupation categories, as well as required education, training, interests, values, and experience. The respondents of O*NET questionnaires – either incumbents of the occupations, or occupational analysts – are asked to rate the importance of various skill dimensions for performance within occupations. We collect the analyst ratings from O*NET releases 2002-2017 for the following groups of O*NET occupation characteristics: Basic Skills, Cross-Functional Skills, Abilities, and Knowledge. Each of these groups in turn consists of multiple narrow skill categories (see Appendix Table 8 for a full list).

Given the richness of occupational characteristics available in O*NET, one could wonder how the four groups of characteristics were selected. While Basic Skills and Cross-Functional Skills seem a natural choice for *skills*, it is less clear whether categories within Abilities and Knowledge can also be interpreted as skills. Upon inspecting the available categories across all groups of characteristics provided by O*NET, we decided to include all four groups, as it is often difficult to separate skills from abilities. For instance, Mathematics is a category that belongs to both Skills and Knowledge groups. We follow Lise and Postel-Vinay (2020), who also include characteristics from Abilities and Knowledge into their skill set. Unlike Lise and Postel-Vinay (2020), however, we do not look at categories within Work Context and Work Activities, as those seem to be much more job-specific. The information on the level of skill requirement in O*NET is collected using the following question: "What level of the skill/ability/knowledge is needed to perform your current job?" (see Appendix Figure 8). Of course, our choice of the groups of O*NET occupational characteristics will determine the reduced skills during the dimensionality-reduction procedure. We will discuss the question of what constitutes skill at length in Section 4.

Previous literature that deals with multidimensional skills typically uses the combination of NLSY79 and O*NET data to estimate individuals' skill endowments and occupational skill requirements (Guvenen *et al.* (2020), Lise and Postel-Vinay (2020)). We diverge from the existing research in that we are using the labor market data from a later NLSY97 survey. The earliest available O*NET survey was released only in 2002. Since the importance of various skills in occupations can change over time, the 1997-2011 NLSY survey is better suited for matching with O*NET data.[12]

3.2 NLSY

To estimate equation (10), we use the 1997 sample of the National Longitudinal Survey of Youth (NLSY97). NLSY is a well-known panel dataset, which follows the lives and careers of young people for twenty years. The respondents were aged 12-17 during the first wave of the survey in 1997. Since NLSY97 samples young adults before they enter the labor market, it allows us to estimate their skill-specific experience precisely, as we know all the jobs across all the occupations the respondents ever worked in.

We use the information on the number of jobs per year, total weeks worked per year and total hours worked per year, as well as tenure, hourly wage and weekly hours worked in each job. Most importantly, NLSY97 provides information about the occupation and industry of every job that the respondent held during each year.[13] This

[12]One obvious example of the change in the skill composition of occupations is substitution from routine to non-routine tasks due to computerization (Autor, Levy and Murnane (2003)). Appendix Figure 9 uses O*NET data to illustrate how occupational skill requirements changed between 2002 and 2017 for Production Occupations, using skill group Abilities as an example.

[13]NLSY97 records occupations and industries using 2002 Census classification. For the purpose of merging NLSY97 with O*NET data, we transformed 2002 Census coding into 2018 SOC classification, using crosswalks provided by US Census Bureau. There are 4 occupations in 2002 Census that don't exist in 2018 SOC. Since we cannot connect these occupations with skill weights λ_{sot}, we drop the respondents who worked in these occupations from our baseline sample.

allows us to accurately measure experience acquired by the respondent even in cases of several parallel jobs. Since NLSY97 collected data only biennially from 2011 onward, we restrict our sample to 1997-2011, which gives us 15 years of observations. We exclude individuals who reported tenure above 53 weeks in 1997, since we have no information about their jobs and occupations prior to 1997. We also drop individuals with gaps in employment history due to non-interview.[14] We do not select our sample based on age: even though some of the respondents are only 12 in 1997, their skill-specific experience is measured correctly, and we still observe 5-10 years of their labor market experience.[15] Nevertheless, it is important to remember that our analysis is based on a young sample of individuals at the beginning of their career.

The most important variable of our analysis is skill-specific experience $\tilde{y}_{ist}$, which measures the time that a particular skill s was practiced. We construct $\tilde{y}_{ist}$ based on NLSY97 data and skill weights λ_{sot} from O*NET. First, we compute hours per year worked in each occupation o for each individual in our baseline sample. We weight these annual hours in occupation o by skill weights λ_{sot}. This transforms occupation-specific hours into skill-specific hours, resulting in $S \times O$ variables measuring hours worked. We proceed by summing up skill-specific hours across all occupations. For each year and for each individual, we have now constructed a measure of annual hours spent practicing skill s. Finally, we sum up annual skill-specific hours across years to get a measure of skill-specific experience $\tilde{y}_{ist}$.

Table 1: *Summary statistics*

	All		Women		Men	
	mean	sd	mean	sd	mean	sd
Age	21.59	4.56	21.63	4.56	21.56	4.56
Male	0.48					
Ever married dummy	0.17		0.20		0.13	
Years of education	11.65	2.68	11.82	2.76	11.46	2.58
Tenure in current job	1.52	1.82	1.48	1.76	1.57	1.89
Wage, 2015 dollars	14.17	9.30	13.47	8.62	14.94	9.94
Total weeks worked per year	28.24	22.33	28.21	22.23	28.28	22.44
Years in labor market (experience)	5.19	4.09	5.21	4.10	5.18	4.07
Observations	83955		44055		39900	

[14]For a more detailed account of imputations, variable construction and cleaning procedures, see Appendix A1.

[15]We conduct robustness checks, in which we exclude those below 14 years of age. Our results remain unchanged.

Our baseline sample is balanced: it includes 5,597 young men and women. Table 1 presents the summary statistics for the selected control variables. The baseline sample is young – the average age in the sample is 21.6 years. Young people are more likely to switch jobs – the average tenure in the sample is only 1.5 years. Since many of the respondents are still studying at the beginning of our sample period, both the average annual number of weeks worked and the average labor market experience are pretty low at 28.2 weeks and 5.2 years respectively.

The main goal of this paper is to measure the average learning rates for different skills. Naturally, the estimates of the learning rates will be driven in part by the characteristics of our sample. Given the average age and labor market experience of the respondents in NLSY97, the values of the parameters β can be treated as average learning rates for the workers at the beginning of their career, and should be generalized to the entire population with caution. The importance of the sample characteristics will become clear in the heterogeneity analysis in Section 6.

4 Constructing skill weights

4.1 What is a skill?

Before continuing, it is useful to briefly discuss what we mean by skill in this paper. The existing literature on skill-based human capital does not provide an unambiguous definition of skill. A common approach is to introduce a dichotomous skill heterogeneity, such as cognitive versus non-cognitive (manual, physical, motor) skills (see, for instance, Yamaguchi (2012)). Other classifications include social or interpersonal skills in addition to cognitive skills (Lise and Postel-Vinay (2020)). Some authors introduce a slightly finer heterogeneity within cognitive or physical skills. Guvenen *et al.* (2020), for instance, distinguish between math and verbal skills. Poletaev and Robinson (2008) split non-cognitive skills into fine motor skills and physical strength, while Ingram and Neumann (2006) separate coordination in addition to fine motor skills and strength. Some non-standard classifications also exist (e.g. Sanders (2016) divides skills into known and unknown to the worker, and Sanders and Taber (2012) suggest the Big Five personality traits as measures of non-cognitive skills). Psychological literature suggests yet another skill classification. For instance, according to Cattell (1963), general ability can be divided into fluid (adaptive) and crystallized (narrow) abilities.

Conceptually, in this paper we take equation (1) in Lazear (2009) seriously in that a skill is defined as any kind of capacity that workers may have that can contribute to their productivity. Furthermore, it is implied by the framework that workers sharing this capacity will see a similar effect on their productivity if they are working in occupations with similar skill requirements. This definition is quite broad, as it could include categories as different as finger dexterity, explosive strength, chemistry, and persuasion.[16] We decided to maintain the term "skills" for lack of a better term to describe these categories, even if it seems as too broad in many cases. In practice, the choice of relevant skill dimensions will arise "endogenously" as an outcome of our dimensionality reduction procedure.

4.2 Dimensionality reduction methods

To construct a measure of skill-specific experience $\tilde{y}$ and perform our regression analysis, we need data on skill requirements – skill weights λ – for each occupation. To that end, we use the 120 skill categories that describe each occupation in the O*NET database, as was described earlier.[17] Clearly, 120 are too many categories for the exercises that we want to perform. Thus, we construct a reduced space of skill weights. Ideally, the dimensionality reduction procedure extracts the true underlying skill weights out of noisy observable O*NET data.

To achieve this, we use Nonnegative Matrix Factorisation (NMF) (Lee and Seung (1999), Févotte and Idier (2011)). This method factorizes the full matrix of skill categories that we get from O*NET into two matrices. The key feature of these matrices is that they have no negative elements. For each of the 120 skill requirements that are recorded in the O*NET database and for each occupation, this underlying process can be described as follows:

$$\lambda'_{s'o} = x_{s',1}\lambda_{1,o} + x_{s',2}\lambda_{2,o} + \ldots + x_{s',s}\lambda_{s,o} \tag{11}$$

where λ's represent the reduced skill requirements, while λ's represent the skill require-

[16] All these examples are also particular categories that are available in the O*NET database.

[17] Specifically, we use *levels* of skills from O*NET.

ments recorded in the O*NET database, and the indices $s = 1, \ldots, S$, $s' = 1, \ldots, 120$ and o respectively indicate reduced skills, the full list of skill requirements and occupations. These equations can be written in matrix form:

$$\underset{(O\times S')}{\mathbf{\Lambda}'} = \underset{(O\times S)}{\mathbf{\Lambda}_S}\underset{(S\times S')}{\mathbf{X}_S} \tag{12}$$

We can now use matrix $\mathbf{\Lambda}_S$ as the reduced space of the skill weights, as it has the desirable dimensions. In particular, each row of this matrix corresponds to an occupation in our sample, and the number of columns corresponds to the number of dimensions that we want our reduced space of skill requirements to have. The main benefit of using NMF for dimensionality reduction is that all the numbers involved are nonnegative. This is a desirable characteristic, because it is difficult to make sense of negative skill requirements or negative combinations of skill requirements. NMF guarantees this result, whereas alternative dimensionality reduction methods such as principal component analysis (PCA) can produce any combination of positive and negative numbers.

NMF produces the matrix of reduced skill weights which we use for our analysis.[18] However, the method by itself does not determine the number of reduced skills. To decide what is the optimal number of skills, we use diagnostics such as the elbow method and the silhouette method. In our case, these methods do not provide an unambiguous answer.[19] We choose four reduced skills, as we believe that this strikes the right balance between obtaining as much explanatory power as possible and maintaining statistical power and interpretability.

To provide informative labels on these reduced skills, we inspect the correlations between NMF components and the skill categories of the original data. Specifically, for each reduced skill, represented by a column of matrix $\mathbf{\Lambda}_S$, we compute the correlation with each of the 120 skill categories in the raw O*NET data.[20]. Table 2 shows the top

[18]Following Lise and Postel-Vinay (2020), we linearly transform the data on the reduced skill requirements so that they fit in the [0,1] interval. Unlike Lise and Postel-Vinay (2020), we do not require our skill weights to sum up to one. This may seem puzzling to the reader as this is a common practice when working with weights. However, in this case, the skill weights do not only reflect the composition of skill requirements of occupations but also the skill level. This crucial dimension would be lost if we forced the weights to always add up to one.

[19]The elbow method (Appendix Figure 10) suggests three reduced skills, but the "elbow" is not that prominent, while the silhouette method (Appendix Figure 11) seems to suggest that anything between two and five reduced skills would be reasonable.

[20]The reduced skills are represented by columns in $\mathbf{\Lambda}_S$, while the skill requirements for each of the 120 categories in the O*NET data are represented by columns in $\mathbf{\Lambda}'$.

five highest correlations for each of the reduced skills.

Table 2: *Correlations between NMF components (reduced skills) and skill categories from O*NET*

	Skill category from O*NET	Correlation
Component 1	Speech Recognition	.778
	Speech Clarity	.768
	English Language	.76
	Psychology	.729
	Sociology and Anthropology	.724
Component 2	Multilimb Coordination	.928
	Static Strength	.918
	Extent Flexibility	.909
	Reaction Time	.896
	Dynamic Strength	.892
Component 3	Equipment Selection	.89
	Technology Design	.838
	Installation	.834
	Troubleshooting	.78
	Equipment Maintenance	.714
Component 4	Engineering and Technology	.91
	Design	.847
	Physics	.81
	Mathematics	.713
	Visualization	.599

The correlation analysis allows us to give intuitive and accurate labels to our reduced skills. The first NMF component – the first reduced skill – is highly correlated with such skill categories as speech recognition, speech clarity, English language, psychology, and sociology. This indicates that the first reduced skill summarizes the social requirements of occupations. The second component is clearly connected to physical skills, as it exhibits close to one correlations with multilimb coordination, strength, flexibility, and reaction time. The third component summarizes information on technical requirements of occupations, such as equipment selection and maintenance, installation, and technology design. Finally, the last NMF component seems to represent creative or cognitive skills, as it shows high correlation with such skill categories as engineering, design, physics, mathematics, and visualization. Based on the correlation analysis, we label the four reduced skills as social, physical, technical, and cognitive.

In summary, NMF is a good dimensionality reduction method for the case of skills. It results in a matrix of nonnegative reduced skills, which is sensible since we naturally think of skills as positive attributes.[21] NMF is used extensively in machine learning applications (for example, in facial recognition), but to our knowledge it has not been used in economics. The final skills – social, physical, technical, and cognitive – are very intuitive, and describe the skill composition of the labor market in a reasonably comprehensive way. They are also comparable with the skills used in the literature. For example, Lise and Postel-Vinay (2020) use three categories of skills: manual, social and cognitive, with manual skills being similar to what we call technical skills. Thus, our dimensionality reduction procedure introduces a somewhat finer heterogeneity in skills, additionally separating physical skills. Of course, the main benefit of our methodology relative to the existing works is that we abstract from making any subjective decisions regarding the choice of skills. The dimensionality reduction methods most commonly used in the O*NET-based literature involve either manual categorization of O*NET characteristics into broader skill groups, or a subjective selection of exclusion restrictions. We simply let the data speak.

Before turning to the estimation of our model, we also check the sensibility of the skill weights λ obtained as a result of NMF by looking at occupation categories. Table 3 shows skill weights λ_{so} in 2011 averaged over broad occupation groups. Indicatively, social, physical, technical, and cognitive skills are required at the highest level by the broad occupation categories Community and Social Service, Construction and Extraction, Computer and Mathematical, and Architecture and Engineering occupations respectively. This is an extra verification that the reduced skills make sense, as the labels were not chosen in association with specific occupations.

Table 3: *Average estimated λ_{so} in 2011 for broad occupation groups.*

Occupation Group	Social	Physical	Technical	Cognitive
Architecture and Engineering	.28618902	.07811863	.23118561	.59912735
Arts, Design, Entertainment, Sports, and Media	.47912365	.13588654	.04198404	.12416819
Building and Grounds Cleaning and Maintenance	.28735389	.35352162	.06457376	.10297743
Business and Financial Operations	.52851543	.02915419	.04442388	.19987429
Community and Social Service	.63872004	.1279843	.00581	.01962201
Computer and Mathematical	.36002418	.02740391	.29165203	.34281122
Construction and Extraction	.19010702	.47489226	.11151757	.21800971
Educational Instruction and Library	.5366491	.09986561	.02359346	.13630168
Farming, Fishing, and Forestry	.2634566	.39553864	.05917655	.1867835
Food Preparation and Serving Related	.33359315	.28807019	.02936215	.04210635
Healthcare Practitioners and Technical	.53958425	.24031693	.03752594	.15054376

[21] Additionally, the matrix of loadings $\mathbf{X}_S$, which describes the way the reduced skills must be combined to recreate the original data, is also nonnegative.

Healthcare Support	.43632113	.28056554	.01671212	.02290295
Installation, Maintenance, and Repair	.1502522	.4426992	.25614258	.24795805
Legal	.59834442	.00826091	.007146	.06432053
Life, Physical, and Social Science	.4361368	.1167108	.10648435	.34711879
Management	.54138941	.08102436	.06109446	.25031494
Office and Administrative Support	.44784238	.11312516	.03105842	.0687637
Personal Care and Service	.42639604	.22663195	.01182458	.03886119
Production	.18667029	.40228343	.16080814	.1718724
Protective Service	.43972485	.36396973	.01049619	.09991131
Sales and Related	.47705116	.0891592	.0341962	.11992848
Transportation and Material Moving	.27406271	.43522831	.06348333	.13892121
Total Average	.40261399	.21865506	.07728415	.16787268

5 Main results

5.1 Skill-specific experience

Before proceeding to the regression results, we present our empirical proxy for unobserved expertise in different skills – skill-specific experience. Using the skill weights λ_{sot}, we follow equation (4) to construct experience $\tilde{y}$ in 4 skills: social, physical, technical, and cognitive.

Figure 1 demonstrates the average skill-specific experience by years in the labor market. Individuals at the beginning of their career accumulate the most experience in social skills. After 15 years of working, the average experience in social skills exceeds 3 units. Experience in physical skills and technical skills comes as a close second, with $\tilde{y}$ around 2. The accumulation of skill-specific experience in social skills accelerates at a later stage of an individual's career, suggesting that young people switch from physical- to social-intensive occupations over time. NLSY97 respondents accumulated the least experience (less than 1) in cognitive skills.

The importance of physical skills at the beginning of workers' career is not surprising: the first job for many young people is often a low-paid entry-level position, which requires no special qualifications. Besides, young people are likely to have a physically-intensive side job while enrolled in college. In fact, the five most common occupations in our young sample are cashiers, retail salespersons, waiters, laborers and material movers, and cooks – all of which are characterized by low skill weights λ in all but physical skills.

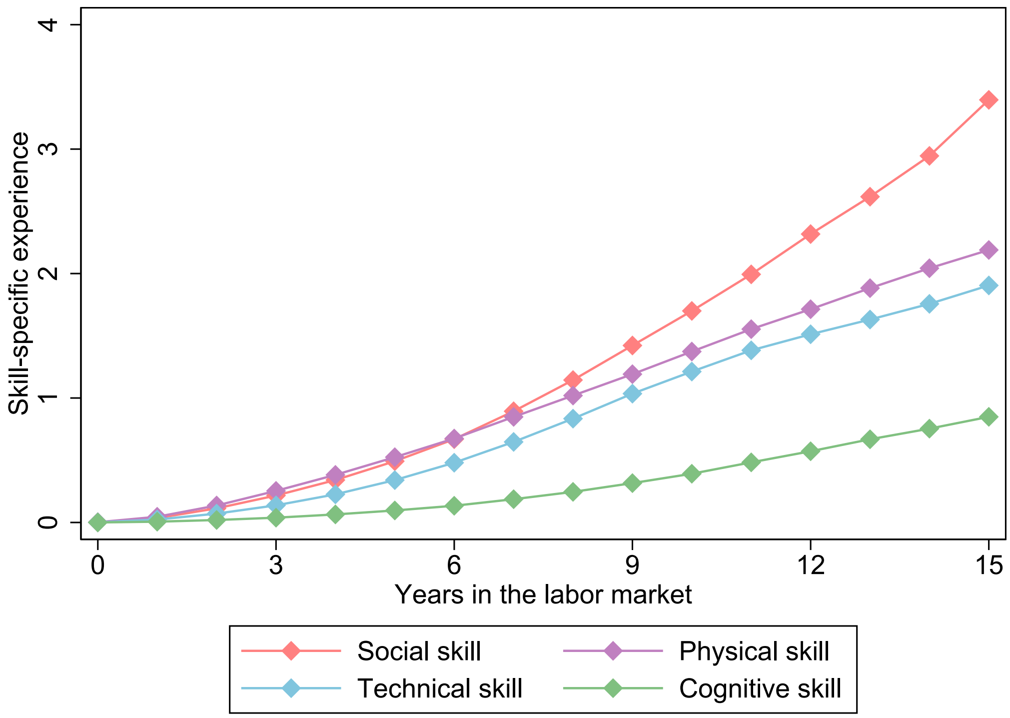


Figure 1: *Average skill-specific experience $\tilde{y}_s$ by years in the labor market.*

As the sample becomes older, individuals switch to occupations which require higher level of social skills. After the first 6 years in the labor market, the paths of experience accumulation in physical and social skills diverge.

Figure 2 highlights important heterogeneities in accumulation of skill-specific experience by gender. The first fact that stands out is that young women specialize in social skills. After 15 years in the labor market, women gain 4 units of experience in social skills, compared to only around 3 units gained by men. Moreover, women start accumulating experience in social skills early on. Men, on the other hand, do not exhibit high concentration in one skill. Over the first 15 years in the labor market, young men reach 3 units of experience in physical and social skills (with physical skills dominating throughout the entire period), and 2 units of experience in technical skills. Men also accumulate slightly more experience in cognitive skills. Since on average young women spend as much time in the labor market as young men (both in terms of years of labor market experience and weeks worked, see Table 1), the difference in experience must be driven by occupation choice and corresponding weights λ. Women concentrate in occupations with higher social skill requirements, and lower requirements for physical, technical, and cognitive skills (see also Appendix Table 9).

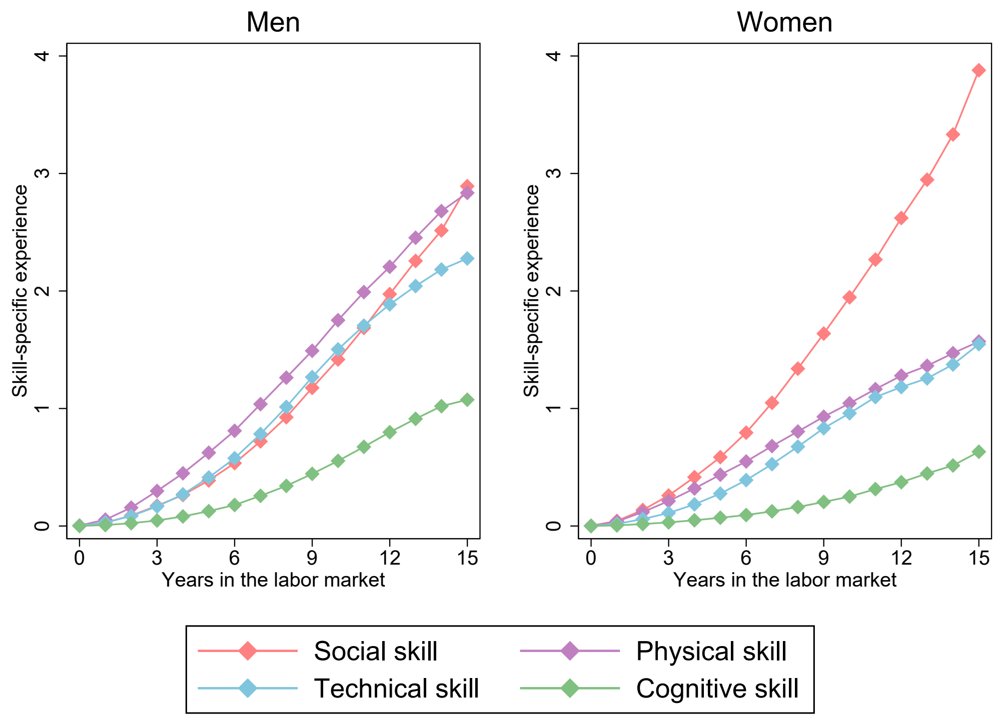


Figure 2: *Average skill-specific experience $\tilde{y}_s$ by years in the labor market for men and women.*

Finally, Figure 3 shows the distribution of skill-specific experience across education groups. The patterns of skill-specific experience by college education mimic those by gender. Perhaps not surprisingly, individuals with no college education achieve higher level of skill-specific experience in physical skills. This result is in line with Yamaguchi (2012), who found that less educated individuals work in occupations with higher motor task requirements. College educated respondents, on the other hand, have significantly more experience in social skills at any stage of their career. Similarly to women, college educated workers seem to specialize in social skills. A less anticipated result is that those without a college education accumulate more experience in technical skills. This result is mainly driven by the differences in working time between college and non-college respondents. Although workers with and without college education are equally likely to select technically-intensive occupations (Appendix Table 9), those with a college degree on average work less hours per week at the beginning of their career, as they probably have a part-time job alongside studying (Appendix Figure 13). Finally, both education groups gain a similar level of experience in cognitive skills. As with technical skills, this result is primarily explained by the differences in working hours, since college-educated

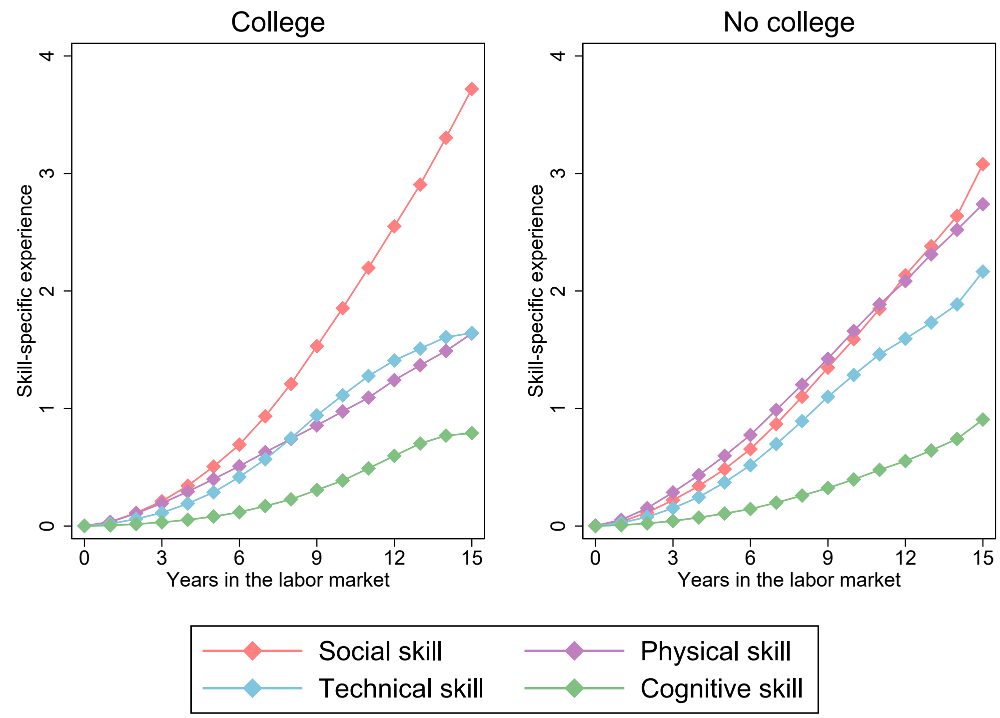


Figure 3: *Average skill-specific experience $\tilde{y}_s$ by years in the labor market. Left panel shows average skill-specific experience for people with some college (Associates, Bachelors, Masters or PhD degree), right panel shows average skill-specific experience for people without college (GED, high-school diploma or no degree at all).*

individuals work in more cognitive-intensive occupations.

Overall, gender heterogeneity is driven by occupational choices, while the differences in skill-specific experience between education groups follow from both occupational selection (in physical and social skills) and hours gap (technical and cognitive skills). The heterogeneity analysis along the education dimension generates another interesting result. It indicates that a simple skilled vs not skilled dichotomy of workers may be inaccurate. Both college and non-college graduates are on average *skilled*. The only difference is the composition of skills that these groups exhibit. Furthermore, we find that college-educated workers specialize in social rather than cognitive skills, and have the same experience in cognitive skills compared to people with less education. This result could have important implications for explaining college wage premium.

5.2 Getting expertise with experience

The main empirical question of our research is how people accumulate expertise in different skills with work experience. To answer this question, we use skill weights λ and skill-specific experience $\tilde{y}$ to estimate equation (10). As discussed earlier, the estimated coefficients β can then be interpreted as the speed of expertise accumulation.

The estimation results are presented in Table 4. Column (1) includes only our measures of skill weights and skill-specific experience.[22] The first five estimates ("weight") correspond to parameters α, and the remaining five estimates ("weighted experience") correspond to parameters β from equation (5). Regressions in columns (2)-(4) additionally control for age and age squared, marital status, years of education, and tenure, as well as for broad occupation and industry groups in column (3), and for narrow occupations and broad industries in column (4).

In the interpretation of the results, we focus on the estimates for the learning rates β as opposed to the estimates for the parameters α on two grounds. First, the parameters α are bound to vary significantly depending on the choice of the sample, as they correspond to the average expertise at the start of workers' career. Second, they are very sensitive to the inclusion of control variables like education. For this reason, we think that the most reliable specification for the estimation of the α's is found in column (1) of Table 4. With respect to the β's, our preferred specification is in column (4), because it shows that our estimates of the learning rates survive both the inclusion of covariates and the use of narrow occupation and industry fixed effects.

According to our theoretical model (equation (2)), the wage of an individual is a weighted sum of her expertise in different skills, with weights λ measuring the importance of skills in a given occupation. In other words, the returns to expertise in different skills are captured by the parameters λ, and are not estimated in this regression. In turn, the estimated parameters α and β describe the relationship between expertise and skill-specific experience. Positive β_s implies that expertise in skill s is increasing with skill-specific experience. The magnitude of the coefficient β_s is interpreted as the learning speed: the higher the coefficient, the quicker people accumulate expertise with work experience.

The estimation results in column (1) indicate that expertise is indeed increasing with skill-specific experience: β coefficients of experience in all skills are positive and

[22] All regressions include individual fixed effects; regressions in columns (2)-(4) also include year fixed effects.

***Table 4:** Estimating learning speed β*

	(1)	(2)	(3)	(4)
Social sk weight	0.135***	-0.109***	0.0617*	-0.185***
	(0.0246)	(0.0243)	(0.0323)	(0.0471)
Physical sk weight	0.00450	0.0176	-0.0182	-0.167***
	(0.0282)	(0.0277)	(0.0294)	(0.0484)
Technical sk weight	0.279***	0.106***	-0.0290	-0.191***
	(0.0215)	(0.0220)	(0.0244)	(0.0301)
Cognitive sk weight	0.449***	0.335***	0.126**	0.0774
	(0.0387)	(0.0372)	(0.0498)	(0.0740)
Weighted experience in Social sk	0.299***	0.193***	0.155***	0.112***
	(0.00670)	(0.00983)	(0.00971)	(0.00976)
Weighted experience in Physical sk	0.177***	0.0681***	0.0340**	0.0301**
	(0.0109)	(0.0140)	(0.0138)	(0.0138)
Weighted experience in Technical sk	0.186***	0.122***	0.143***	0.155***
	(0.0148)	(0.0143)	(0.0146)	(0.0142)
Weighted experience in Cognitive sk	0.237***	0.247***	0.239***	0.220***
	(0.0250)	(0.0257)	(0.0261)	(0.0278)
Age		0.0776***	0.0691***	0.0552***
		(0.0135)	(0.0130)	(0.0124)
Age squared		-0.00167***	-0.00148***	-0.00123***
		(0.000255)	(0.000245)	(0.000233)
Ever married dummy		0.0453***	0.0423***	0.0361***
		(0.00851)	(0.00815)	(0.00765)
Years of education		0.0386***	0.0332***	0.0296***
		(0.00257)	(0.00248)	(0.00237)
Tenure in current job		0.00997***	0.0140***	0.0159***
		(0.00170)	(0.00167)	(0.00161)
Constant	2.192***	0.801***	0.922***	1.553***
	(0.0126)	(0.150)	(0.155)	(0.214)
Broad occupations	No	No	Yes	No
Narrow occupations	No	No	No	Yes
Industry FE	No	No	Yes	Yes
Observations	53682	52724	52657	52657
Adjusted R^2	0.270	0.320	0.349	0.401

Clustered standard errors in parentheses. FE estimator. All regressions include year fixed effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

statistically significant at the 1% level. Every additional unit of skill-specific experience leads to accumulation of human capital. The coefficients of skill-weights – parameters α from equation (5) – are also positive and significant (apart from physical skills), suggesting that individuals enter the labor market with some positive expertise.

Gradually adding control variables does not affect the significance of the β coefficients, which means that our measures of skill-specific experience have high explanatory power even after taking into account workers' education, tenure, industry, and occupation. The R^2 in column (1) is also relatively high, suggesting that skill-specific experience alone explains more than a fourth of the variation in wages.

The magnitudes of the coefficients differ across skills. According to the results in column (4), experience in cognitive skills is associated with the quickest increase of expertise. Expertise in technical and social skills is acquired somewhat slower: it will take twice as much time to reach the same level of expertise in social skills compared to cognitive skills for the same level of skill-specific experience. Experience in physical skills is somewhat of an outlier. The β coefficient of physical skills is rather small in magnitude, suggesting that workers accumulate little expertise in physical skills with experience. This result is important for two reasons. First, it shows that our model, combined with the applied dimensionality-reduction procedure, distinguishes between different types of skills. Our measures of skill-specific experience do not simply proxy for the overall labor market experience of the individual – on the contrary, they contain useful information about the accumulated level of skill.[23] Second, it tells us something about the nature of physical skills. We will return to this last point later.

We can now use the estimated coefficients β from column (4) of Table 4 and our empirical measure of skill-specific experience to compute unobserved expertise e_{is} in different skills. Specifically, we calculate expertise e_{is} for every individual in our sample, thus taking into account her employment history. We then average expertise across individuals with the same total labor market experience.[24] Figure 4 plots average expertise in our sample over the years in the labor market.

Overall, the ranking of skills in terms of accumulated expertise changed compared to skill-specific experience. According to Figure 4, individuals in our sample accumulate the most expertise in social and technical skills. Cognitive skills, on the other

[23] We further confirm this conclusion by including the total years of labor market experience as an additional control variable. Our estimates of β remain unchanged. Results are available upon request.

[24] We normalize α to zero for all skills, assuming that everybody enters the labor market with the same level of initial expertise.

hand, are relatively slow-growing: in 15 years of labor market experience NLSY97 respondents reached only half as much expertise in cognitive as in social skills. Finally, expertise in physical skills is rather flat, increasing only slightly with experience.

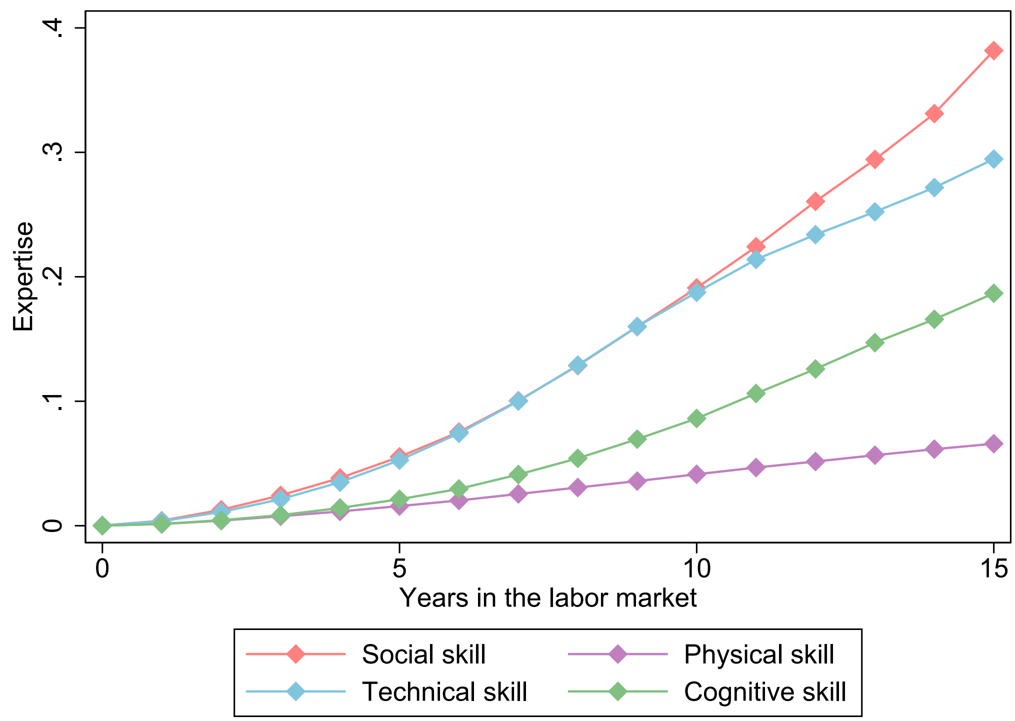


Figure 4: *Average expertise $\bar{e}_s$ against total years of labor market experience y.*

Expertise is determined by skill-specific experience $\tilde{y}$ (which in turn depends on occupational choices and skill-weights λ) and the learning rate β. Low beta coefficient means that the speed of learning of the skill is low, and it takes a lot of experience to accumulate expertise. High learning speed means that even low levels of experience translate into relatively high expertise. Skill-specific experience is determined by the choices the individuals make about their careers – which occupations they choose, when and how much they work, whereas expertise is defined by the nature of the skill itself. Comparing our estimates of expertise in Figure 4 with skill-specific experience presented in Figure 1, we see that high level of experience in social skills accumulated in our sample translates into high expertise in social skills. On average, workers practice social skills a lot ($\tilde{y}$ is high). This, combined with a relatively high learning rate in social skills (β) results in high expertise.

At the same time, we see some striking differences between the figure for skill-specific experience and the figure for expertise. Although individuals in our sample get a similar level of experience in physical and technical skills, they do not transform this experience into expertise in the corresponding skills in the same way. Physical skill was the second most important skill in terms of experience – as discussed previously, young individuals often work in physically-intensive occupations. All this experience, however, does not contribute to expertise accumulation. Since the learning rate in physical skills is very low, young workers do not gain additional human capital by practicing physical tasks, and the expertise profile in physical skills is flat. On the contrary, high learning rate in technical skills suggests that workers transform their experience into expertise quicker. During the first decade in the labor market, individuals in our sample accumulate equally high expertise in social and technical skills, despite the fact that experience in technical skills is always below experience in social skills. High learning rates also compensate for low accumulated experience in cognitive skills.

Expertise in physical skills diverges strongly from experience. Every additional unit of labor market experience in physical skills leads to only a minor increase in expertise. In other words, practicing physical skills does not increase human capital. This result seems intuitive, as one could expect that there is less to learn in physical skills compared to, for example, cognitive skills. Another potential reason for low learning rates in physical skills is that physical labor leads to deterioration of a worker's health. Moreover, workers with a lot of experience in physical skills are less likely to find a non-manual job. Although individuals in our sample accumulated a lot of experience in physical skills, they did not gain expertise. Low expertise translates into lower wages, since wage is a weighted sum of expertise in different skills. Hence, workers who work in occupations that demand physical skills are not rewarded significantly more with the passage of time. More generally, the heterogeneity in speed of expertise accumulation has important implications for wages: all else being equal, high speed of expertise accumulation translates into higher wage growth.

6 Heterogeneities in expertise

6.1 Gender

Our analysis of the estimation results in the previous section highlights some important heterogeneities between skills, both in terms of level (measured by skill-specific experience) and speed of accumulation. We find that individuals accumulate the most experience in social skills, but cognitive skills are characterized by the highest learning speed. Expertise in physical skills exhibits a flat profile – experience in physical skills does not increase human capital. We now turn to the analysis of heterogeneities in skill accumulation between individuals across two dimensions: gender and education.

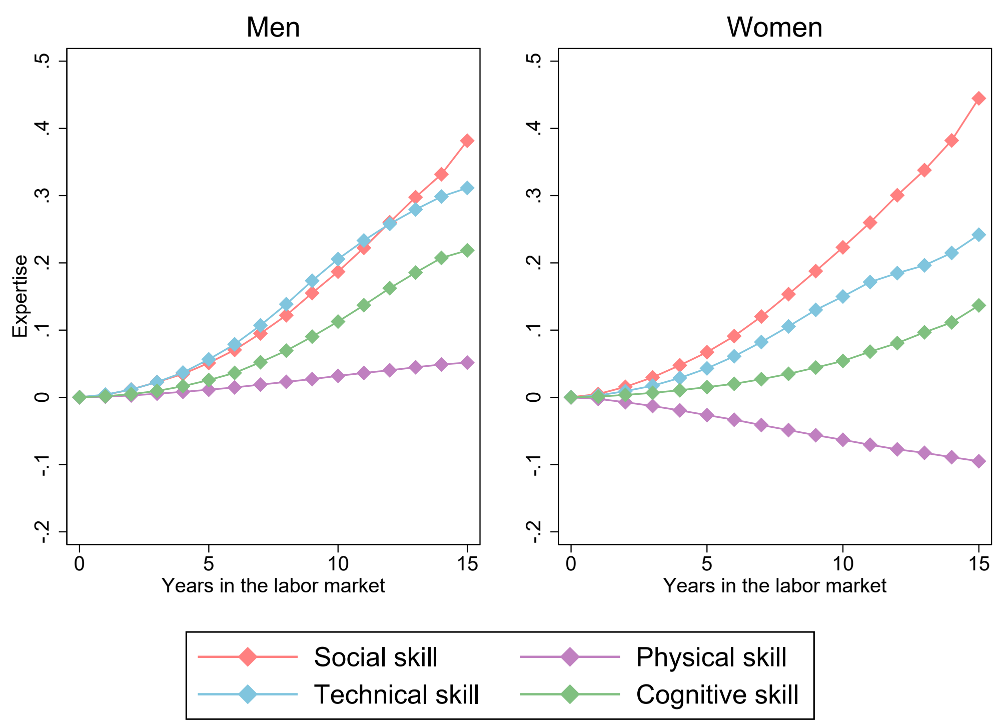


Figure 5: *Average expertise $\bar{e}_s$ against total years of labor market experience y.*

Figure 5 plots expertise in skills by years in the labor market for men and women. The expertise profiles of men and women are qualitatively similar – the ranking of skills in terms of expertise is the same for both groups. Over the first 15 years of labor market work, young men and women accumulate the most expertise in social skills, and the least expertise in physical skills. However, expertise is more dispersed for women. Men seem to gain more expertise in technical and cognitive skills, while women gain more expertise in social skills. Furthermore, women may in fact be losing expertise in physical skills with experience.

Comparing Figure 5 with Figure 2, we see a similar pattern in skill-specific experience: men are more experienced in technical and cognitive skills, and less experienced in social skills than women. This similarity indicates that there are no significant differences in the speed of skill accumulation between men and women – they gain expertise at the same rate with labor market experience. The beta coefficients of the regression confirm this conclusion.[25] The only skill that has a significantly different β for men and women is physical skill: the learning rate of physical skills for women is negative (although the coefficient is not statistically significant in both regressions), and the expertise profile is declining. Practicing physical skills depletes human capital for women.

The analysis of the expertise profiles of men and women allows us to reach an important conclusion. The differences in expertise accumulation between men and women are predominantly the consequence of occupational choices: women accumulate less expertise in technical and cognitive skills because they concentrate in occupations with lower requirement of these skills.[26] According to Appendix Table 9, the average technical (cognitive) skill weight for women is 0.15 (0.05), compared to 0.21 (0.09) for men. However, the speed with which men and women accumulate expertise is the same. Physical skill is an outlier: although women work in occupations that require physical skills (experience in physical skills is on average similar to experience in technical skills), this experience does not contribute to their expertise, and might even have a negative impact on their wage.

6.2 Education

Finally, we analyze the patters of expertise accumulation across education groups. We split our sample into two groups based on education level, with high education group

[25] If anything, women learn cognitive and technical skills somewhat quicker than men, although the difference in learning rates is not statistically significant. Appendix Table 10 shows the regression results.

[26] Our conclusions do not say anything regarding gender discrimination in the labor market. In this section, we document the differences in the realized skill-specific experience between men and women in our sample. We do not attempt to identify the reason that men and women make different choices, nor to provide evidence on the existence of discrimination by employers.

comprising of individuals with some college education. The expertise profiles by years in the labor market are presented in Figure 6 (see Appendix Table 11 for regression results).

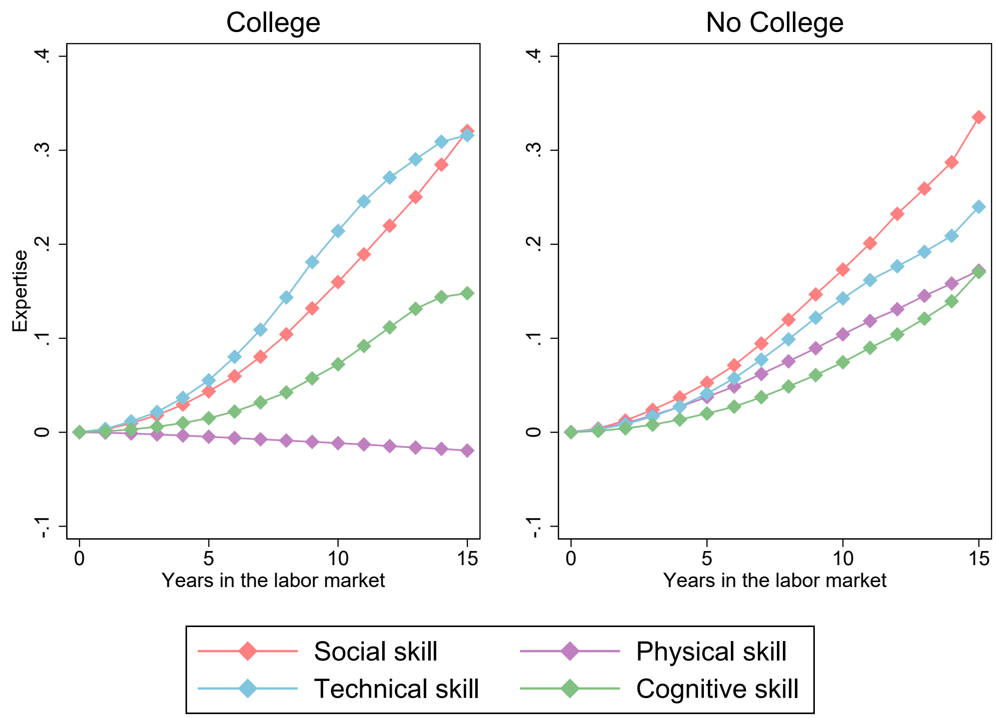


Figure 6: *Average expertise $\bar{e}_s$ against total years of labor market experience y.*

At first glance, the results presented in Figure 6 are somewhat counterintuitive.[27] The less educated respondents attain slightly higher level of expertise in social and cognitive skills, and a significantly higher level of expertise in physical skills. College-educated individuals have an advantage only in technical skills. The distribution of expertise for highly educated respondents is more dispersed.

Comparing Figure 6 with Figure 3, we see that the expertise profiles are quite distinct from the patterns of skill-specific experience. This means that, in contrast to the gender results, there are crucial differences in the speed of expertise accumulation between individuals with different educational background. According to the regression estimation in Appendix Table 11, college-educated individuals gain expertise in social skills slightly slower than the non-college educated. Thus, although individuals with more education have more experience in this skill (and on average self-select into occupations with higher

[27] We cannot of course exclude the possibility that some of these estimates are biased according to the processes described in Lemma 3.

social skill requirement), less of this experience is transformed into expertise. On the other hand, the speed of expertise accumulation in technical skills is significantly higher for college graduates, which results in a high level of expertise despite lower skill-specific experience. As before, physical skills represent a special case: the beta coefficient is slightly negative for college graduates and significantly positive for the less educated. This means that experience in physical skills may lead to depletion of human capital for the college educated, while those with less years of education gain expertise (albeit slowly) while practicing physical skills.

One important driver of our results is the fact that people with college education put in less hours of work in the beginning of their working career (see Appendix Figure 13). An individual with a high-school degree is more likely to start working full-time and have a strong attachment as soon as she enters the labor market, while a college student is more likely to only work part-time at the beginning, and have a lot of labor market interruptions during her studies. If this is the case, a less educated individual with 5 years of labor market experience is not directly comparable with a college graduate with 5 years of experience, as the latter is not as attached to the labor market at the beginning.

Finally, the regression results presented in Appendix Table 11 lead to several interesting observations. First, the R^2 of the regression for the highly educated is much higher than for the less educated. It seems that our measures of skill explain the variation in wages of the college educated much better. This suggests a possibility of other dimensions of skills that are more important for the less educated workers but are not accounted for in the current model. Second, the analysis of skill-specific experience and the speed of expertise accumulation across education groups points us to the conjecture that "highly educated" does not necessarily mean "highly skilled". As mentioned earlier, education does not determine expertise that a person can accumulate. Nevertheless, the heterogeneity analysis of expertise along education dimension presents some puzzling results that merit further exploration.

7 Additional results

7.1 Robustness checks

The results of our analysis, summarized in Table 4, suggest that expertise grows with skill-specific experience, and that the learning rate is heterogeneous across the four skills. However, as highlighted in Section 2.1, the reliability and the interpretation of our estimates depend on two important assumptions: (1) weights λ capture market price of skills and (2) the correlation between individual-specific learning rate and λ or $\tilde{y}$ is zero. Here, we explore the robustness of our results in cases when the above assumptions do not hold.

The first column of Table 5 replicates our baseline results with broad occupation groups as control variables.[28] Both the magnitude and the significance of the coefficients β_s are similar to the ones reported as our main results.

Column (2) of Table 5 includes Occupation×Year fixed effects. This specification relaxes assumption (1). If occupation weights λ do not fully reflect the market price of skills, then using Occupation×Year fixed effects can help to control for the relative price changes across occupations. The results remain unchanged.

In column (3), we follow Gathmann and Schönberg (2010) and replicate our analysis with bootstrapped standard errors to correct for the generated regressor bias. This procedure leaves our baseline estimates unchanged.

Finally, in column (4) of Table 5 we attempt to relax the assumption (2) described above. Specifically, we try to take care of the correlation between individual-specific learning rates and occupation selection (Lemma 3). We implement an adaptation of a well-known instrumental variable (IV) estimator first suggested by Altonji and Shakotko (1987). We instrument $\lambda_{sot}\tilde{y}_{ist}$ from equation (10) with $\lambda_{sot}\hat{y}_{ist}$, where $\hat{y}_{ist} = \tilde{y}_{ist} - \bar{y}_{is}$ and $\bar{y}_{is}$ is the average skill-specific experience in skill s of individual i over her entire career. The usefulness of this instrumental variable rests on the fact that it is not correlated with the individual-specific component of learning ability, because the latter should be driving only the average skill-specific experience over the career. For example, a person with high individual learning ability in cognitive skills may be more

[28] In all of the robustness checks we use broad occupations instead of narrow occupations to control for occupation fixed effects. This is because the number of narrow occupations is too high to conduct most of the robustness checks for a given number of observations.

Table 5: *Robustness checks: estimating learning speed β with alternative methodologies*

	Baseline	Occ × Year FE	Bootstrap SE	IV
Social sk weight	0.0616*	0.126***	0.0616*	0.0683**
	(0.0323)	(0.0364)	(0.0365)	(0.0324)
Physical sk weight	-0.0183	-0.00747	-0.0183	-0.0105
	(0.0294)	(0.0305)	(0.0279)	(0.0294)
Technical sk weight	-0.0290	0.0274	-0.0290	-0.0262
	(0.0245)	(0.0262)	(0.0221)	(0.0244)
Cognitive sk weight	0.127**	0.0912*	0.127**	0.135***
	(0.0498)	(0.0532)	(0.0559)	(0.0496)
Weighted experience in Social sk	0.155***	0.143***	0.155***	0.146***
	(0.00971)	(0.0118)	(0.0109)	(0.00965)
Weighted experience in Physical sk	0.0340**	0.0374**	0.0340**	0.0243*
	(0.0138)	(0.0169)	(0.0149)	(0.0137)
Weighted experience in Technical sk	0.143***	0.120***	0.143***	0.139***
	(0.0146)	(0.0160)	(0.0135)	(0.0145)
Weighted experience in Cognitive sk	0.238***	0.212***	0.238***	0.230***
	(0.0261)	(0.0346)	(0.0236)	(0.0255)
Broad occupations	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	52657	52657	52657	52657
Adjusted R^2	0.349	0.357	0.349	

All regressions include year fixed effects.

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

likely to have a high skill-specific experience in cognitive skills on average in her career, but the deviation of skill-specific experience from the average will not be likewise related to the individual learning ability. Implementing this IV estimator also leaves our baseline results unchanged.

Several robustness checks are conducted pertaining to sample selection. Specifically, we drop the youngest individuals (<14 or <16 in 1997) from our sample. It is possible, for instance, that the importance of experience in physical skills in a wage regression is an artifact of the young age of the sample, and that these skills would not matter in a more mature sample. In a similar spirit, we replicate our results disregarding any work experience gained before graduating from high school. In both exercises, our main results remain valid.[29]

The final robustness checks once again look into the potential bias described in Lemma 3. In Table 6, we try to directly measure the quality of the match between worker and occupation. Good match between a worker and an occupation can be positively correlated with skill-specific experience, and at the same time can have a positive effect on wage through channels other than experience. Using the data on college attendance from NLSY97, we compare the last chosen college specialization (major) with a worker's current occupation. We then create a dummy variable *Good match* equal to 1 if

[29]Results are available upon request.

Table 6: *Estimating equation (10) for those with college degree, controlling for match quality.*

	Baseline high	Match quality	Firm size	Both with Occ × Year FE
Social sk weight	0.164***	0.177***	0.268***	0.340***
	(0.0493)	(0.0501)	(0.0576)	(0.0657)
Physical sk weight	0.0203	0.0348	0.0838	0.136**
	(0.0487)	(0.0488)	(0.0578)	(0.0635)
Technical sk weight	-0.0203	-0.0106	-0.0333	0.0339
	(0.0375)	(0.0382)	(0.0414)	(0.0465)
Cognitive sk weight	0.172**	0.177**	0.129	0.142
	(0.0741)	(0.0753)	(0.0815)	(0.0906)
Weighted experience in Social sk	0.134***	0.125***	0.153***	0.141***
	(0.0143)	(0.0144)	(0.0157)	(0.0185)
Weighted experience in Physical sk	0.00821	-0.00461	-0.00157	-0.0220
	(0.0295)	(0.0284)	(0.0285)	(0.0336)
Weighted experience in Technical sk	0.179***	0.167***	0.183***	0.140***
	(0.0206)	(0.0208)	(0.0216)	(0.0257)
Weighted experience in Cognitive sk	0.211***	0.209***	0.195***	0.189***
	(0.0304)	(0.0305)	(0.0292)	(0.0461)
Good match		0.0848***		0.0776***
		(0.0152)		(0.0178)
Big firm			0.0751***	0.0769***
			(0.0108)	(0.0109)
Broad occupations	Yes	Yes	Yes	Yes
Occupation × Year FE	No	No	No	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	21581	21115	14618	14318
Adjusted R^2	0.453	0.457	0.488	0.504

Clustered standard errors in parentheses. FE estimator. All regressions include year fixed effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

occupation fits with previous college specialization, and zero otherwise. For example, if a worker majored in Business management, and her current occupation belongs to a broad group 11 (Management occupations) or 13 (Business and Financial Operations occupations), she will be considered a good match with her current occupation.[30] Since college specializations are only available for those who actually attended college, for this exercise we restrict our sample to individuals with college education.

The first column of Table 6 shows the baseline results for a college-educated sample. Here, experience in physical skills is no longer statistically significant, as physical skills do not contribute to expertise of highly educated workers. Column (2) includes our proxy for match quality. The coefficient of *Good match* dummy is positive and statistically significant: workers receive 8.5 percentage points higher wage when they work in occupations that fit their college specialization. This is a within-individual effect. Nevertheless, our main results remain unchanged even controlling for match quality.

Another potential source of correlation analyzed in Section 2.1 are firm-specific

[30]The crosswalk between majors and broad occupations is available upon request.

skill weights λ. By construction, occupational skill requirements in our model vary only across occupations and time. It is easy, however, to imagine a scenario in which skill requirements are both occupation- and firm-specific. This would be the case if firms that are more productive systematically require higher level of skills. The unobserved firm-specific heterogeneity in skill weights λ will then enter the error term, and will generate bias if correlated with both wages and skill-specific experience $\tilde{y}$. In other words, if workers in more productive firms accumulate more skill-specific experience and receive higher wages, the estimates of learning rates β might be upward biased. To check whether firm productivity affects our results, in column (3) of Table 6 we include a dummy variable *Big firm* equal to 1 if the number of employees in the worker's current firm exceeds 249. The dummy variable is statistically significant and high in magnitude: workers in bigger firms on average receive 7.5 percentage points higher wages. Our coefficients of interest, however, remain qualitatively unchanged.[31]

Finally, in column (4) we add both the match quality dummy and the firm size dummy, and include Occupation×Year fixed effects. Although the magnitude of β coefficient for experience in technical and cognitive skills is slightly reduced, overall both qualitatively and quantitatively our results are unaffected.

The robustness checks performed in this section are designed to investigate the presence of potential biases. On the one hand, weights λ might not fully incorporate the market price of skills, which means that β_s captures not only the learning rate, but also the part of skill price. On the other hand, we cannot exclude the possibility of an omitted variable bias due to correlation between the error term and λ or $\tilde{y}$ (such as unobserved match quality or firm size). According to the estimation in column (4) of Table 6, neither of the biases significantly influences our results. Of course, the dummy variables *Good match* and *Big firm* might be naive proxies, which only control for one potential source of the correlation between the error term and occupation selection or skill-specific experience. Nevertheless, the evidence presented in Tables 5 and 6 suggests that our main results are robust.

[31]Interestingly, controlling for firm size does not affect the magnitude of β coefficients symmetrically. The coefficient of experience in cognitive skills becomes smaller, suggesting the presence of a small but anticipated upward bias – bigger firms require higher level of cognitive skills. On the other hand, the coefficient of experience in social skills goes up, indicating a downward bias when firm size is not taken into account – bigger firms require less social skills.

7.2 Square of skill-specific experience

In our theoretical setup, skill-specific productivity of workers – expertise – is a linear function of skill-specific experience $\tilde{y}$. According to equation (5), expertise e_{is} of individual i in skill s depends on her skill-specific experience $\tilde{y}_{is}$ in a linear fashion, with parameters α_s and β_s guiding the relationship. While such a specification brings simplicity and interpretability to the estimation, it is possible that the true relationship between expertise and experience is more complex. It is common, for instance, to control for a polynomial of (potential) experience in Mincer-type wage regressions to capture the diminishing returns to expertise. In this section, we introduce a second-order polynomial to equation (5). Specifically, expertise of worker i in skill s can now be described as follows:

$$e_{is} = \alpha_s + \beta_s \tilde{y}_{is} + \gamma_s \tilde{y}_{is}^2 + \epsilon_{is} \tag{13}$$

The resulting regression equation then takes the following form:

$$w_{io} = \sum_{s=1}^{S} \alpha_s \lambda_{so} + \sum_{s=1}^{S} \beta_s \lambda_{so} \tilde{y}_{is} + \sum_{s=1}^{S} \gamma_s \lambda_{so} \tilde{y}_{is}^2 + \sum_{s=1}^{S} \lambda_{so} \epsilon_{is} + u_{io} \tag{14}$$

Table 7 shows the estimation results for the polynomial specification. Column (1) contains our baseline estimates of the learning speed. The squared terms of skill-specific experience are added in column (2). As before, the regressions control for age and age squared, marital dummy, years of education, and tenure in current job, and include individual, year, narrow occupation and broad industry fixed effects.

According to Table 7, coefficients γ in all four skills are statistically significant (at the 1% level except for physical skills) and negative. The negative sign of the coefficients suggests that skill-specific experience can be characterized by diminishing marginal returns: every additional unit of skill-specific experience contributes to the accumulation of expertise at a decreasing rate. Of course, the introduction of a squared experience has affected the magnitude of the β coefficients. However, the original ranking is preserved. Even with the polynomial specification individuals accumulate expertise in cognitive skills and the highest rate, followed by technical, social, and lastly physical skills.

Adding a second-order polynomial somewhat complicates the interpretation of the coefficients in terms of learning speed, but it may also be making the dynamics significantly more realistic. In particular, cognitive skills are accumulated very quickly

Table 7: *Estimating β and γ*

	(1)	(2)
Social sk weight	-0.185***	-0.210***
	(0.0471)	(0.0474)
Physical sk weight	-0.167***	-0.180***
	(0.0484)	(0.0503)
Technical sk weight	-0.191***	-0.223***
	(0.0301)	(0.0339)
Cognitive sk weight	0.0774	-0.119
	(0.0740)	(0.0781)
Weighted experience in Social sk	0.112***	0.194***
	(0.00976)	(0.0209)
Weighted experience in Physical sk	0.0301**	0.0706***
	(0.0138)	(0.0260)
Weighted experience in Technical sk	0.155***	0.274***
	(0.0142)	(0.0279)
Weighted experience in Cognitive sk	0.220***	0.654***
	(0.0278)	(0.0662)
Weighted experience in Social sk, sq		-0.0166***
		(0.00364)
Weighted experience in Physical sk, sq		-0.00884*
		(0.00460)
Weighted experience in Technical sk, sq		-0.0423***
		(0.00652)
Weighted experience in Cognitive sk, sq		-0.110***
		(0.0158)
Broad occupations	No	No
Narrow occupations	Yes	Yes
Industry FE	Yes	Yes
Observations	52657	52657
Adjusted R^2	0.401	0.403

Clustered standard errors in parentheses. FE estimator. All regressions include year fixed effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(based on the β coefficient), but the returns to experience in cognitive skills diminish at a high rate as well. In order to evaluate the speed of expertise accumulation while taking into account both β and γ, we examine the expertise profiles in Figure 7, and compare them with those in Figure 4.

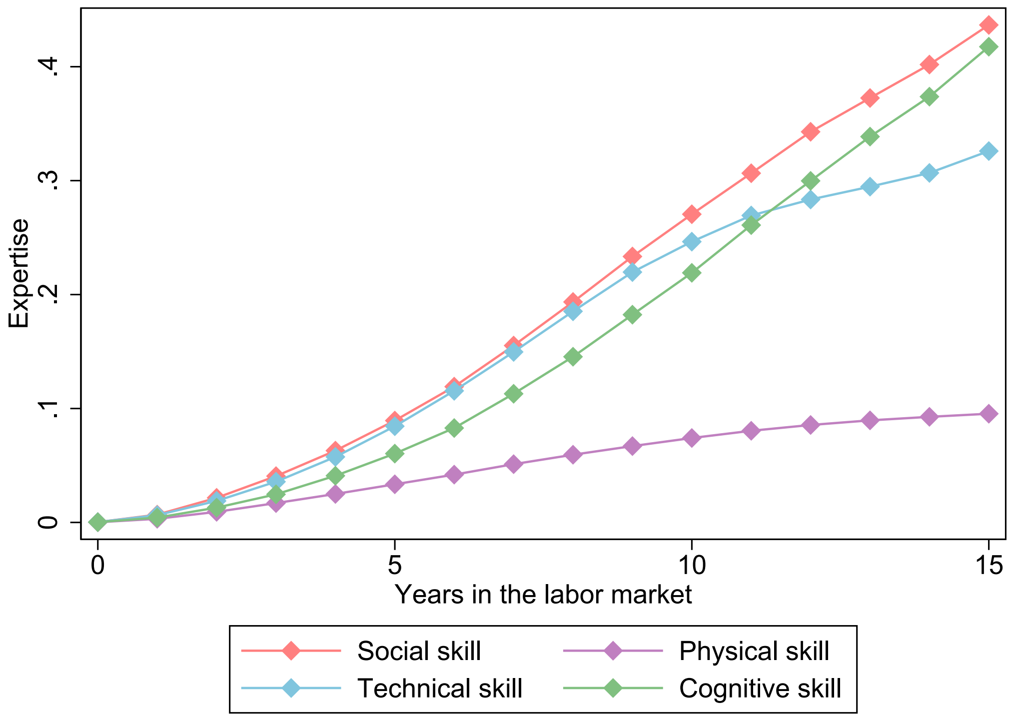


Figure 7: *Average expertise $\bar{e}_s$ against total years of labor market experience y.*

The inclusion of the second-order polynomial drastically changes the expertise profile in cognitive skills. According to Figure 4, expertise in cognitive skills was accumulated rather slowly. The expertise profile was driven by the low level of skill-specific experience in cognitive skills, which resulted in low level of expertise despite the highest learning rate. Adding a square of experience boosts the speed of acquisition of cognitive skills significantly. As depicted in Figure 7, expertise in cognitive skills is gained very quickly at the beginning of a worker's career. After 15 years in the labor market, the workers in our sample accumulate similar levels of expertise in social and cognitive skills, despite having very little experience in the latter. The behavior of other skills is not affected by the square of experience.

The results presented above emphasize the potential importance of allowing for a more complex relationship between expertise and skill-specific experience. Linear dependency underestimates the speed of learning of cognitive skills at the beginning of a worker's career (and potentially overestimates the learning rate at a later stage).

8 Conclusion

In this paper, we introduced a new concept of skill-specific experience. We measured this variable by combining the data on occupational skill requirements from O*NET with occupational histories from NLSY97. We found that the skill-specific experience profiles can vary both across skill categories and across different demographic groups. In the regression setting, we showed that skill-specific experience is a good proxy for expertise in different skills. The resulting estimates correspond to skill-specific learning rates while on the job. To our knowledge, we are the first to explicitly estimate learning rates in multidimensional skills. At the same time, we classified possible sources and directions of bias and showed ways to correct for the bias. Conceptually, our research introduced a framework that is flexible and mostly agnostic regarding the structure of the labor market in general and occupational choice in particular. Thus, a wide range of labor market structures can fit our framework for the purpose of studying skill-specific experience or expertise.

At the beginning of their career young individuals accumulate the most experience in social and physical skills. They enter the labor market by working part-time in occupations that require no special qualifications, such as cashiers, waiters, and retail salespersons. Over the course of their career they gradually switch away from physically-intensive occupations, gaining the most experience in social skills. Moreover, there exist heterogeneities in the level of skill-specific experience across gender and education. Men and non-college educated respondents seem to specialize more in technically- and physically-intensive occupations, while women and individuals with some college education concentrate in occupations with higher requirements of social skills.

Expertise in different skills grows with skill-specific experience. Our estimates of parameter β_s are positive and statistically significant for all skills. Cognitive skills can be characterized by the highest speed of learning: individuals quickly gain expertise in these skills by working in cognitively demanding occupations. The expertise profile in cognitive skills is steep and increasing despite the fact that the workers in our sample accumulate little experience in such skills. Physical skills represent an outlier, as experience in occupations that require physical labor seems to affect expertise in these skills only marginally. This result highlights the unproductive nature of physical skills, as practicing these skills does not lead to accumulation of human capital. The heterogeneity analysis of the learning speed shows that there are considerable differences in expertise accumulation between men and women despite the fact that their learning speed in most skills is very similar. Hence, past occupational choices are an important determinant of expertise in different skills.

An important assumption underlying the interpretation of our main results is that individual-specific variables (related to the learning rate) do not drive occupational choice. We have addressed the concerns about the bias with a range of robustness exercises, including an IV estimator and controls for match quality and firm productivity. The impressive robustness of our results suggests that any bias is small and negligible, and we cautiously interpret our estimates as true learning rates of skills in a young sample. Further research could introduce explicit occupation matching functions to our setup, and use simulation methods in estimation. This may lead to less general results, as the occupation matching function has to be chosen, but may allow to estimate skill-specific depreciation and skill-specific learning rates in education.

Finally, a vital component of our paper is the measures of weights that are associated with skills for each occupation. O*NET provides the best possible data for the construction of such weights. However, these data are produced through surveys and are prone to framing effects and potentially complex judgment calls. Given the inevitable noise in the data, we consider the statistical significance of our estimates a success. In addition, the study of multidimensional skills poses the challenge of the reduction of available data on skill requirements to a manageable set. We have suggested NMF as a valid approach to select the most representative skills within O*NET, which also allows us to interpret the dimensions of skill. Apart from nonnegativity of the components, our method is superior to the standard PCA methodology, as it makes the selection of skills more objective. However, the question of dimensionality reduction remains important, as there is no clear method to follow. A comparative analysis based on different dimensionality reduction approaches could be fruitful for the literature.

References

Altonji, J. G. and Shakotko, R. A. (1987). Do Wages Rise with Job Seniority? *The Review of Economic Studies*, **54** (3), 437–459.

Autor, D. H., Levy, F. and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, **118** (4), 1279–1333.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, **70** (5), 9–49.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, **75** (4), 352–365.

Blandin, A. (2018). Learning by Doing and Ben-Porath: Life-cycle Predictions and Policy Implications. *Journal of Economic Dynamics and Control*, **90**, 220–235.

Cattell, R. B. (1963). Theory of Fluid and Crystallized Intelligence. *Journal of Educational Psychology*, **54** (1), 1–22.

Cossa, R., Heckman, J. J. and Lochner, L. (1999). Wage Subsidies and Skill Formation: A Study of the Earned Income Tax Credit. *Unpublished paper.*

Engbom, N. (2021). Labor Market Fluidity and Human Capital Accumulation. *Working paper.*

Flinn, C., Gemici, A. and Laufer, S. (2017). Search, matching and training. *Review of Economic Dynamics*, **25**, 260–297, Special Issue on Human Capital and Inequality.

Févotte, C. and Idier, J. (2011). Algorithms for Nonnegative Matrix Factorization with the β-Divergence. *Neural Computation*, **23** (9), 2421–2456.

Gathmann, C. and Schönberg, U. (2010). How General Is Human Capital? A Task-Based Approach. *Journal of Labor Economics*, **28** (1), 1–49.

Guvenen, F., Kuruscu, B., Tanaka, S. and Wiczer, D. (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics*, **12** (1), 210–44.

Imai, S. and Keane, M. P. (2004). Intertemporal Labor Supply and Human Capital Accumulation. *International Economic Review*, **45** (2), 601–641.

Ingram, B. F. and Neumann, G. R. (2006). The returns to skill. *Labour Economics*, **13** (1), 35–59.

Kambourov, G. and Manovskii, I. (2009). Occupational Specificity of Human Capital. *International Economic Review*, **50** (1), 63–115.

Killingsworth, M. R. (1982). "Learning by Doing" and "Investment in Training": A Synthesis of Two "Rival" Models of the Life Cycle. *The Review of Economic Studies*, **49** (2), 263–271.

Kuruscu, B. (2006). Training and Lifetime Income. *American Economic Review*, **96** (3), 832–846.

Lazear, E. P. (2009). Firm-Specific Human Capital: A Skill-Weights Approach. *Journal of Political Economy*, **117** (5), 914–940.

Lee, D. and Seung, H. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, **401**, 788–791.

Lise, J. and Postel-Vinay, F. (2020). Multidimensional Skills, Sorting, and Human Capital Accumulation. *American Economic Review*, **110** (8), 2328–76.

Poletaev, M. and Robinson, C. (2008). Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000. *Journal of Labor Economics*, **26** (3), 387–420.

Sanders, C. (2016). Skill Accumulation, Skill Uncertainty, and Occupational Choice. *Unpublished manuscript, Washington University in St. Louis.*

— and Taber, C. (2012). Life-Cycle Wage Growth and Heterogeneous Human Capital. *Annual Review of Economics*, **4**, 399–425.

Shaw, K. L. (1989). Life-Cycle Labor Supply with Human Capital Accumulation. *International Economic Review*, **30** (2), 431–456.

Yamaguchi, S. (2012). Tasks and Heterogeneous Human Capital. *Journal of Labor Economics*, **30** (1), 1–53.

Appendix

A1 Baseline sample

To construct our baseline sample, we implemented the following cleaning procedures:

- introduce a cap on hours worked per week at 112, assuming that individual can work a maximum of 16 hours per day 7 days a week;
- introduce a cap on hours worked per year at 5,824 hours, assuming that individual can work a maximum of 16 hours per day 7 days a week;
- replace missing hours worked per year with weeks worked per year, multiplied by 40 (in case we know weeks worked but not the hours)
- correct hourly wages for inflation, using CPI provided by OECD, with 2015 as base year (Source: https://data.oecd.org/price/inflation-cpi.htm). Replace wage with missing for wages above 99$^{\text{th}}$ and below 1$^{\text{st}}$ percentiles.

Given the relatively small number of individuals with complete working histories, we have attempted to impute missing values when possible. The following variables contain imputed values:

- marital status. We use reverse cascade (filling in values at t based on values at $t+1$) for the cases of "never married". We also replace missing values with "never married" for those with age below 17, as they are likely to be below the legal marriage age.
- age. Replace missing at t with age at $t-1$ plus one.
- highest degree received yearly (`degree_year`). We fill in the missing value at t if value at $t-1$ is equal to the value at $t+1$. We use the same technique in cases of several missing values. Finally, we compare `degree_year` with highest degree received ever (a time-invariant variable `degree_ever`). In cases when `degree_year` is not missing and equal to `degree_ever`, we use cascade to replace all the following missing values in `degree_year`.

- job identifier (`jobid`). We fill in the missing value at t if `jobid` at $t-1$ is equal to `jobid` at $t+1$, and individual was employed at t. We use the same technique in cases of several missing values. We also impute `jobid` on a case-by-case basis (for example, based on the information on labor force participation status and job identifiers in the following year, we can deduce the year in which the job switch occurred). Note, however, that these imputations affect our baseline sample only in cases when they are *not* accompanied by missing information on occupations. All the individuals with missing occupation information get dropped even if we imputed `jobid`.

- occupation. We fill in the missing value at t if occupation at $t-1$ is equal to occupation at $t+1$, and individual was employed at t. We use the same technique in cases of several missing values, but only if `jobid` doesn't change (it is unlikely that individual would change her occupation, and then return to her previous occupation while working for the same employer).

The following control variables were created:

- years of education (time-varying). This variable was generated based on the information on highest grade, highest degree received yearly (`degree_year`) and highest degree received ever (`degree_ever`). We assume that one needs 14 years of studying (12 years of school and 2 years of college) to obtain an Associates degree, 16 years of studying to get a Bachelor degree, 18 years to get a Master degree, and 20 years to get a PhD.

- labor force participation status. This variable was generated based on weekly employment arrays. We create two variables capturing labor force participation status: if respondent worked for at least 12 weeks during the year, and if respondent worked during the last week of the year.

- labor market experience (time-varying). This variable was generated based on the labor force participation status. We set labor market experience equal to one during the first year of employment. Experience then mechanically increases by one in each year with positive labor force participation status, and remains the same for years in which respondent was not employed.

A2 Skill categories from O*NET

Table 8: *List of narrow Abilities, Skills and Knowledge categories, downloaded from O*NET to compute* λ_{sot}

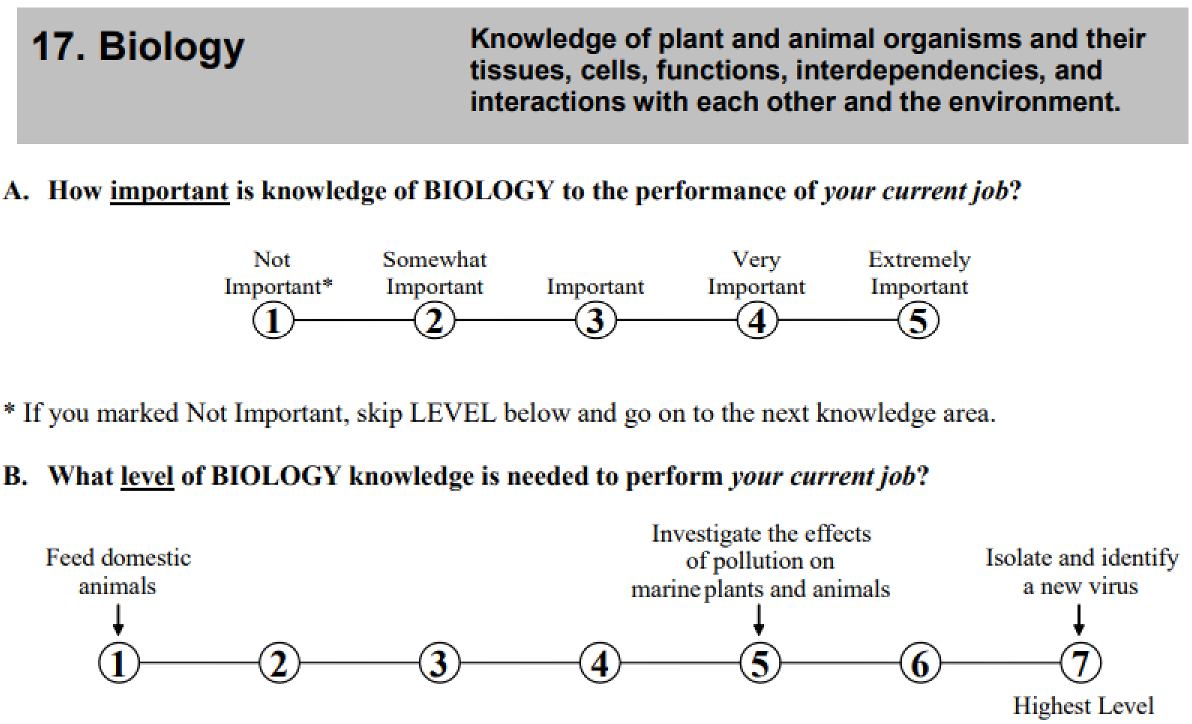
Group	Subgroup	Skill Title
Abilities	Cognitive Abilities	Oral Comprehension
Abilities	Cognitive Abilities	Written Comprehension
Abilities	Cognitive Abilities	Oral Expression
Abilities	Cognitive Abilities	Written Expression
Abilities	Cognitive Abilities	Fluency of Ideas
Abilities	Cognitive Abilities	Originality
Abilities	Cognitive Abilities	Problem Sensitivity
Abilities	Cognitive Abilities	Deductive Reasoning
Abilities	Cognitive Abilities	Inductive Reasoning
Abilities	Cognitive Abilities	Information Ordering
Abilities	Cognitive Abilities	Category Flexibility
Abilities	Cognitive Abilities	Mathematical Reasoning
Abilities	Cognitive Abilities	Number Facility
Abilities	Cognitive Abilities	Memorization
Abilities	Cognitive Abilities	Speed of Closure
Abilities	Cognitive Abilities	Flexibility of Closure
Abilities	Cognitive Abilities	Perceptual Speed
Abilities	Cognitive Abilities	Spatial Orientation
Abilities	Cognitive Abilities	Visualization
Abilities	Cognitive Abilities	Selective Attention
Abilities	Cognitive Abilities	Time Sharing
Abilities	Psychomotor Abilities	Arm-Hand Steadiness
Abilities	Psychomotor Abilities	Manual Dexterity
Abilities	Psychomotor Abilities	Finger Dexterity
Abilities	Psychomotor Abilities	Control Precision
Abilities	Psychomotor Abilities	Multilimb Coordination
Abilities	Psychomotor Abilities	Response Orientation
Abilities	Psychomotor Abilities	Rate Control
Abilities	Psychomotor Abilities	Reaction Time
Abilities	Psychomotor Abilities	Wrist-Finger Speed
Abilities	Psychomotor Abilities	Speed of Limb Movement
Abilities	Physical Abilities	Static Strength
Abilities	Physical Abilities	Explosive Strength
Abilities	Physical Abilities	Dynamic Strength
Abilities	Physical Abilities	Trunk Strength
Abilities	Physical Abilities	Stamina
Abilities	Physical Abilities	Extent Flexibility
Abilities	Physical Abilities	Dynamic Flexibility
Abilities	Physical Abilities	Gross Body Coordination
Abilities	Physical Abilities	Gross Body Equilibrium
Abilities	Sensory Abilities	Near Vision
Abilities	Sensory Abilities	Far Vision
Abilities	Sensory Abilities	Visual Color Discrimination
Abilities	Sensory Abilities	Night Vision

Abilities	Sensory Abilities	Peripheral Vision
Abilities	Sensory Abilities	Depth Perception
Abilities	Sensory Abilities	Glare Sensitivity
Abilities	Sensory Abilities	Hearing Sensitivity
Abilities	Sensory Abilities	Auditory Attention
Abilities	Sensory Abilities	Sound Localization
Abilities	Sensory Abilities	Speech Recognition
Abilities	Sensory Abilities	Speech Clarity
Basic Skills	Content	Reading Comprehension
Basic Skills	Content	Active Listening
Basic Skills	Content	Writing
Basic Skills	Content	Speaking
Basic Skills	Content	Mathematics
Basic Skills	Content	Science
Basic Skills	Process	Critical Thinking
Basic Skills	Process	Active Learning
Basic Skills	Process	Learning Strategies
Basic Skills	Process	Monitoring
Cross-Functional Skills	Social Skills	Social Perceptiveness
Cross-Functional Skills	Social Skills	Coordination
Cross-Functional Skills	Social Skills	Persuasion
Cross-Functional Skills	Social Skills	Negotiation
Cross-Functional Skills	Social Skills	Instructing
Cross-Functional Skills	Social Skills	Service Orientation
Cross-Functional Skills	Complex Problem-Solving Skills	Complex Problem Solving
Cross-Functional Skills	Technical Skills	Operations Analysis
Cross-Functional Skills	Technical Skills	Technology Design
Cross-Functional Skills	Technical Skills	Equipment Selection
Cross-Functional Skills	Technical Skills	Installation
Cross-Functional Skills	Technical Skills	Programming
Cross-Functional Skills	Technical Skills	Operation Monitoring
Cross-Functional Skills	Technical Skills	Operation and Control
Cross-Functional Skills	Technical Skills	Equipment Maintenance
Cross-Functional Skills	Technical Skills	Troubleshooting
Cross-Functional Skills	Technical Skills	Repairing
Cross-Functional Skills	Technical Skills	Quality Control Analysis
Cross-Functional Skills	Systems Skills	Judgment and Decision Making
Cross-Functional Skills	Systems Skills	Systems Analysis
Cross-Functional Skills	Systems Skills	Systems Evaluation
Cross-Functional Skills	Resource Management Skills	Time Management
Cross-Functional Skills	Resource Management Skills	Management of Financial Resources
Cross-Functional Skills	Resource Management Skills	Management of Material Resources
Cross-Functional Skills	Resource Management Skills	Management of Personnel Resources
Knowledge	Business and Management	Administration and Management
Knowledge	Business and Management	Clerical
Knowledge	Business and Management	Economics and Accounting
Knowledge	Business and Management	Sales and Marketing
Knowledge	Business and Management	Customer and Personal Service
Knowledge	Business and Management	Personnel and Human Resources
Knowledge	Manufacturing and Production	Production and Processing
Knowledge	Manufacturing and Production	Food Production
Knowledge	Engineering and Technology	Computers and Electronics

Knowledge	Engineering and Technology	Engineering and Technology
Knowledge	Engineering and Technology	Design
Knowledge	Engineering and Technology	Building and Construction
Knowledge	Engineering and Technology	Mechanical
Knowledge	Mathematics and Science	Mathematics
Knowledge	Mathematics and Science	Physics
Knowledge	Mathematics and Science	Chemistry
Knowledge	Mathematics and Science	Biology
Knowledge	Mathematics and Science	Psychology
Knowledge	Mathematics and Science	Sociology and Anthropology
Knowledge	Mathematics and Science	Geography
Knowledge	Health Services	Medicine and Dentistry
Knowledge	Health Services	Therapy and Counseling
Knowledge	Education and Training	Education and Training
Knowledge	Arts and Humanities	English Language
Knowledge	Arts and Humanities	Foreign Language
Knowledge	Arts and Humanities	Fine Arts
Knowledge	Arts and Humanities	History and Archeology
Knowledge	Arts and Humanities	Philosophy and Theology
Knowledge	Law and Public Safety	Public Safety and Security
Knowledge	Law and Public Safety	Law and Government
Knowledge	Communications	Telecommunications
Knowledge	Communications	Communications and Media
Knowledge	Business and Management	Transportation

A3 O*NET questionnaires

For each occupational characteristic within Skills, Knowledge and Abilities, O*NET questionnaires contain two questions: one estimates the *importance* of the characteristic, and the other evaluates the *level* of the characteristic required to perform a job. The level of characteristic can vary between 1 and 7, with 7 being the highest skill requirement. Furthermore, the level scale has anchors to help the respondents identify the level of skill requirement (in the example below the anchors are on levels 1, 5 and 7). We use only the level information to generate skill weights λ.



17. Biology — **Knowledge of plant and animal organisms and their tissues, cells, functions, interdependencies, and interactions with each other and the environment.**

A. How important is knowledge of BIOLOGY to the performance of *your current job*?

Not Important* (1) — Somewhat Important (2) — Important (3) — Very Important (4) — Extremely Important (5)

* If you marked Not Important, skip LEVEL below and go on to the next knowledge area.

B. What level of BIOLOGY knowledge is needed to perform *your current job*?

Feed domestic animals → (1) — (2) — (3) — (4) — (5) ← Investigate the effects of pollution on marine plants and animals — (6) — (7) ← Isolate and identify a new virus

Highest Level

Figure 8: *Example question from O*NET Knowledge questionnaire*

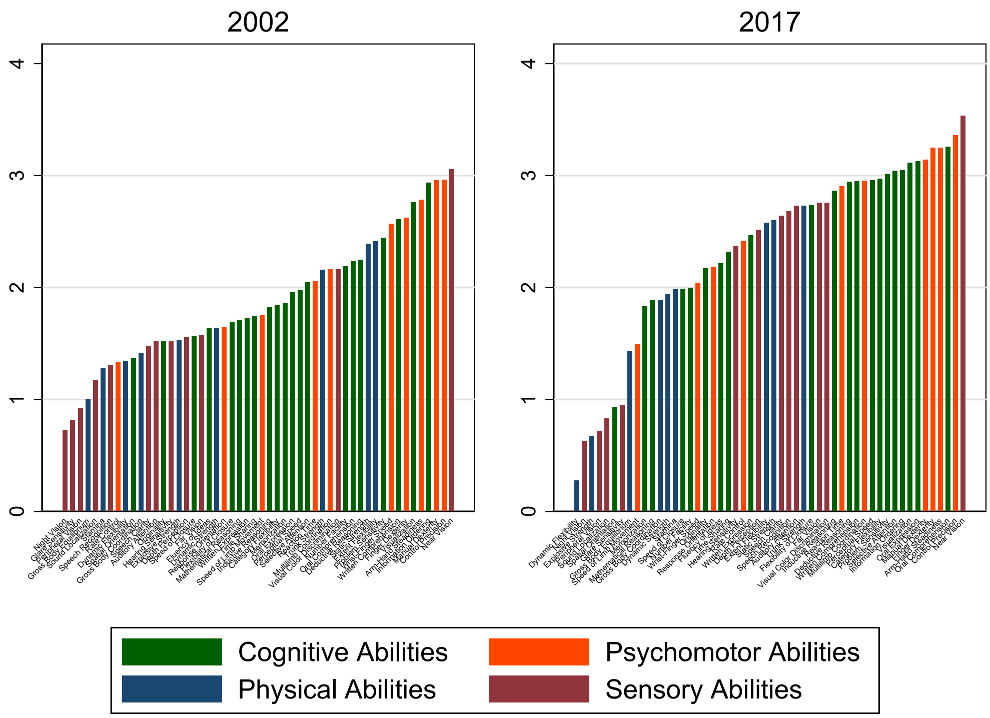


Figure 9: *Level of skills from group Ability that were required in Production Occupations in 2002 and 2017. Production required more cognitive and fewer physical skills in 2017 compared to 2002.*

A4 Dimensionality reduction procedure

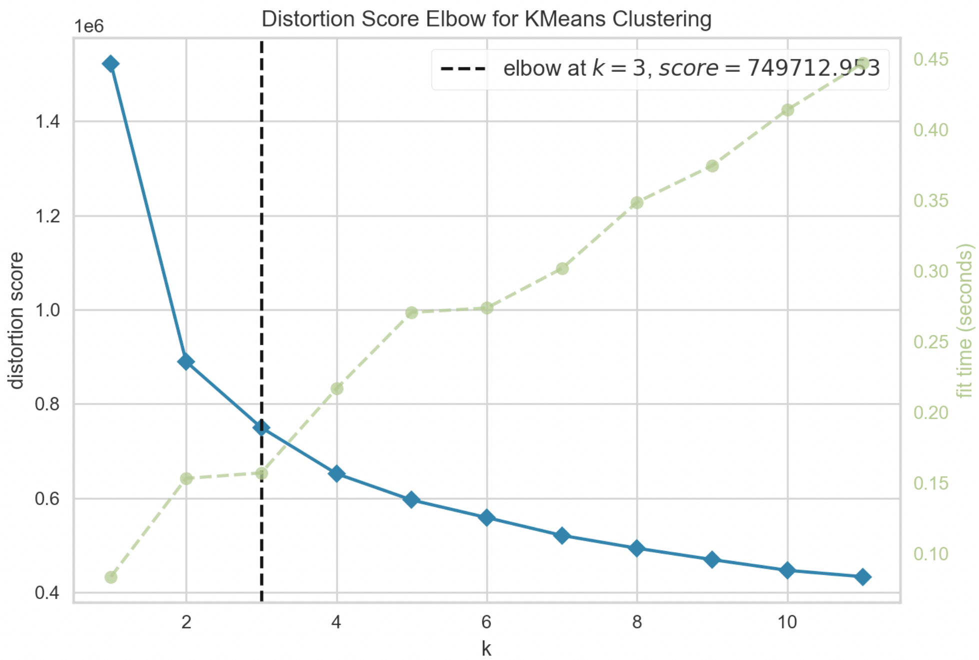


Figure 10: *Elbow method. According to the elbow method, the optimal number of skills is 3 (elbow at k=3).*

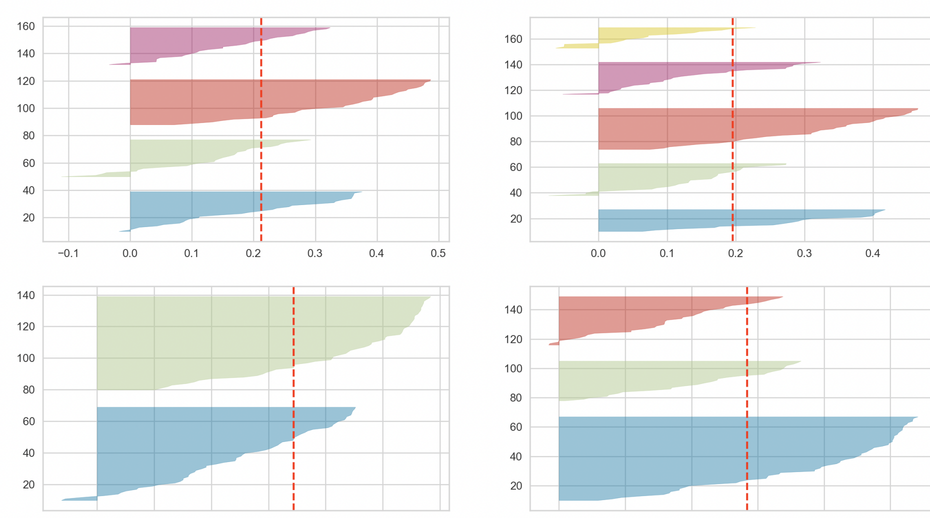


Figure 11: *Silhouette method. According to the silhouette method, any number of skills between 2 and 5 is acceptable.*

A5 Reduced skills

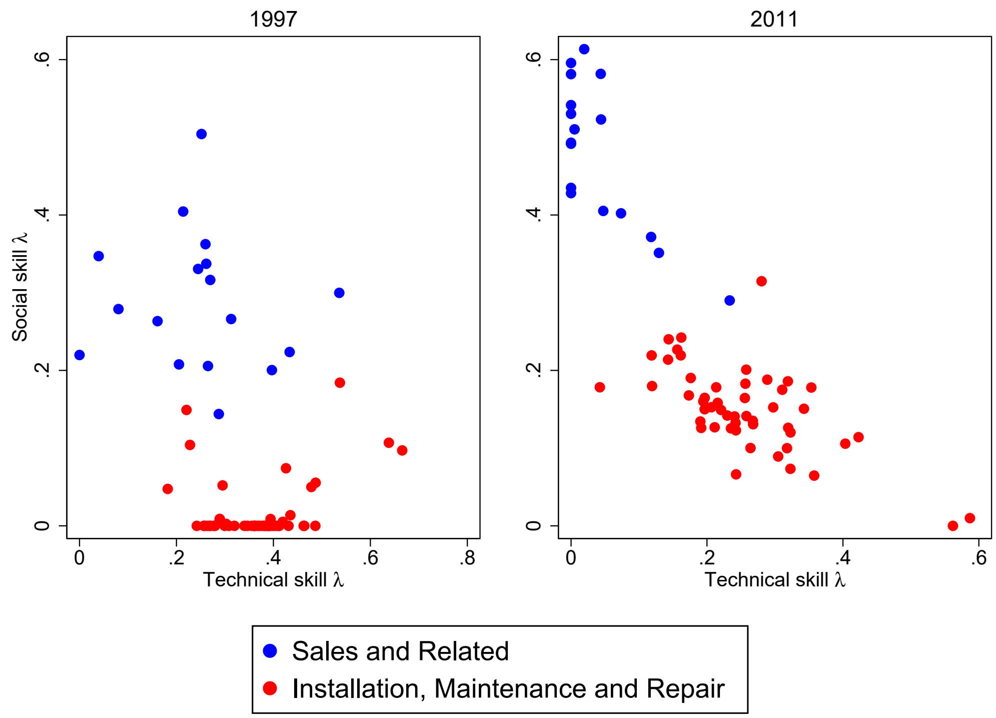


Figure 12: *Estimated λ in social (y-axis) and technical (x-axis) skills for two occupation groups: Sales and Installation. Each dot represents a narrow occupation within broad occupation groups.*

A6 Average skill weights

Table 9: *Average skill weights λ in the sample.*

	All		Women		Men		No college		College	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Social sk weight	0.31	0.17	0.36	0.14	0.24	0.17	0.27	0.16	0.36	0.17
Physical sk weight	0.24	0.15	0.20	0.12	0.29	0.16	0.27	0.14	0.19	0.14
Technical sk weight	0.18	0.16	0.15	0.15	0.21	0.17	0.18	0.15	0.18	0.17
Cognitive sk weight	0.07	0.10	0.05	0.07	0.09	0.11	0.06	0.08	0.08	0.11
Observations	59528		31116		28412		35366		23973	

A7 Weekly hours worked

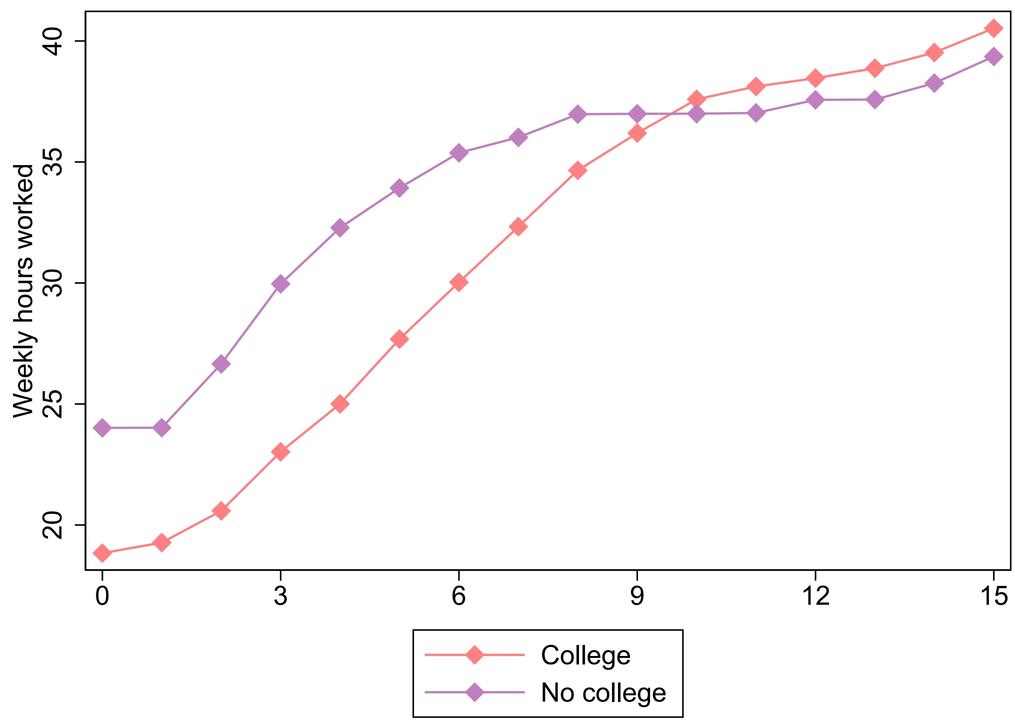


__Figure 13:__ Weekly hours worked over years in the labor market by education groups. College stands for people with some college degree (Associates, Bachelors, Masters or PhD degree), No college stands for people without college degree (GED, high-school diploma or no degree at all).

A8 Heterogeneities

Table 10: *Estimating equation (10) by gender.*

	(1)	(2)
	Women	Men
Social sk weight	-0.251***	-0.152**
	(0.0671)	(0.0666)
Physical sk weight	-0.170**	-0.106
	(0.0730)	(0.0668)
Technical sk weight	-0.186***	-0.183***
	(0.0464)	(0.0400)
Cognitive sk weight	0.117	-0.0509
	(0.129)	(0.0925)
Weighted experience in Social sk	0.115***	0.132***
	(0.0125)	(0.0157)
Weighted experience in Physical sk	-0.0605	0.0182
	(0.0386)	(0.0162)
Weighted experience in Technical sk	0.156***	0.137***
	(0.0246)	(0.0174)
Weighted experience in Cognitive sk	0.216***	0.203***
	(0.0656)	(0.0314)
Age	0.0361**	0.0829***
	(0.0167)	(0.0182)
Age squared	-0.000808**	-0.00182***
	(0.000317)	(0.000338)
Ever married dummy	0.0119	0.0744***
	(0.00991)	(0.0117)
Years of education	0.0358***	0.0199***
	(0.00313)	(0.00355)
Tenure in current job	0.0149***	0.0178***
	(0.00228)	(0.00225)
Constant	1.602***	1.413***
	(0.270)	(0.299)
Narrow occupations	Yes	Yes
Industry FE	Yes	Yes
Observations	27702	24955
Adjusted R^2	0.406	0.421

Clustered standard errors in parentheses. FE estimator. All regressions include year fixed effects.

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

Table 11: *Estimating equation (10) by education.*

	(1)	(2)	(3)
	Baseline	Low	High
Social sk weight	-0.185***	-0.100*	-0.310***
	(0.0471)	(0.0585)	(0.0783)
Physical sk weight	-0.167***	-0.165***	-0.235***
	(0.0484)	(0.0614)	(0.0807)
Technical sk weight	-0.191***	-0.136***	-0.250***
	(0.0301)	(0.0377)	(0.0480)
Cognitive sk weight	0.0774	0.103	-0.105
	(0.0740)	(0.0970)	(0.123)
Weighted experience in Social sk	0.112***	0.109***	0.0862***
	(0.00976)	(0.0143)	(0.0137)
Weighted experience in Physical sk	0.0301**	0.0628***	-0.0119
	(0.0138)	(0.0168)	(0.0291)
Weighted experience in Technical sk	0.155***	0.111***	0.192***
	(0.0142)	(0.0180)	(0.0210)
Weighted experience in Cognitive sk	0.220***	0.188***	0.187***
	(0.0278)	(0.0501)	(0.0330)
Age	0.0552***	0.0969***	0.0118
	(0.0124)	(0.0152)	(0.0219)
Age squared	-0.00123***	-0.00211***	-0.000399
	(0.000233)	(0.000281)	(0.000419)
Ever married dummy	0.0361***	0.0284***	0.0390***
	(0.00765)	(0.0101)	(0.0117)
Years of education	0.0296***	0.00941**	0.0255***
	(0.00237)	(0.00460)	(0.00441)
Tenure in current job	0.0159***	0.0228***	0.00786***
	(0.00161)	(0.00203)	(0.00259)
Constant	1.553***	1.271***	2.105***
	(0.214)	(0.303)	(0.292)
Narrow occupations	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	52657	31076	21581
Adjusted R^2	0.401	0.330	0.507

Clustered standard errors in parentheses. FE estimator. All regressions include year fixed effects.

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