Soft Skills and the Wage Progression of Low-Educated Workers*

Philippe Aghion Antonin Bergeaud

Richard Blundell Rachel Griffith

February 2022

Abstract

*Aghion, College de France, LSE and CEPR, p.aghion@lse.ac.uk; Bergeaud, Banque de France and CEP, antonin.bergeaud@banque-france.fr; Blundell, UCL, IFS and CEPR, r.blundell@ucl.ac.uk; Griffith, University of Manchester, IFS and CEPR, rgriffith@ifs.org.uk. We thank Nick Bloom, Juan Jimeno, Francis Kramarz, Attila Lindler, Stefanie Stantcheva and participants at the CIFAR, ENS-Cachan, PSE, NBER Summer Institute, ECB labor market workshop, ESCB Research Cluster, Bank of Italy-CEPR conference, LAGV and ASSA 2019 special session in Atlanta for helpful comments and suggestions. We thank the ESRC Centre for the Microeconomic Analysis of Public Policy (CPP) at IFS for funding under grant number ES/M010147/1.

1 Introduction

There is increasing interest in social and income mobility as relevant (inverse) measures of inequality in developed economies (e.g. see Chetty et al., 2014). This partly reflects the growing concern that the "American dream", - i.e. the possibility for all individuals to climb the social ladder no matter their social origins - is no longer operating.[1]

One response to the decline in social and income mobility, and more generally to the surge in income inequality over the past decades, has been to increase taxes and subsidies in order to foster redistribution. And in some countries such as the UK, taxes and benefits have been quite effective at boosting incomes at the bottom of the income distribution until quite recently (e.g. see Blundell et al., 2018). However, continuing to rely on the tax/subsidy lever alone may not be sufficient to restore social and income mobility.[2]

In this paper we explore another potential channel to foster social and income mobility, namely the "good jobs" channel. A good job is one that provides workers, even those with a low level of formal education, with favorable prospects for tenure, pay progression and promotion within the firm. This will be the case when there is the potential for the worker's marginal contribution to the firm's performance to increase over time. Measurable cognitive skills certainly play an important role, but particularly for low education workers, good jobs are those that enhance those workers' soft skills, in particular their ability to take initiatives and to interact and coordinate with other workers in the same firm.[3] Our main finding in this paper is that workers in occupations that require no, or only very low, levels of formal education indeed experience stronger pay progression in occupations where soft skills are more important.

To gain intuition of what we mean by soft skills, think of a worker in a low skilled occupation, for example a maintenance worker, a personal assistant or a sales or a telephonist, who shows outstanding initiative and reliability. These attributes may be difficult to measure and verify. Yet, they allow the worker to perform tasks which complement the tasks performed by workers in high skilled occupations within the firm in the sense that if performed well they can increase the productivity of the high

[1] For example, Figure 5 shows the strong pay progression that educated workers experience, compared to the lack of progression experienced by low educated workers in the UK.

[2] In the UK, spending on working age benefits, as a percentage of GDP, have nearly doubled between the end of the 1990s and the mid-2010s; while this has kept inequality from increasing it is a difficult level of expenditure to sustain.

[3] Barrera-Osorio et al. (2020) show that providing vocational training on soft skills increases employment and wages (in a randomised experiment in the field in Colombia).

skilled employees. Soft skills refer to a worker's ability to communicate and interact effectively with other actors in the firm.

We use O*NET Survey data to construct an index of occupations for which soft skills are important. The O*NET data describes the mix of knowledge, skills and abilities required in an occupation and the activities and tasks performed. The data is collected through surveys of US workers and occupational workers. Examples of low educated occupations where soft skills are important include receptionist, medical or school secretary, air transport operative, assembler, and where soft skills are less important include cleaner, bar staff, caretaker, packer, process operator. In Section 2.1 we provide empirical evidence, based on employment information in the UK, suggesting that low educated workers in high soft skill occupations: (i) experience steeper wage progression profiles with age; (ii) have higher prospects for career advancement; (iii) tend to be more satisfied with their working conditions.

To formalize the notion of soft skill tasks or jobs, in Section 3 we develop a model of wage bargaining with complementarity between low educated workers and the firm's other assets, to explain why workers in some low educated occupations with high soft skills get a higher premium and more wage progression than low educated workers with low soft skills. The main assumptions of the model are that: (i) a low educated worker's productivity on each task depends upon both an asset Q (which includes high skill workers) and her own quality q; (ii) in the model the tasks where soft skills are important - call them the "good jobs" - are those that involve more complementarity between q and Q and/or for which it is more difficult to replace the low educated worker on the spot by another low educated worker. Thus low educated workers on these tasks will command a higher bargaining power than low educated workers on tasks where soft skills matter less. Moreover, on tasks where soft skills matter the firm will want to increase the quality of the low educated worker, because this quality matters more to the firm's total surplus, hence the firm will invest more in training the low educated worker in order to increase that worker's quality.

In Section 4, we use matched employee-employer panel data to show that both the wage level and wage progression of workers in low educated occupations are higher in occupations where soft skills are important. Our main regressions control for a number of potentially confounding factors. Moreover, we show that the wage and wage progression in high soft skill occupations, are higher in more innovative firms.

Our work relates to several strands of literature. First, to the literature on wage inequality and skill-biased technical change. This literature looks at both, returns to cognitive skills (e.g. see Krusell et al., 2000; Acemoglu, 2002; Goldin and Katz, 2010) and returns to non-cognitive skills (see Acemoglu and Autor, 2011). We contribute to

this literature by pointing at a premium to soft skills among low educated occupation workers, and by providing evidence linking innovation to the rate at which the returns to soft skills increase with tenure. We compare those returns to returns from cognitive skills in Section 6.3.

Second, to a labor and wage literature (Gibbons and Katz, 1992; Groshen, 1991; Abowd et al., 1999; Bonhomme et al., 2019 among others) which emphasizes firm heterogeneity as an important source of wage differences across workers. This labor and wage literature has also pointed at the fact that in many countries there is considerable wage inequality among seemingly similar workers (see e.g. Card et al., 2016). Our analysis brings soft skills and firms' ability to enhance them by creating good jobs as another important source of wage heterogeneity across firms and among low educated workers.

Third, to a literature on soft skills (Brunello and Rocco, 2017; Barrera-Osorio et al., 2020; Carruthers and Jepsen, 2020; Silliman and Virtanen, 2019; Hanushek et al., 2017; Rodrik and Stantcheva, 2021; Battiston et al., 2017) that looks at how the development of soft skills in firms affects workers' satisfaction on the job and also their long-term career outcomes. We contribute to this literature by looking at how soft skills affect the wage level and wage progression of low educated workers, and how this depends upon characteristics of tasks/occupations – e.g. the extent to which these complement hard skills or other firm's assets – and upon characteristics of the firm, in particular its degree of innovativeness.

Finally, we draw on the literature on wage inequality and the organization of the firm (e.g. see Kremer, 1993; Kremer and Maskin, 1996; Garicano, 2000; Garicano and Rossi-Hansberg, 2006). We contribute to this literature by looking at the complementarity between workers in low educated occupations with high soft skills and the firm's other assets.[4]

The paper is organized as follows. In Section 2 we describe the data and our method to identify the occupations in which soft skills are important, and we show some initial correlations. In Section 3 we develop our theoretical framework and we lay out its main predictions. In Section 4 we confront our predictions to the data:

[4]By using information on the R&D intensity of firms, we also relate to recent papers look at the effects of innovation on income inequality using aggregate data (e.g. Aghion et al., 2018a; Akcigit et al., 2017). Taking a more microeconomic approach, Kline et al. (2018) for the US and Aghion et al. (2018b) for Finland, use administrative tax data merged with patent data to look at the individual returns from innovation to the inventors and to their co-workers. Both papers find significant returns to innovation, most of which accrue to other employees or stakeholders within the inventor's firm. We add to this literature by focusing on the returns to soft skills for workers in low educated occupations, and on how innovativeness increases the steepness of the wage progression of these workers.

we present our core regression analysis and we discuss the robustness of our main findings. Section 6 discusses potential threats to our identification strategy. Section 7 collects our concluding remarks.

2 Measuring wage progression and job characteristics

2.1 Wages and wage progression

Our main source of information on wages is the longitudinal dataset generated by the Annual Survey of Hours and Earnings (ASHE), which follows a random sample of 1% of the UK working population and is collected by the Office of National Statistics (ONS). ASHE contains detailed information on earnings, hours of work, gender, age, tenure and occupation. It records the firm that employs the workers, and this can be matched to information about the firm. We use data from the Business expenditures on Research and Development (BERD) survey, which contains detailed information on the firm's innovation activity. BERD is a census of firms with 400 or more employees, and a stratified random sample of firms below that size.

We use data for the period 2004-2019. We use information on male workers aged 18-49 who work in occupations that do not require any formal education in firms in the private sector that have 400 or more employees. Our main sample consists of *212,428* observations on *63,407* employees who work in *5,966* firms. In Section 6 we show that our results are robust to alternative sample selections, including using all firms and all workers in ASHE.

2.2 Occupation characteristics

We classify occupations by the importance of formal educational requirements, cognitive skills and soft skills.

Educational requirements

To measure the importance of qualifications we use the UK Regulatory Qualifications Framework (RQF). This framework is regulated by Ofqual (the regulator of qualifications and exams), and Appendix J defines the education level required for each 4-digit

occupation that is used for UK immigration purposes.[5] We aggregate these to three categories.

- **Low educated**, no formal qualifications necessary operatives: this includes assemblers, clerical, secretaries, cleaners, security drivers, technicians, sales
- **Medium educated**, typically requires A-level or some basic professional qualification: this includes trades, specialist clericals, associate professionals, medical or IT technicians, some managerial occupations
- **High educated**, typically requires higher education or an advanced professional qualification: this includes most managerial and executive occupations, engineers, scientists, R&D manager, bankers, other professions

Cognitive skills

We measure cognitive skills using the O*NET data to identify occupations where cognitive skills are important. The O*NET data describe the mix of knowledge, skills and abilities required in an occupation and the activities and tasks performed on that occupation. Workers are surveyed across occupations and asked to grade various characteristics or "dimensions" from 1 (when this dimension is not relevant to the workers' occupation) to 5 (when this dimension is very relevant to the workers' occupation). The O*NET data is based on surveys of workers and experts in the US. Our analysis is performed at the 4-digit SOC 2010 occupation level, this includes 361 occupations, 124 of which have little or no formal education requirements.

Many other papers have used the O*NET data to measure characteristics of occupations, most notable Acemoglu and Autor (2011). In section 6.3, we discuss how our results are related to this work.

We consider the following dimensions in the O*NET data to capture the importance of cognitive skills for workers in occupations with little or no formal educational requirements:

1. **Category Flexibility:** The ability to generate or use different sets of rules for combining or grouping things in different ways.
2. **Deductive Reasoning:** The ability to apply general rules to specific problems to produce answers that make sense.

[5] We use the version that was accessed from https://www.gov.uk/guidance/immigration-rules/immigration-rules-appendix-j-codes-of-practice-for-skilled-work on 4 June 2020.

3. **Fluency of Ideas:** The ability to come up with a number of ideas about a topic (the number of ideas is important, not their quality, correctness, or creativity).

4. **Inductive Reasoning:** The ability to combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events).

5. **Mathematical Reasoning:** The ability to choose the right mathematical methods or formulas to solve a problem.

6. **Information Ordering:** The ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules (e.g., patterns of numbers, letters, words, pictures, mathematical operations).

7. **Number Facility** The ability to add, subtract, multiply, or divide quickly and correctly.

We aggregate these dimensions into a single score **using factor** analysis. Table 1 presents the dimensions and their relative importance.

Table 1: Importance of cognitive skills

O*NET code	Characteristic description	Weight
abLV.1.A.1.b.7	Category Flexibility	0.3822
abLV.1.A.1.b.4	Deductive Reasoning	0.3935
abLV.1.A.1.b.1	Fluency of Ideas	0.3674
abLV.1.A.1.b.5	Inductive Reasoning	0.3799
abLV.1.A.1.c.1	Mathematical Reasoning	0.3753
abLV.1.A.1.c.2	Number Facility	0.3629
abLV.1.A.1.b.6	Information Ordering	0.3838

Notes: coordinates of the first eigen vector in a principal component analysis based on 7 characteristics taken from O*NET. Each characteristic has been aggregated at the SOC 4 digit occupation level using an employment weighted mean from the original O*NET-SOC occupation classification.

Soft skills

Soft skills are difficult to observe and to measure directly, both by the econometrician and the firm. To construct a proxy for the importance of soft skills we look at the task and skill content of occupations using the O*NET data.

We combine 10 items in O*NET to create a single index of occupations for which soft skills are important. These 10 dimensions are:

1. Problem Sensitivity: how big is the worker's ability to tell when something is wrong or is likely to go wrong?
2. Active listening: to which extent does the worker devote full attention to what other parties are saying, and how much time does she devote to understand the points that are made by other parties, asking questions whenever appropriate and not interrupting at inappropriate times?
3. Social Perceptiveness: to which extent is the worker aware of other parties' reactions and to which extent does she understand why the other parties react as they do?
4. Coordination: to which extent does the worker adjust her actions to the actions taken by the other parties?
5. Work With Work Group or Team: How important is it to work with others in a group or team in this job?
6. Coordinate or Lead Others: How important is it to coordinate or lead others in accomplishing work activities in this job?
7. Responsibility for Outcomes and Results: How responsible is the worker for work outcomes and results of other workers?
8. Consequence of Error: how serious would the result usually be if the worker made a mistake that was not readily correctable?
9. Importance of Being Exact or Accurate: How important is being very exact or highly accurate in performing this job?
10. Impact of decisions on Co-workers or Company Results: What results do your decisions usually have on other people or the image or reputation or financial resources of your employer?

We aggregate the grades in all these these dimensions into a single score using factor analysis; we denote this score by λ. Table 2 presents the dimensions and their relative importance in the definition of λ.

Table 2: Construction of λ

O*NET code	Characteristic description	Weight
abLV.1.A.1.b.3	Problem Sensitivity (Level)	0.3660
skLV.2.A.1.b	Active Listening (Level)	0.3384
skLV.2.B.1.a	Social Perceptiveness (Level)	0.3515
skLV.2.B.1.b	Coordination (Level)	0.3800
wc.4.C.1.b.1.e	Work With Work Group or Team	0.3153
wc.4.C.1.b.1.g	Coordinate or Lead Others	0.3629
wc.4.C.1.c.2	Responsibility for Outcomes and Results	0.2773
wc.4.C.3.a.1	Consequence of Error	0.2110
wc.4.C.3.b.4	Importance of Being Exact or Accurate	0.1031
wc.4.C.3.a.2.a	Impact of Decisions on Co-workers or Company Results	0.3471

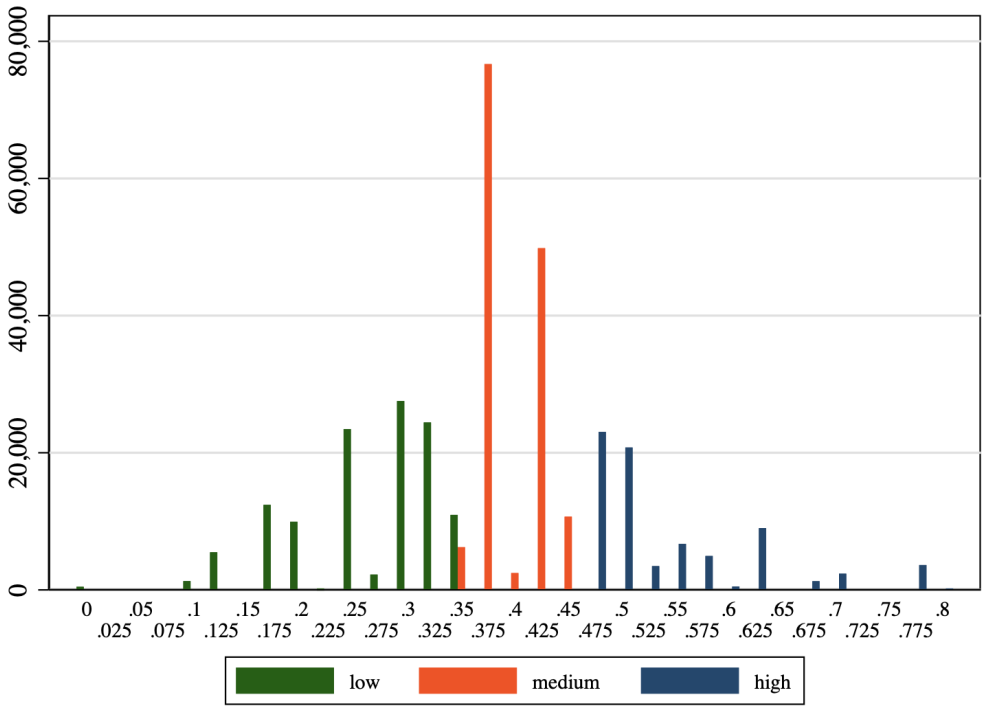
Notes: coordinates of the first eigen vector in a principal component analysis based on 9 characteristics taken from O*NET. Each characteristic has been aggregated at the SOC 4 digit occupation level using an employment weighted mean from the original O*NET-SOC occupation classification.

We normalize λ so as to lie between 0 and 1 and discretize it into three categories according to where it lies in the overall distribution of workers in low educated occupations in the UK. We refer to a high λ occupation as an occupation in the top 33% of the (unweighted) distribution of λ's, and similarly to a low λ occupation as an occupation in the bottom 33%. Examples of low educated occupations with low λ include: cleaner, bar staff, caretaker, packer, process operator; examples of low educated occupations with medium λ include: finance officer, book-keeper, plasterer, clerk, sales assistant; examples of low educated occupations with high λ include: receptionist, medical or school secretary, air transport operative, assembler.[6]

Figure 1 shows the distribution of λ's amongst workers in our sample of males aged 18-49 working in low educated occupations in private sector firms with 400 or more employees.

[6]We provide a complete list of occupations in the low, medium, and high λ categories on this webpage https://www.rachelgriffith.org/soft-skills-and-wage-progression-of.

Figure 1: Distribution of workers in low educated occupations by λ

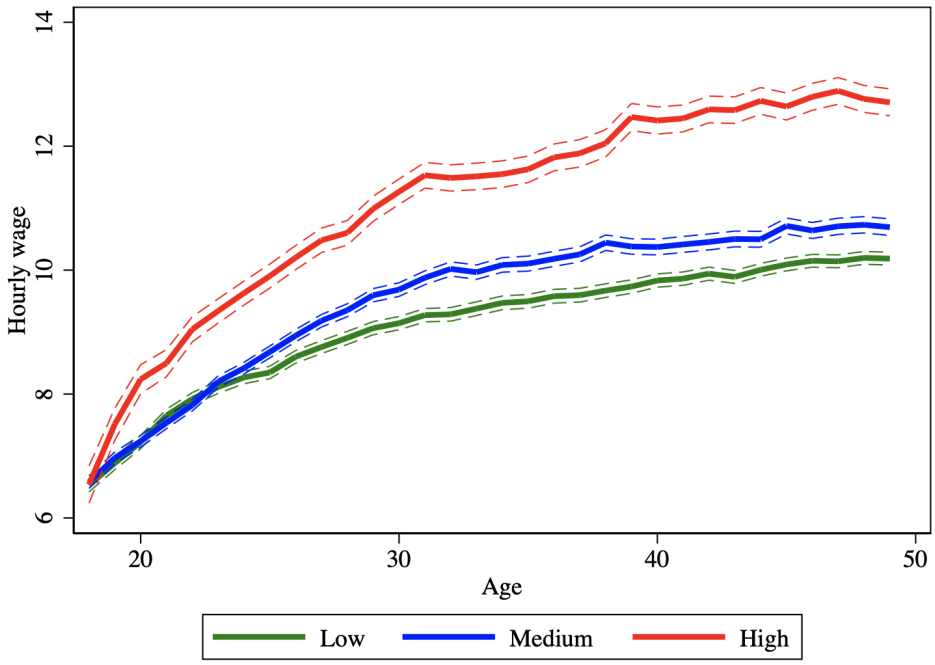


Notes: distribution of λ across male workers in low educated occupations aged 18-49 in private firms.

2.3 Evidence that high soft skill occupations are 'good jobs'

The most direct evidence that jobs in high skilled occupations represent 'good jobs' is that we observe more wage progression in workers in high λ occupations. Figure 2 plots the average wage at different ages for workers in low educated occupations in high, medium and low λ occupations respectively. Those in the higher λ occupations: (i) enjoy higher wages than workers in lower λ occupations at all age; (ii) experience sharper wage progression with age than workers in lower λ occupations. In addition to enjoying stronger wage growth, workers in high λ occupations also differ on other characteristics, shown in Table 3. In particular, they work in larger firms, are more likely to work full-time, and have longer tenure.

Figure 2: Average wage by importance of soft skills in occupation



Notes: Data from Annual Survey of Hours and Employment (ASHE) 2004-2019. Figure shows average hourly wage at each age for male workers in private sector firms in occupations with low educational requirements categorised by a measure of the importance of soft skills using O*NET data, see section 2.2.

Building on recent work by Rodrik and Stantcheva (2021), in this subsection we provide suggestive evidence to the effect that low educated workers in high λ occupations indeed perceive these as good jobs. For this purpose, we use the 2015 wave of the European Working Conditions Survey (EWCS). The EWCS is a European survey on working conditions. 43,000 European workers from 35 countries are interviewed and asked about their jobs.[7] We use data from Belgium, Denmark, Germany, Spain, France, Ireland, Italy, Luxembourg, Netherlands, Austria, Portugal, Finland, Sweden, United Kingdom, Norway, and Switzerland. Information in EWCS are reported at the 2-digit ISCO80 level. We recalculate a value of λ at this level.

EWCS participants are asked to indicate their answers to various questions on a Likert scale. We use the following questions that measure how workers view the quality of their job:

1. Does your job offer good prospects for career advancement?

2. On the whole, are you satisfied with working conditions?

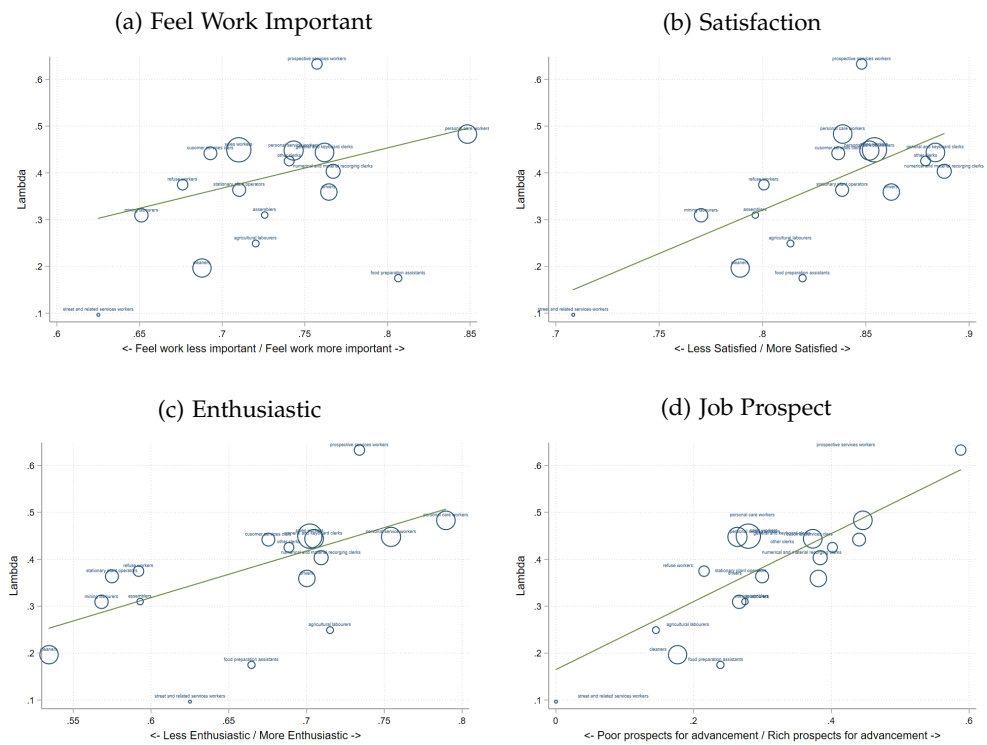
[7]This included aspects of their working life such as employment status, working time, work organisation, work-life balance, physical and psychosocial risks factors, learning and training, voice and participation health and well-being as well as earnings.

3. Are you enthusiastic about your job?

4. Do you doubt the importance of your work?

Figure 3 plots the mean answer to each of these questions (horizontal axis) against the mean λ at the two-digit ISCO80 level across occupations (each dot corresponds to a particular occupation) and weighted by employment. We see a positive correlation between the λ level of an occupation and the extent to which the occupation is perceived by workers in the EWCS survey as important (Figure 3(a)), satisfying (Figure 3(b)), enthusiastic (Figure 3(c)) and as offering good career advancement prospects (Figure 3(d)).

Figure 3: Correlation between λ and training



Notes: Correlation between λ and four questions form the European Working Condition Survey. Data from Belgium, Denmark, Germany, Spain, France, Ireland, Italy, Luxembourg, Netherlands, Austria, Portugal, Finland, Sweden, United Kingdom, Norway, and Switzerland and consider the 2015 vintage of the survey. The sizes of the bins correspond to the employment level in each ISCO 2008 2-digit occupation.

3 A model of good jobs for low educated workers

To rationalize our findings in the previous section and to guide our empirical analysis in the next sections, in this section we develop a simple model in which: (i) low skilled occupations can be ranked according to their degree of reliance on soft skills λ; (iii)

the degree of reliance on soft skills λ depends upon both, the complementarity of the task with the firm's other assets and the difficulty for the firm to find a replacement on that task. Hence workers in low skilled occupations draw bargaining power for two reasons. First, from the fact that they are more complementary to the firm's other assets. Second, from the fact that it is hard for the firm to find alternative workers in low skilled occupations with relatively high soft skills: instead, firms need time to find and/or train workers to get equal levels of soft skills. As a result, workers in low skilled occupations with higher soft skills will command a higher wage. We now proceed to formalize our argument.

3.1 Model setup

Production function

We consider a representative firm which employs an asset of quality Q,[8] which it combines with tasks, each of which is performed by a different worker in each low educated occupation. A task is a pair $(\mu, q_L) = \Gamma$, where $\mu \in [0,1]$ measures the degree of complementarity between the worker's quality and the quality of the firm's asset Q, and where q_L denotes the quality of an outside worker hired on the spot on that task. Different Γ's correspond to different λ's in our above analysis: a higher μ or a lower q_L both correspond to a task which commands a higher degree of soft skills.

More formally, if q denotes the quality of the low educated worker on task μ, then the output produced on that task is assumed to be determined by the following "partially O'Ring" production function (see Kremer, 1993 and Kremer and Maskin, 1996):

$$f(\mu, q, Q) = \mu q Q + (1 - \mu)(q + Q).$$

The value $\mu = 0$ corresponds to full substitutability between the qualities of low educated workers and of the firm's other asset. The value $\mu = 1$ corresponds to the case where these qualities are fully complementary. Let $\phi(\mu)$ denote the density distribution on μ and $\phi(\mu)$ denote the corresponding cumulative distribution function.

For each task $\Gamma = (\mu, q_L)$, the firm will optimally choose the quality level $q(\Gamma)$ it wants workers on task Γ to achieve for each Γ, and the difference $q(\Gamma) - q_L$ captures the degree of soft skill required on task Γ.

[8] This complementary asset may just boil down to the high educated employee(s) in the firm, in which case we can think of Q as the level of hard skills of these high educated employees.

The firm's total output is then taken to be the sum of the outputs on all individual tasks. More formally:

$$F(\vec{q}, Q) = \int_{\Gamma} f(\mu, q, Q)\phi(\Gamma)d\Gamma.$$

where:

$$\vec{q} = (q(\Gamma))_{\Gamma} \text{ and } \int_{\Gamma} \phi(\Gamma)d\Gamma = 1.$$

3.1.1 Wage negotiation

We follow Stole and Zwiebel (1996) and assume separate Nash bargaining of the firm with each individual worker on each occupation (or task) Γ. This negotiation leads to the equilibrium wage $w_q(\Gamma)$. In the Appendix, we show that this bargaining yields similar results as in the case where all workers in a given task bargain together. We further assume that the firm and workers on each task Γ equally share the net surplus generated on that task.

In this bargaining, the firm has the outside option of replacing a worker in the occupation on that task - this worker has quality $q(\Gamma)$ and is paid wage $w_q(\Gamma)$ - by an outside worker with reservation quality q_L and reservation wage w_L.[9] Similarly, the low educated workers have outside option $\bar{w}^L$ which is also exogenous.

The firm's ex-post net surplus from employing low educated workers with quality q on task Γ, is equal to:

$$S^F = [\mu Q + (1 - \mu)]\,(q - q_L) + w_L - w_q(\Gamma).$$

The surplus of the low educated workers on that task is equal to

$$S^{LS} = w_q(\Gamma) - \bar{w}^L.$$

Using the fact that $S^F = S^{LS}$, we get that the equilibrium wage of a low educated worker on task Γ satisfies:

$$w_q(\Gamma) = \frac{1}{2}[(\mu Q + (1 - \mu))\,(q - q_L) + w_L - \bar{w}^L].$$

The firm's total wage bill is then equal to

$$W(\vec{q}) = \int w_q(\Gamma)\phi(\Gamma)d\Gamma.$$

[9] An alternative interpretation is that absent a wage agreement the worker in the low skilled occupation chooses to underperform at quality level q_L.

Net profits and training

The firm's *ex-post* profit is equal to:

$$\tilde{\Pi}(\vec{q}) \equiv F(\vec{q}) - W(\vec{q}).$$

We assume that prior to the wage negotiation, the firm can train the worker in the low skilled occupation on each task μ, so that the expected quality of the worker moves up from q_L to some higher quality level $q(\Gamma)$ at a quadratic cost. The firm's *ex-ante* training investment will seek to maximize *ex-post* profit minus the training cost, namely:

$$\tilde{\Pi}(\vec{q}) - \int_{\Gamma} C(q(\Gamma))^2 \phi(\Gamma) d\Gamma,$$

with respect to $\vec{q} = (q(\Gamma))_{\Gamma}$. The training cost $C(q(\Gamma))^2$ can be reinterpreted as a cost of learning about the worker's underlying soft skills as as the time required for the firm to learn and thereby fully take advantage of the worker's quality $q(\Gamma)$ on task Γ.

3.2 Solving the model

To simplify the analysis we assume that the training cost (or learning cost) parameter C is independent of Γ.

3.2.1 Optimal training decision

For each task Γ we consider the firm's optimal choice of qualities $q^* = q^*(\Gamma)$.

The firm chooses q^* by solving:

$$q^* = \arg\max_{q_L \leq q} \left\{ \tilde{\Pi}(\vec{q}) - \int C(q(\Gamma))^2 \phi(\Gamma) d\Gamma \right\}$$

Pointwise maximization yields, for each Γ:

$$q^*(\Gamma) = \frac{\mu Q + 1 - \mu}{4C},$$

which naturally implies the following proposition.

Proposition 1. *$q^*(\Gamma) - q_L$, the degree of soft skill required on task Γ and which translates into an optimal training or learning cost, is increasing in μ.*

3.2.2 Equilibrium wages

Substituting for $q = q^*(\Gamma)$ in the above expression for $w_q(\Gamma)$ yields (up to a constant):

$$w^*(\Gamma) \equiv w_{q^*(\Gamma)}(\Gamma) = \frac{1}{2}[\mu Q + 1 - \mu]\left[\frac{\mu Q + 1 - \mu}{4C} - q_L\right].$$

We then immediately establish:

Proposition 2. *The equilibrium wage of a low educated worker satisfies:*

$$\frac{\partial w^*}{\partial \mu} > 0; \frac{\partial w^*}{\partial q_L} < 0.$$

3.2.3 Skill intensity

Skill intensity is captured by the quality Q of the firm's other assets which are combined with low educated labor in the production process. We have:

Proposition 3. *The equilibrium wage of a low educated worker increases faster with μ, and decreases faster with q_L, the more skill-intensive the firm is:*

$$\frac{\partial^2 w^*}{\partial \mu \partial Q} > 0; \frac{\partial^2 w^*}{\partial q_L \partial Q} < 0.$$

3.2.4 Innovativeness

There are several ways to capture innovativeness in this model. A first way is to use the fact that more innovative firms tend to be more skill-intensive and therefore have a bigger Q. A second way is to use the fact that more innovative firms tend to draw higher rents from production. More formally, we assume that:

$$f(\mu, q, Q) = M\ \left[\mu q Q + (1-\mu)(q+Q)\right],$$

where a higher M reflects a more innovative activity. Both approaches lead to the prediction that the wage premium for low educated workers should increase faster with μ and decrease faster with q_L in more innovative firms.

3.2.5 Outsourcing

Suppose that training low-educated workers up to quality level q costs time δq, and that the firm faces a time constraint at T. Then if this time constraint is binding, it

is optimal for the firm to fix $q^*(\Gamma) = q_L$ for some values of Γ, which we interpret as outsourcing the corresponding occupation. Formally, the firm chooses q^* by solving:

$$q^* = \arg\max_q \left\{ \tilde{\Pi}(\vec{q}) - \int C(q(\Gamma))^2 \phi(\Gamma) d\Gamma \right\}$$

subject to the time constraint

$$\int_\Gamma \delta q(\Gamma) \phi(\Gamma) d\Gamma \leq T.$$

In the Appendix we prove:

Proposition 4. *In other words, for a given q_L, there exists a cut-off value $\overline{\mu}$ such that $q(\Gamma) = q_L$ for $\mu \leq \overline{\mu}$. Moreover we have:*

$$\frac{d\overline{\mu}}{dQ} > 0.$$

In other words, for a given q_L, all task with a degree of complementarity $\mu \leq \overline{\mu}$ are being outsourced by the firm, and the higher Q the higher the fraction of tasks that will be outsourced.

3.3 Main predictions

In the next section, we confront the following predictions of the model with the data.

Fact 1: The wage premium for low educated workers is higher in "higher λ" jobs, i.e. in jobs with higher μ and/or lower q_L.
Fact 2: Low educated workers on higher λ jobs benefit from more training or more learning by the firm.
Fact 3: The wage premium for low educated workers increases faster with λ in more skill-intensive or more innovative firms.
Fact 4: More innovative firms outsource more tasks.

4 An empirical model of wage progression

In this section we lay out our empirical approach to investigate the main predictions of the model. From our analysis in the previous section we expect workers in low educated occupations where soft skills are important to experience higher wages and

stronger wage progression with tenure than workers in occupations where soft skills matter less.

We first describe our empirical strategy for characterising tenure-wage profiles for workers in soft skill occupations using panel data on the matched employee-employer panel. Our aim here is to check whether the steeper profiles for lower educated workers in soft skill occupations that we saw in Figure 2 hold up once we allow for other differences across firms and workers. We then look in detail at the role of unobserved worker heterogeneity, productivity shocks and firm characteristics. Finally, before taking the model to data, we discuss the interpretation of the estimated parameters.

4.1 A panel data framework

We index workers by i, occupations by j, firms by f and time (years) by t. We denote the wage of a worker i in occupation j in firm f at time t as w_{ijft}. Each worker will have a different ability to perform the soft skill intensive tasks that comprise high λ occupations. Let κ_i be the *potential* level of soft skills of worker i, and λ_j be a binary indicator selecting high λ occupations. Other skills, like cognitive skills, will matter too and we will account for these in the empirical application, but they will typically be qualification-based and more verifiable.

A main focus of our empirical analysis, is on how the returns for workers in high λ occupations evolve with the worker's tenure in the firm. We define

$$\phi_f(\kappa_i, \lambda_j, T_{if}) \tag{1}$$

as the fraction of the joint surplus recovered by worker i with tenure T_{if} in firm f working in a λ_j type occupation. This surplus will depend on the extent to which the occupation relies on soft skills λ_j and on the value of the worker's underlying ability to perform soft skill tasks κ_i. As soft skills are hard to verify, it will also depend on the firm's knowledge of the level of soft skills which will increase with the worker's tenure T_{if} in firm f. For workers in occupations that make use of soft skills, firms will wish to invest in learning about those skills and/or in enhancing them through training. As we argue below, this is a key reason why firm tenure and training are important in determining the surplus in soft skill occupations. Our model therefore predicts a complementarity between κ_i, λ_j, and T_{if}.

The panel data (log) wage equation for worker i in period t is the sum of the surplus $\phi_f(\kappa_i, \lambda_j, T_{ift})$ and the standard arguments that enter a panel data wage spec-

ification, resulting in

$$\ln w_{ijft} = \phi_f(\kappa_i, \lambda_j, T_{ift}) + g(A_{it}, FT_{i0}, S_f, T_{ift}) + \gamma_i + \eta_t + e_{ijft} \tag{2}$$

where $g(A_{it}, FT_{i0}, S_f, T_{ift})$ is a flexible function of worker age A, a binary indicator FT equal to one if the worker is hired as a full time employee, the size of the firm S, and the worker's tenure T in the firm. In our empirical analysis $g(.)$ will also contain measures of cognitive skills. The γ_i, η_t and e_{ijft} components in (2) represent unobserved individual heterogeneity, time effects and transitory shocks, respectively. Firm size and full-time will be measured at outset of a job to avoid temporal endogeneity through correlation with e_{ijft}. Although λ_j, T_{ift}, A_{it}, FT_{i0}, and S_f are observable to the econometrician, the κ_i, γ_i, η_t and e_{ijft} terms are not. To make further progress with the empirical specification and choice of estimator we need to look closer at the surplus $\phi_f(\kappa_i, \lambda_j, T_{ift})$ and examine the role of the unobservable components in estimation.

When a worker first joins a firm we expect the surplus term to be small and to grow with tenure. It takes time for the firm to discover and utilize the worker's underlying soft skill ability κ_i in a high λ occupation.[10] The firm may also enhance the worker's soft skills over time through training. To focus the discussion we rewrite the surplus term:

$$\phi(\kappa_i, \lambda_j, T_{if}) = \theta(\kappa_i, T_{if})\lambda_j \tag{3}$$

where $\theta(\kappa_i, T_{if})$ measures the level of soft skills κ_i for worker i that is revealed at tenure T_{if}. Note that, for the sake of exposition, we have assumed that differences across firms are captured by λ_j and T_{if} only. Below we will allow the surplus function to also differ across firms with different levels of R&D, with different proportions of high educated workers, and with the amount of training the worker receives in the firm. We will also allow the returns to vary over time.

At the start of a new job in firm f, worker i's tenure equals zero $T_{if} = 0$. At this point the revealed level of κ_i will be small, possibly zero. As tenure increases the firm will take better advantage of the worker's ability and thus we write:

$$\theta(\kappa_i, T_{ift}) = \alpha_1\kappa_i + \alpha_2 k(T_{if})\kappa_i \tag{4}$$

where $\alpha_1\kappa_i$ is the initial level of observed soft skills and where $k(T)$ is a function

[10] Think for example of "reliability": reliability is a soft skill which is largely unverifiable across firms. Typically, it takes time for a firm to discover how reliable a worker is once the worker has been hired on a high λ occupation. It is thus usually through small steps of experimenting with the worker that the firm will assess the extent to which the worker can be relied upon to operate in the occupation autonomously.

that lies in the unit interval, is increasing in T and is equal to zero when $T_{if} = 0$. Combining (3) and (4), gives

$$\phi(\kappa_i, \lambda_j, T_{if}) = \alpha_1 \kappa_i \lambda_j + \alpha_2 \kappa_i k(T_{if}) \lambda_j. \tag{5}$$

This generates our log wage specification:

$$\ln w_{ijft} = \alpha_1 \kappa_i \lambda_j + \alpha_2 \kappa_i k(T_{if}) \lambda_j + g(A_{it}, FT_{if}, S_f, T_{if}) + \gamma_i + \eta_t + e_{ijft}. \tag{6}$$

4.2 Worker heterogeneity

Note that wage equation (6) contains two dimensions of unobserved individual worker heterogeneity, κ_i and γ_i, as well as a transitory shock to wages e_{ijft}. The properties of these error components will have important implications for the estimated parameters in equation (6). We consider the implications of each separately.

Turning first to κ_i, the standard least squares panel data estimator of the coefficient on λ_j will identify $\alpha_1 \mathbb{E}\left(\kappa_i | T = 0, i \in j\right)$ and the estimator of the coefficient on $k(T_{if})\lambda_j$ will identify $\alpha_2 \mathbb{E}\left(\kappa_i | T = T_{if}, i \in j\right)$. Note that $\mathbb{E}\left(\kappa_i | T = T_{if}, i \in j\right)$ measures the average value of soft skills used in occupation j for those workers still in firm f and occupation j with tenure $T = T_{if}$. This term could therefore enhance the increase in wage with tenure that is already implicit in $k(T)$ if firms retain workers in occupation j as they learn more about their ability κ_i. Although it would be informative to separate out this dynamic selection on κ_i from the pure tenure effect $k(T)$, an estimate of the combined term $\mathbb{E}\left(\kappa_i | T, i \in j\right) k(T)$ as a function of T is sufficient for our purposes. It tells us the degree to which workers with higher soft skills, working in occupations where these skills are important (high λ_j occupations), experience higher wage progression and longer tenures.

The second dimension of heterogeneity in (6), the additive heterogeneity term γ_i, is a standard individual heterogeneity term in a log wage equation reflecting additive unobserved ability. This term can also induce a bias in our estimated tenure coefficients if individuals with higher γ_i are more likely to be retained in the firm. It will also likely be correlated with κ_i. We can interpret γ_i as capturing initial underlying levels of ability, unobserved to the econometrician, but observed in the market. To avoid bias in the estimated parameters we replace γ_i by a measure of the initial wage. We assume this captures the level of skills of the worker at *entry*. That is, we assume that T_{if} is orthogonal to γ_i conditional on the initial wage. Conveniently our data contains pre-sample observations on wages. We use the first wage that we observe. This pre-sample wage will reflect the worker's initial skill level and is not influenced

by the evolution of skills during the observation period.[11]

The final unobservable determinant of individual worker wages in (6) is the idiosyncratic shock e_{ijft}. This can induce a bias in our estimate of the tenure effects in the wage equation if these transitory shocks are correlated with worker exits from the firm. Examining the wage equation (6) we notice that the key tenure term of interest, $\alpha_1\kappa_i\lambda_j + \alpha_2\kappa_i k(T_{if})\lambda_j$, measures the impact of tenure for those workers in high lambda occupations *relative* to the impact of tenure on those workers in lower lambda occupations which is captured in the term $g(A_{it}, FT_{if}, S_f, T_{if})$. Thus, provided the bias from selective exit on e_{ijft} is the same across lambda occupations, the estimates of *the relative impact* will remain unbiased.

5 The impact of soft skills on wage progression

Before turning to the estimates of the empirical model describe above, Table 3 shows means of the main variables, for the whole sample, and separately for workers in high and low λ occupations. Workers in high λ occupation have on average higher wages, are more likely to work full time, and have longer average tenure.

Table 3: Descriptive statistics

	(1)	(2)	(3)
		importance of soft skills (λ)	
	All	low	high
Wage (£ hourly)	9.73	9.23	11.61
Full-time (%)	74.6	70.7	89.5
Age	32	32	33
Tenure	5.21	4.96	6.17
Firm size (number employees)	37,917	43,121	18,511
Number in our sample:			
workers	63,407	53,035	17,097
firms	5,966	5,202	3,411
firm-years	34,032	28,656	14,992
worker-firm-years	212,428	167,506	44,922

Notes: Data from Annual Survey of Hours and Employment (ASHE) 2004-2019 for male workers in private sector firms with 400+ employees in occupations with low educational requirements categorised by a measure of the importance of soft skills using O*NET data, see section 2.2.

[11]This is similar to an idea developed in Blundell et al. (1999) and Blundell et al. (2002).

5.1 Basic regression results

Table 4 shows the estimates of the parameters in equation (6). We see a significantly positive effect of working in a high λ_j occupation on wages, and this effect increases with the worker's tenure, particularly over the first five years of tenure. This reflects the fact that soft skills take time to be valued by the firm. Note that our measure of cognitive skills is included throughout. These results are robust to including time effects that vary within travel to work areas (TTW), these are areas defined by the Office of National Statistics as those in which most of the residents work (local labour markets), and to including individual worker effects (column (3)).

The estimates in column (4) suggest that workers in high λ occupations receive around 4.6% higher wages on average, and they achieve faster pay progression in their first five years within the firm at around 1.3% and at around 0.4% in the years after that.

Table 4: Wage progression in high λ occupations

Dependent variable: $\log(w_{ijkft})$

	(1)	(2)	(3)	(4)
High lambda (λ)	0.1157***	0.0589***	0.0278***	0.0461***
	(0.0032)	(0.0041)	(0.0047)	(0.0033)
High lambda $\times$ tenure		0.0081***	0.0021***	0.0038***
		(0.0004)	(0.0004)	(0.0003)
High lambda $\times$ tenure 0-5 years		0.0070***	0.0055***	0.0086***
		(0.0012)	(0.0008)	(0.0010)
High cognitive	0.0762***	0.0745***	0.0044***	0.0450***
	(0.0033)	(0.0033)	(0.0029)	(0.0023)
Initial firm size (S)	-0.0056***	-0.0053***	-0.0004	-0.0023***
	(0.0007)	(0.0007)	(0.0008)	(0.0005)
Full-time (FT)	0.1144***	0.1173***	-0.0103***	0.0966***
	(0.0031)	(0.0030)	(0.0028)	(0.0024)
Age (A)	0.0219***	0.0219***	0.0104***	0.0185***
	(0.0003)	(0.0003)	(0.0015)	(0.0003)
Age squared (A^2)	-0.00047***	-0.00047***	-0.00066***	-0.00046***
	(0.00001)	(0.00001)	(0.00002)	(0.00001)
Tenure (T)	0.0244***	0.0234***	0.0086***	0.0230***
	(0.0005)	(0.0005)	(0.0005)	(0.0004)
Tenure squared	-0.00042***	-0.00047***	-0.00021***	-0.00056***
	(0.00002)	(0.00002)	(0.00002)	(0.00001)
Initial wage				0.0463***
				(0.0010)
TTW-Year	✓	✓		✓
Worker effects			✓	
Year effects			✓	
R^2	0.285	0.288	0.340	0.474
Observations	212,389	212,389	212,389	212,389

*Notes: Data is from ASHE 2004-2019 and BERD 2001-2019. Sample is male workers aged 18-49 in low-educated occupations in private sector firms with 400+ employees. Numbers are coefficients with robust standard errors in parentheses. TTW: travel to work area. Stars indicate * p<0.1, ** p<0.05, *** p<0.01.*

5.2 Allowing the surplus to vary across firms

So far we have assumed that the returns to workers' soft skills are homogeneous across firms (see equation (5)). However, our analysis in Section 3 also predicts that wages and wage progression for low educated workers in higher λ occupations should increase with the *quality* of the firm's other assets Q. In particular this quality may refer to the hard skills of high educated workers in the firm, or to the innovative-

ness of the firm.

To explore this we allow (5) to vary with whether the firm does R&D, the quality of workers in high educated occupations (proxied by their mean wage), and the quantity of workers in high educated occupations (share of total workforce). In particular we add the following interaction terms to the log wage equation (6):

$$\alpha_3 R_{ft}\lambda_j + \alpha_4 k(T_{ift}) R_{ft}\lambda_j, \tag{7}$$

where R_{ft} refers either to the quality of the workers in high educated occupations in firm f or to whether firm f does R&D. We also include a term in the level of R_{ft}.

Table 5 shows the estimates of the parameters. Columns (2) - (3) include the quality of the workers in high educated occupations while columns (4) - (5) refers to whether firm f does R&D. Column (1) is a repeat of column (4) from Table 4 for ease of comparison. Columns (2) and (4) include individual worker effects, while columns (3) and (5) use the initial wage to control for unobserved worker characteristics. In column (6) we repeat the specification in column (1) using the smaller sample of data for which data on the entire workforce from WERS is available. In column (7) we include the interaction with the share of the entire workforce that is in high skilled occupations.

The results show that pay progression is higher for workers in high λ occupations in firms that are more innovative, measured either by whether the firm does R&D, has skilled workers who are higher paid, or where a higher share of its workforce are skilled workers. Based on our model of production and bargaining within the firm, we interpret this to reflects the complementarity between high levels of soft skills among low-educated workers and the firms other assets.

The estimates suggest that pay progression in the first five years in a job is higher for workers in high *lambda* occupations if they work in a firm that does R&D (columns 2 and 3), if the skilled workers in the firm are higher quality (columns 4 and 5), and if the there is a higher share of skilled workers in the firm (column 7).

Table 5: Wage progression of workers in high λ occupations in higher tech firms

Dependent variable: $\log(w_{ijkft})$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High lambda	0.0461***	0.0340***	0.0499***	0.0275***	0.0447***	0.0485***	0.0699***
	(0.0033)	(0.0048)	(0.0033)	(0.0048)	(0.0035)	(0.0088)	(0.0083)
High lambda × tenure	0.0038***	0.0024***	0.0043***	0.0022***	0.0036***	0.0046***	0.00072***
	(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0004)	(0.0006)	(0.0006)
High lambda × tenure 0-5 years	0.0086***	0.0045***	0.0070***	0.0054***	0.0085***	0.0044**	0.0072***
	(0.0010)	(0.0009)	(0.0011)	(0.0008)	(0.0010)	(0.0022)	(0.0024)
High lambda × tenure 0-5 years × R&Dfirm		0.0055**	0.0096***				
		(0.0022)	(0.0024)				
High lambda × tenure 0-5 years × wage high educated				0.0001	0.0004***		
				(0.0001)	(0.0001)		
High lambda × tenure 0-5 years × share high educated							0.0451***
							(0.0107)
High lambda × R&D firm		-0.0381***	-0.0226***				
		(0.0086)	(0.0057)				
High lambda × wage high educated				0.0014***	0.0030***		
				(0.0004)	(0.0004)		
High lambda × share high educated							0.1678***
							(0.0365)
R&D firms		0.0705***	0.0768***				
		(0.0053)	(0.0043)				
R&D firms × tenure		0.0006	-0.0003				
		(0.0005)	(0.0003)				
Wage high educated				0.0002	0.0007***		
				(0.0002)	(0.0001)		
Wage high educated × tenure				0.00002	0.00007***		
				(0.00003)	(0.00002)		
Share high educated						0.1672***	
						(0.0201)	
Share high educated × tenure						0.0051***	
						(0.0016)	
High cognitive	0.0450***	0.0046	0.0466***	0.0033	0.0406***	0.0458***	0.0402***
	(0.0023)	(0.0028)	(0.0023)	(0.0030)	(0.0021)	(0.0043)	(0.0037)
Initial wage	0.0463***		0.0458***		0.0459***	0.0449***	0.0433***
	(0.0010)		(0.0010)		(0.0010)	(0.0016)	(0.0016)
TTW-Year	✓		✓		✓	✓	✓
Worker effects		✓		✓			
Year effects		✓		✓			
R^2	0.474	0.343	0.479	0.340	0.480	0.497	0.511
Observations	212,389	212,389	212,389	198,446	198,446	53,592	53,592

Notes: Data is from ASHE 2004-2019, BERD 2001-2019 and WERS 2011. Sample is male workers aged 18-49 in low educated occupations in private sector firms with 400+ employees. Numbers are coefficients with robust standard errors in parentheses. TTW: Travel To Work area. All columns include firm size, full-time indicator, age, age-squared, tenure and tenure-squared. Numbers are coefficients with robust standard errors in parentheses. Stars indicate * p<0.1, ** p<0.05, *** p<0.01.

5.3 Firm investment in training

A further implication of our model is that firms will invest more in training workers in occupations where soft skills are more important (high λ occupations).

We do not have information on training in ASHE. Instead, we use data from the LFS and EWCS. The LFS is a household survey on the UK which provides details on employment conditions. In particular, it provides detailed information on individuals' education and skills as well as some information on training.

Table 6 reports training of individual UK workers. The table shows that workers in occupations with λ's above the median: are more likely to have been offered training by their employer (row 1), are more likely to be in training (row 2), and are more likely to have received training during work (row 3).

Table 6: Training and soft skills (λ)

	λ of occupation		
	Below median	Above median	Diff
Whether employer has offered training	13.9	15.7	1.7***
	(0.17)	(0.18)	(0.24)
In education or training (of any kind)	9.5	10.9	1.5***
	(0.12)	(0.13)	(0.18)
Training during work	4.9	5.8	0.9***
	(0.29)	(0.31)	(0.42)

Notes: Data are from the Labour Force Survey (LFS). Sample is male workers aged 18-49 in low skilled occupations in private firms with 400+ employees. Numbers are coefficients with negative numbers underneath standard errors. Stars indicate * p<0.1, ** p<0.05, *** p<0.01.

Figure 4 shows mean responses by workers in EWCS to the question "have you undergone any training paid for or provided by your employer?" (horizontal axis) across occupations ranked according to their λ's (vertical axis). The figure shows that workers in higher λ occupations are more likely to receive training than workers in lower λ occupations.

Figure 4: Correlation between λ and training

Notes: Data are from European Working Condition Survey (EWCS). Mean response to the question "have you undergone any training paid for or provided by your employer?" against the value of λ. Data from Belgium, Denmark, Germany, Spain, France, Ireland, Italy, Luxembourg, Netherlands, Austria, Portugal, Finland, Sweden, United Kingdom, Norway, and Switzerland and consider the 2015 vintage of the survey. The size of the bins correspond to the employment level in each ISCO 2008 2-digit occupation.

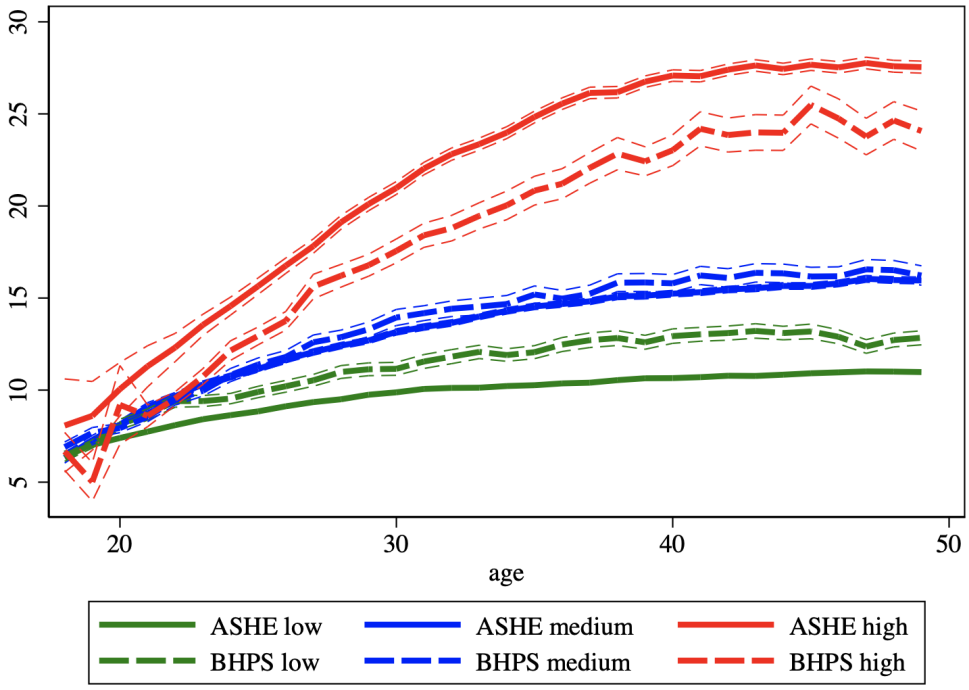
6 Robustness and additional results

In this section we test the robustness of our results to various extensions of our baseline identification strategy.

6.1 Measurement of education level of occupations

Our categorisation of occupations relied on the qualification list used by the immigration regulation authority. Is this categorization accurate? How does it compare to the qualifications that workers actually have? In this subsection we use data from the British Household Panel Survey (BHPS) and Understanding Society (USoc) to show that the pattern of wage progression remains unchanged when we use actual qualifications in these two datasets instead of the RFQ information; see Figure 5 which reports the average hourly wage at each age.

Figure 5: Wage progression by occupation (ASHE) and qualification (BHPS)



Notes: Data are from ASHE 2004-2019 for men working in the private sector and from BHPS for men working in the private and public sector.

6.2 Measurement of soft skills

One potential concern regarding our measure of soft skills (λ) is that it might be correlated with the actual education level of the worker. In particular, this could be a

problem if low educated occupation workers in high λ occupations are more educated than those in low λ occupations. We check that this is not the case using the Labour Force Survey in which we can look at the average level of education by λ. Results are presented in Table 7 and show that if anything, workers in occupations that are below the median in terms of λ are slightly more educated than workers in other low educated occupations.

Table 7: Average level of education for above and below median λ

Workers in low educated occupations

	Lambda		
	Below median	Above median	diff
Age left education	17.8	17.7	-0.09***
	(0.02)	(0.02)	(0.03)
Has higher education degree	12.9	11.9	-1.0***
	(0.14)	(0.14)	(0.20)
N	55,546	52,818	109,364

Notes: Authors' calculations using LFS, 2011-2016, males 18-49 in work

6.3 Cognitive skills and non-routine jobs

We showed above that workers without formal qualifications get positive and significant returns to soft skills. In all of our regressions we controlled for the level of cognitive skills required in the occupation, measured analogously to the way we measure soft skills. In this subsection we consider how pay progression might be affected by cognitive skills and compare that to the way pay progression varies with soft skills.

Table 8 shows estimates of the coefficients for equation (2), as in Table 4; the first column repeats to final column in that table. In subsequent columns we allow pay progression to also vary with cognitive skills. Unsurprisingly, there are positive returns to cognitive skills, and these increase with tenure. The magnitude of the returns to cognitive skills is similar to the magnitude of returns to soft skills.

One concern with our λ measure is that it may amount to little more than the extent to which an occupation is routine versus non-routine. In Table 8 we include our measure of λ on top of the routine/non-routine job indicator that Acemoglu and Autor (2011) have shown to (also) be correlated with the worker's wage. We see that the premium from working in a high lambda occupation for workers in occupation

with low educational requirements remains positive and significant, even after controlling for the routine versus non-routine nature of the occupation and allowing these to affect pay progression.

Table 8: Cognitive skills and occupation characteristics from Acemoglu and Autor (2011)

Dependent variable: $\log(w_{ijkft})$

	(1)	(2)	(3)	(4)	(5)
High lambda	0.0644***	0.0551***	0.0737***	0.0513***	0.0593***
	(0.0034)	(0.0033)	(0.0035)	(0.0035)	(0.0038)
High lambda × tenure	0.0041***	0.0030***	0.0043***	0.0047***	0.0050***
	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0004)
High lambda × tenure 0-5 years	0.0085***	0.0055***	0.0077***	0.0050***	0.0044***
	(0.0010)	(0.0012)	(0.0011)	(0.0012)	(0.0012)
High cognitive		0.0309***			
		(0.0030)			
High cognitive × tenure		0.0014***			
		(0.0003)			
High cognitive × tenure 0-5 years		0.0045***			
		(0.0009)			
Non-routine cognitive analytical			-0.0209***		
			(0.0027)		
Non-routine cognitive analytical × tenure			-0.0209***		-0.0137***
			(0.0027)		(0.0028)
Non-routine cognitive analytical × tenure 0-5 years			0.0012		0.0009
			(0.0007)		(0.0007)
Routine cognitive				0.0307***	0.0269***
				(0.0030)	(0.0031)
Routine cognitive × tenure				-0.0013***	-0.0013***
				(0.0003)	(0.0003)
Routine cognitive × tenure 0-5 years				0.0055***	0.0054***
				(0.0009)	(0.0009)
Initial wage	0.0468***	0.0463***	0.0467***	0.0468***	0.0467***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
R^2					
Observations	212,389	212,389	212,389	212,389	

Notes: Data is from ASHE 2004-2019. Sample is male workers aged 18-49 in low educated occupations in private sector firms with 400+ employees. Numbers are coefficients with robust standard errors in parentheses. All columns include Travel To Work (TTW) times year effects, firm size, full-time indicator, age, age-squared, tenure and tenure-squared. Numbers are coefficients with robust standard errors in parentheses. Stars indicate * p<0.1, ** p<0.05, *** p<0.01.

6.4 Allowing surplus function to vary over time

One potential threat to our identification is that the interaction between λ and tenure could capture the occupation specific wage dynamics that are correlated with λ. For example, if demand for workers in occupations that require soft skills is increasing, this might be driving relative increases in wages over time in these occupations, which could be captured in the interaction with tenure, which also increases over time. In Table 9, we add an interaction between λ and time effects. Column (1) repeats the results in the final column in Table 4. In column (2) we include a time trend interacted

with the high λ indicator. In column (3) we include year dummies interacted with the high λ indicator, and in the column (4) we include 2-digit occupation - year effects. In all of these cases the result that pay progression is higher in high λ occupations holds up.

Table 9: Allowing the surplus function to vary over time

Dependent variable: $\log(w_{ijkft})$	(1)	(2)	(3)	(4)
High lambda	0.0461***	0.0461***	0.0288***	0.0461***
	(0.0033)	(0.0048)	(0.0058)	(0.0033)
High lambda × tenure	0.0038***	0.0038***	0.0040***	0.0038***
	(0.0003)	(0.0003)	(0.0003)	(0.0034)
High lambda × tenure 0-5 years	0.0086***	0.0086***	0.0084***	0.0086***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
High lambda × time trend		0.00001		
		(0.0005)		
High lambda ×2005			0.0084	
			(0.0075)	
High lambda ×2006			0.0011	
			(0.0085)	
High lambda ×2007			0.0024***	
			(0.0085)	
High lambda ×2008			0.0230***	
			(0.0079)	
High lambda ×2009			0.0433***	
			(0.0078)	
High lambda ×2010			0.0282***	
			(0.0076)	
High lambda ×2011			0.0378***	
			(0.0082)	
High lambda ×2012			0.0317***	
			(0.0095)	
High lambda ×2013			0.0190	
			(0.0139)	
High lambda ×2014			-0.0011	
			(0.0098)	
High lambda ×2015			-0.0023	
			(0.0108)	
High lambda ×2016			-0.0054	
			(0.0111)	
High lambda ×2017			0.0217**	
			(0.0103)	
High lambda ×2018			0.0259**	
			(0.0111)	
High lambda ×2019			0.0202	
			(0.0123)	
High cognitive	0.050***	0.050***	0.0449***	0.00450***
	(0.0023)	(0.0023)	(0.0023)	(0.0023)
Initial wage	0.0463***	0.0463***	0.0463***	0.0463***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
TTW-Year	✓	✓	✓	✓
2-digit occupation-Year				✓
R^2	0.474	0.474	0.474	0.474
Observations	212,389	212,389	212,389	212,389

Notes: Data is from ASHE 2004-2019. Sample is male workers aged 18-49 in low educated occupations in private sector firms with 400+ workers. Numbers are coefficients with robust standard errors in parentheses. All columns include Travel To Work (TTW) times year effects, firm size, full-time indicator, age, age-squared, tenure and tenure-squared. Column (4) includes 2-digit occupation times year effects. Numbers are coefficients with robust standard errors in parentheses. Stars indicate * p<0.1, ** p<0.05, *** p<0.01.

6.5 Alternative samples of workers

Here we show that our main results are robust to considering different samples of workers. In column (1) we repeat the results in the final column in Table 4, which uses a sample of male workers aged 18-49 in low educated occupations in private sector firms with 400+ employees. In column (2) we include workers in smaller firms. Column (3) includes male and female workers. Column (4) includes only female workers. Column (5) includes private and public sector firms for male workers only. Column (6) includes male workers of all ages. Our main result, that pay progression is higher in high λ occupations, holds for all of these samples, though with different magnitudes.

Table 10: Wage premium for working in high λ and tenure

Dependent variable: $\log(w_{ijkft})$

	(1)	(2)	(3)	(4)	(5)	(6)
High lambda (λ)	0.0461***	0.0368***	0.0560***	0.0687***	0.0440***	0.0495***
	(0.0033)	(0.0024)	(0.0016)	(0.0021)	(0.0023)	(0.0019)
High lambda × tenure	0.0038***	0.0034***	0.0024***	0.0019***	0.0031***	0.0018***
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0001)
High lambda × tenure 0-5 years	0.0086***	0.0085***	0.0059***	0.0039***	0.0088***	0.0069***
	(0.0010)	(0.0008)	(0.0005)	(0.0007)	(0.0007)	(0.0006)
High cognitive	0.0450***	0.0576***	0.0468***	0.0410***	0.0561***	0.0597***
	(0.0023)	(0.0020)	(0.0019)	(0.0021)	(0.0020)	(0.0021)
Firm size (S)	-0.0023***	0.0041***	0.0012***	-0.0011***	0.0064***	0.0057***
	(0.0005)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Full-time (FT)	0.0966***	0.0981***	0.0950***	0.0855***	0.1009***	0.1099***
	(0.0024)	(0.0024)	(0.0021)	(0.0021)	(0.0026)	(0.0022)
Age (A)	0.0185***	0.0205***	0.0187***	0.0174***	0.0210***	0.0145***
	(0.0003)	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0002)
Age squared	-0.0005***	-0.0005***	-0.0005***	-0.0005***	-0.0005***	-0.0003***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	
Tenure (T)	0.0230***	0.0216***	0.0199***	0.0181***	0.0215***	0.0182***
	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0002)
Tenure squared	-0.0006***	-0.0005***	-0.0005***	-0.0004***	-0.0006***	-0.0003***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	
Male			0.0584***			
			(0.0018)			
Initial wage	0.0463***	0.0453***	0.0436***	0.0399***	0.0452***	0.0435***
	(0.0010)	(0.0009)	(0.0009)	(0.0012)	(0.0009)	(0.0008)
R^2	0.474	0.444	0.404	0.318	0.449	0.445
Observations	212,389	339,803	650,890	311,087	405,047	547,649

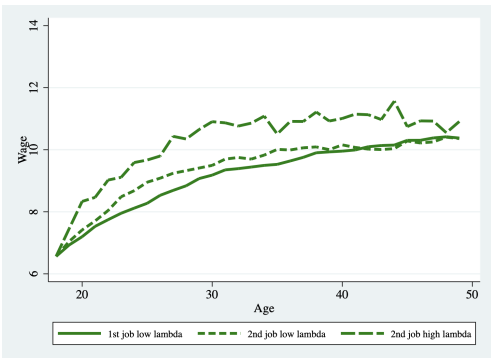
Notes: Data is from ASHE 2004-2019. Numbers are coefficients with robust standard errors in parentheses. All columns include Travel To Work (TTW) times year effects. Numbers are coefficients with robust standard errors in parentheses. Stars indicate * p<0.1, ** p<0.05, *** p<0.01. Samples include in col (1) male workers aged 18-49 in low educated occupations in private sector firms with 400+ employees, col (2) male workers aged 18-49 in low educated occupations in private firms, col (3) male and female workers aged 18-49 in low educated occupations in private firms, col (4) female workers aged 18-49 in low educated occupations in private firms, col (5) male workers aged 18-49 in low educated occupations in all firms (private and public), col (6) male workers of all ages in low educated occupations in all firms (private and public).

6.6 Occupation and job mobility

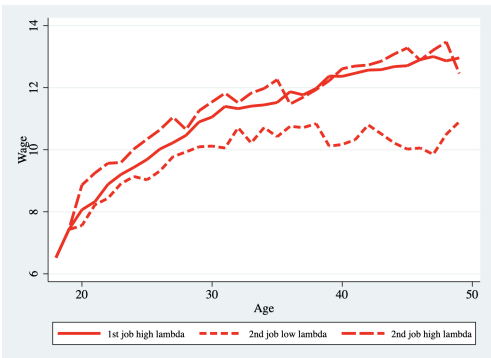
Figure 6 shows wage progression for workers in their first and second jobs. Figure (a) plots wages for workers whose first job is in an occupation where soft skills are relatively less important (low λ), while Figure (b) plots wages for workers whose first job is an occupation where soft skills are relatively more important (high λ). The solid lines are wages for workers in their first job. The short dashed line is wages in their second job for workers who move to a low λ occupation, the long dashed line for those whose second job is in a high λ occupation.

Figure 6

(a) 1st job low λ



(b) 1st job high λ



Notes: Data is from ASHE 2004-2019. The figure uses wages for males working in private sector firms in their first and second job (observed in that period). We take the mean wage for six groups: (solid green line) workers in their first job, where that job is in a low λ occupation; (dashed green line) workers in their second job, where their first and second jobs were both in low λ occupations; (long dashed green line) workers in their second job, where the first job was in a low λ occupation and the second job is in a high λ occupation; (solid red line) workers in their first job, where that job is in a high λ occupation; (dashed red line) workers in their second job, where their first job was in a high λ occupation and the second job is in a low *lambda* occupation; (long dashed red line) workers in their second job, where both the first and second jobs were in a high λ occupation.

In Table 11 we repeat the analysis presented in Table 4 using only data on a workers' first job. Again this supports our finding that pay progression is higher in high λ occupations.

Table 11: Wage premium for working in high λ and tenure, using first job only

Dependent variable: $\log(w_{ijkft})$

	(1)	(2)	(3)	(4)
High lambda (λ)	0.1342***	0.0820***	-0.0435***	0.0534
	(0.0031)	(0.0048)	(0.0070)	(0.0036)
High lambda × tenure		0.0070***	0.0048***	0.0025***
		(0.0004)	(0.0005)	(0.0003)
High lambda × tenure 0-5 years		0.0055***	0.0047***	0.0078***
		(0.0015)	(0.0009)	(0.0012)
High cognitive	0.0764***	0.0748***	0.0117***	0.0358***
	(0.0033)	(0.0033)	(0.0034)	(0.0020)
Firm size (S)	-0.0057***	-0.0054***		-0.0017***
	(0.0008)	(0.0008)		(0.0005)
Full-time (FT)	0.1075***	0.1105***	-0.0645***	0.0827***
	(0.0034)	(0.0033)	(0.0035)	(0.0022)
Age (A)	0.0216***	0.0217***	0.0041**	0.0154***
	(0.0004)	(0.0004)	(0.0017)	(0.0003)
Age squared	-0.0005***	-0.0005***	-0.0004***	-0.0004***
	(0.00001)	(0.00001)	(0.00002)	(0.00001)
Tenure (T)	0.0257***	0.0248***	0.0067***	0.0253***
	(0.0005)	(0.0005)	(0.0009)	(0.0004)
Tenure squared	-0.0005***	-0.0005***	-0.0003***	-0.0006***
	(0.00002)	(0.00002)	(0.00003)	(0.00001)
Initial wage				0.0528***
				(0.0011)
TTW-Year	✓	✓		✓
Worker effects			✓	
Year effects			✓	
R^2	0.316	0.318	0.292	0.560
Observations	158,809	158,809	158,809	158,809

Notes: Data is from ASHE 2004-2019. Sample is male workers aged 18-49 in low educated occupations in private sector firms with 400+ employees in their first job in our data. Numbers are coefficients with robust standard errors in parentheses. Travel To Work (TTW) times year, or TTW time 2-digit occupation times year are included as indicated. Stars indicate * p<0.1, ** p<0.05, *** p<0.01.

6.7 Outsourcing

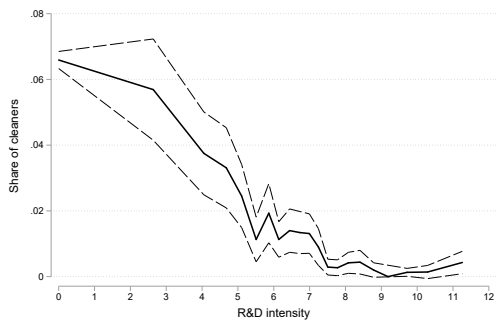
Our model predicts that more innovative firms tend to outsource a higher fraction of tasks than less innovative firms, in particular those tasks with lower complementarity (associated with a smaller μ in the model). Unfortunately, it is not easy to directly measure outsourcing in our data for at least two reasons. First, because outsourced workers do not necessary appear in the ASHE data, and even if they do, they won't

be linked to the firm that use their services. Second, because we conjecture that most of the outsourcing occurred before 2004, which prevents us from following workers in low-skilled occupations that are outsourced from innovative firms as in Goldschmidt and Schmieder (2017).

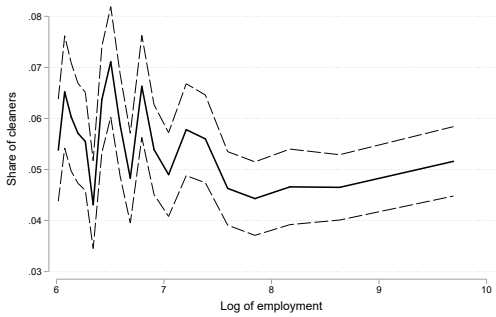
We therefore proceed indirectly. We start from the assumption that, because the technology of cleaning does not vary much across firms, all firms need the same share of cleaners, which can be arguably seen as a low λ task. The only reason that the share of cleaners amongst low-skilled workers would be lower than average in some firms is because those firms outsource cleaning. In Figure 7, we plot the share of cleaners among all workers in low-skilled occupations against R&D intensity in the left-hand side panel and against total employment in the right-hand side panel. Figure 7(a) clearly shows that innovative firms employ fewer cleaners than non innovative firms. Our interpretation is that innovative firms are outsourcing them, and Figure 7(b) suggests that this is not a firm size effect.

Figure 7: Share of workers in low-skilled occupations that are cleaners

(a) by R&D



(b) by employment



Notes: The y-axis shows the share of cleaners over the total number of workers in low-skilled occupations. R&D intensity if plotted by dividing firm into 20 percentiles of the R&D intensity distribution, with one category for non innovative firms; the number shown on the horizontal axis is the mean R&D intensity in that percentile (left-hand side panel) and the average value of employment for each quantile of employment of the firm with 20 quantiles (right-hand side panel).

7 Conclusion

In this paper, we use matched employee-employer administrative data from the UK together with the O'NET survey dataset, to analyze the effect of soft skills on the wage level and wage progression of workers in low educated occupations. We showed that low educated occupations which rely more on soft skills, lead to higher wage levels and to sharper wage progression. Moreover this is particularly true in more skill-intensive and in more innovative firms.

Our analysis can be extended in several directions. First, it would be interesting to look at whether the (low educated) occupations that yield more return to soft skills are more "relational". A second idea is to explore whether our main effects are stronger in more competitive sectors or in areas where potential replacements for incumbent workers in low skilled occupations are of lower quality.

Finally, in this work we used data on a 1% sample of employees. It is likely that workers at different parts of the earning distribution and company hierarchy are differentially affected. We could then look at subgroups of agents within the high and low educated occupation categories. In particular, we would expect the premium to soft skills to be higher at the very top end of the occupation distribution. A first place to look is at CEOs, taking into account their total revenues (wage income plus capital income). These and other extensions of the analysis in this paper await further research.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, March 1999, *67* (2), 251–334.

Acemoglu, Daron, "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, mar 2002, *40* (1), 7–72.

_ **and David H. Autor**, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of Labor Economics" number 4. In 'Orley Ashenfelter and David E. Card (eds.),.', Amsterdam: Elsevier, 2011, pp. 1043–1171.

Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hémous, "Innovation and top income inequality," *The Review of Economic Studies*, 2018, *86* (1), 1–45.

_ **,** _ **, Ari Hyytinen, and Otto Toivanen**, "On the returns to invention within firms: Evidence from Finland," in "AEA Papers and Proceedings," Vol. 108 2018, pp. 208–12.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas, "Immigration and the Rise of American Ingenuity," Technical Report feb 2017.

Barrera-Osorio, Felipe, Adriana D Kugler, and Mikko I Silliman, "Hard and Soft Skills in Vocational Training: Experimental Evidence from Colombia," Working Paper 27548, National Bureau of Economic Research July 2020.

Battiston, Diego, Jordi Blanes i Vidal, and Tom Kirchmaier, "Is distance dead? Face-to-face communication and productivity in teams," Discussion Paper DP11924, CEPR 2017.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer, "Individual effects and dynamics in count data models," *Journal of econometrics*, 2002, *108* (1), 113–131.

_ **,** _ **, and John Van Reenen**, "Market share, market value and innovation in a panel of British manufacturing firms," *The Review of Economic Studies*, 1999, *66* (3), 529–554.

_ **, Robert Joyce, Agnes Norris Keiller, and James P Ziliak**, "Income inequality and the labour market in Britain and the US," *Journal of Public Economics*, 2018, *162*, 48–62.

Bonhomme, Stephane, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, May 2019, *87* (3), 699–739.

Brunello, Giorgio and Lorenzo Rocco, "The labor market effects of academic and vocational education over the life cycle: Evidence based on a British cohort," *Journal of Human Capital*, 2017, *11* (1), 106–166.

Card, David, Ana Rute Cardoso, Jorg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," NBER Working Papers 22850, National Bureau of Economic Research, Inc November 2016.

Carruthers, Celeste and Christopher Jepsen, "Vocational education: an international perspective," Working Paper 8718, CESifo 2020.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, October 2000, *108* (5), 874–904.

_ **and Esteban Rossi-Hansberg**, "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 2006, *121* (4), 1383–1435.

Gibbons, Robert and Lawrence Katz, "Does unmeasured ability explain inter-industry wage differentials?," *The Review of Economic Studies*, 1992, *59* (3), 515–535.

Goldin, Claudia and Lawrence F. Katz, *The Race between Education and Technology*, Belknap Press, 2010.

Goldschmidt, Deborah and Johannes F Schmieder, "The rise of domestic outsourcing and the evolution of the German wage structure," *The Quarterly Journal of Economics*, 2017, p. qjx008.

Groshen, Erica L, "Sources of intra-industry wage dispersion: How much do employers matter?," *The Quarterly Journal of Economics*, 1991, *106* (3), 869–884.

Hanushek, Eric A, Guido Schwerdt, Ludger Woessmann, and Lei Zhang, "General education, vocational education, and labor-market outcomes over the lifecycle," *Journal of human resources*, 2017, *52* (1), 48–87.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who Profits from Patents? Rent-Sharing at Innovative Firms," Working Paper 25245, National Bureau of Economic Research November 2018.

Kremer, Michael, "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, 1993, *108* (3), 551–575.

_ **and Eric Maskin**, "Wage Inequality and Segregation by Skill," Working Paper 5718, National Bureau of Economic Research August 1996.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante, "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

Rodrik, Dani and Stefanie Stantcheva, "A Policy Matrix for Inclusive Prosperity," Working Paper 28736, National Bureau of Economic Research April 2021.

Silliman, Mikko and Hanna Virtanen, "Labor market returns to vocational secondary education," Working Papers 65, ETLA 2019.

Stole, Lars and Jeffrey Zwiebel, "Intra-firm Bargaining under Non-binding Contracts," *Review of Economic Studies*, 1996, *63* (3), 375–410.

A Theoretical appendix

Bargaining

We follow Stole and Zwiebel (1996) and assume that n workers in an occupation Γ with quality $q(\Gamma)$ are bargaining separately for a wage $w(n, q(\Gamma), \Gamma)$. In Stole and Zwiebel (1996)'s framework, if the nth worker in an occupation refuses the wage offer $w(n, q(\Gamma), \Gamma)$, then she is replaced by a worker of quality q_L with wage w_L and the remaining $n-1$ worker the same occupation renegotiate a wage $w(n-1, q(\Gamma), \Gamma)$. By induction, this provides a generic expression for the equilibrium wage.

Formally,the nth worker's surplus is equal to:

$$w(n, q(\Gamma), \Gamma) - \bar{w}^L$$

while firm surplus is:

$$\mu Q + (1-\mu)(q - q_L) - nw(n, q(\Gamma), \Gamma) + (n-1)w(n-1, q(\Gamma), \Gamma) + w_L$$

This implies:

$$w(n, q(\Gamma), \Gamma) = \frac{\mu Q + (1-\mu)(q - q_L) + w_L - \bar{w}^L}{2} \equiv w(q(\Gamma), \Gamma).$$

Importantly, $w(q(\Gamma), \Gamma)$ does not depend upon n, this means that the resulting equilibrium wage is the same that in a simpler case in which all workers in occupation Γ bargain together for a wage $w(q(\Gamma), \Gamma)$, solution to the equalization of the total firm surplus:

$$(\mu Q + (1-\mu)(q - q_L) + w_L)\, \phi(\Gamma) - w(q(\Gamma), \Gamma)$$

and the workers' surplus

$$\phi(\Gamma) \left(w(q(\Gamma), \Gamma) - \bar{w}^L \right)$$

Outsourcing

Consider the maximization problem:

$$q^* = \arg\max_q \left\{ \tilde{\Pi}(\vec{q}) - \int C(q(\Gamma))^2 \phi(\Gamma) d\Gamma \right\}$$

subject to the time constraint

$$\int_{\Gamma} \delta q(\Gamma)\phi(\Gamma)d\Gamma \leq T.$$

If ν denotes the Lagrange multiplier associated with the constraint, then as long as it is larger than q_L, the optimal value of $q(\Gamma)$ is defined by:

$$\nu\delta = \frac{\mu(Q-1)+1}{2} - 2Cq(\Gamma),$$

or equivalently:

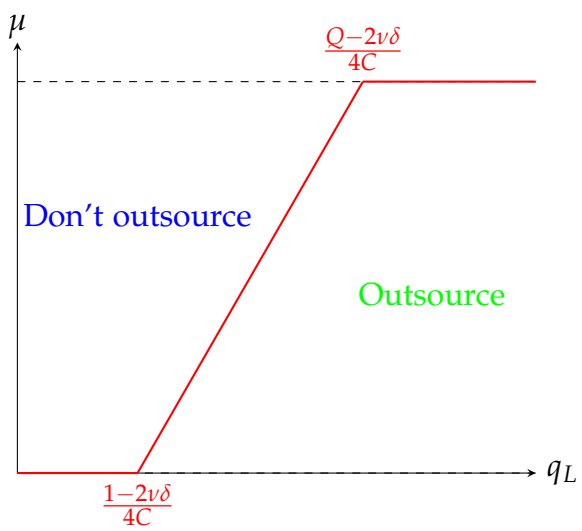
$$q(\Gamma) = \max\left[q_L, \frac{\mu(Q-1)+1}{4C} - \frac{\nu\delta}{2C}\right]$$

This defines an implicit cutoff space $\bar{\Gamma}$ defined by $\bar{\mu}(q_L)$ such that:

$$4Cq_L = \bar{\mu}(q_L)(Q-1) + 1 - 2\nu\delta,$$

and which is represented in Figure A1.

Figure A1: Outsourcing space



To find the value of ν, we use the fact that the constraint is binding, that is:

$$\frac{T}{\delta} = \frac{Q-1}{4C}\int_{\Gamma\in\mathcal{R}} \mu\phi(\Gamma)d\Gamma + \frac{1-2\nu\delta}{4C}\int_{\Gamma\in\mathcal{R}} \phi(\Gamma)d\Gamma + \int_{\Gamma\notin\mathcal{R}} q_L\phi(\Gamma)d\Gamma,$$

where $\mathcal{R}$ denotes the area of Γ in which the firm does not outsource. To simplify

the exposition, we assume that there is only one value of q_L for all occupations and all workers to focus on the effect of μ. The problem therefore becomes one dimensional and the previous equation can be rewritten as:

$$4C\left(\frac{T}{\delta} - q_L\right) = (Q-1)\left(\int_{\mu>\bar{\mu}} \mu\phi_\mu(\mu)d\mu - \bar{\mu}(1-\Phi_\mu(\mu))\right),$$

where Φ_μ is the cumulative distribution function of μ, associated with a density ϕ_μ. The right hand side is decreasing in $\bar{\mu}$ (its derivative is equal to $\Phi_\mu(\bar{\mu}) - 1 < 0$) and the left-hand side is a constant. Hence increasing the value of Q implies an increase in $\bar{\mu}$ which means that more occupation are outsourced.