CVER Discussion Paper Series - ISSN 2398-7553

CENTRE FOR VOCATIONAL EDUCATION RESEARCH

The Changing Demand for Skills in the UK

Andy Dickerson and Damon Morris

Research Discussion Paper 020

April 2019

CENTRE *for* ECONOMIC PERFORMANCE

National Institute of Economic and Social Research

LE London Economics

The University Of Sheffield.

The Centre for Vocational Education Research (CVER) is an independent research centre funded by the UK Department for Education (DfE). CVER brings together four partners: the LSE Centre for Economic Performance; University of Sheffield; National Institute of Economic and Social Research and London Economics.

Published by:
Centre for Vocational Educational Research
London School of Economics & Political Science
Houghton Street
London WC2A 2AE

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Department for Education

THE CHANGING DEMAND FOR SKILLS IN THE UK

Andy Dickerson* and Damon Morris

April 2019

Abstract

We present estimates of changes in skills utilisation and in the returns to skills in the UK for 2002-2016 using new measures of skills derived from a systematic and detailed matching between the US O*NET system and UK SOC. Over the period, there is strongly increasing utilisation of both analytical skills and interpersonal skills, and declining use of physical skills. A decomposition analysis reveals that most of the change in skills utilisation is within occupations rather than between occupations, suggesting that the changes are pervasive throughout employment. The returns to skills are estimated using a standard Mincerian earnings function. We find positive and significantly increasing returns to analytical skills throughout the period. While the returns to interpersonal skills are lower than to analytical skills, they are also increasing over time, and are significant especially post-2010. Finally, the returns to physical skills are significantly negative over the whole period. The results suggest that the UK labour market is strongly increasing its demand for both analytical and interpersonal skills.

JEL Classification: J20; J24; J31

Keywords: Skills; Occupations; Earnings; O*NET

Acknowledgements: The Centre for Vocational Education Research (CVER: cver.lse.ac.uk) is an independent research centre funded by the UK Department for Education (DfE). Any views expressed are those of the authors, and do not represent the views of DfE. Part of the research for this paper was undertaken while Dickerson was Visiting Research Fellow in the Department of Economics at the University of Melbourne, and he would like to thank them for their kind hospitality. Useful comments received from participants at the CVER Annual Conference (LSE, September 2017), the Royal Economic Society Annual Conference (Sussex, March 2018), and the 2nd IZA Labor Statistics Workshop on the Returns to Skill (Bonn, April 2018) are also gratefully acknowledged.

Affili`ations: Department of Economics, University of Sheffield and CVER

* University of Sheffield, 9 Mappin Street, Sheffield S1 4DT.
a.p.dickerson@sheffield.ac.uk

THE CHANGING DEMAND FOR SKILLS IN THE UK

1. Introduction

The literature has produced two dominant canonical explanations for the observed changes in employment and wages over the last 30 years. First, proponents of skill-biased technical change (SBTC) argued that technology has monotonic effects throughout the skills distribution, and can therefore explain the observed increased returns to education for example (Berman et al, 1994, 1998; Katz and Murphy, 1992). Second, those favouring a rather more nuanced interpretation of the impact of technology which distinguishes tasks and skills have argued that routine tasks in the middle of the skills distribution are increasingly becoming automated as compared to jobs at either ends of the skills spectrum, and this has led to job 'polarisation' (Autor et al, 2003; Goos and Manning, 2007; Goos et al, 2009; Autor et al, 2008). More recently, Beaudry et al (2016) present evidence for a 'great reversal' in the US, with stagnating or decreasing returns to cognitive skills since 2000 for young workers 25-35, and higher-skilled workers displacing lower-educated workers in less-skilled jobs. Finally, Deming (2017) argues that there is a growing importance of 'social skills' in the US labour market, with an increasing share of US jobs requiring high levels of social interaction. He provides evidence for this hypothesis in the form of increasing returns to social skills post-2000. Similar findings are reported for Swedish prime-aged males by Edin et al (2017). One possible explanation is that this may be the flip-side of the increased automation and routinisation of jobs.

Most of the tasks vs skills literature is US-focussed and typically utilises measures of tasks and skills derived from the US Dictionary of Occupational Titles (DOT) and/or its successor O*NET (Occupational Information Network) (e.g. Autor et al, 2003; Abraham and Spletzer, 2009), although there are also a few bespoke surveys (e.g. Autor and Handel, 2013; Handel, 2016b). In contrast, for the UK, while 'skills' have long been a major policy priority (e.g. DEE, 2000; Leitch, 2006; DBIS, 2009; UKCES, 2009, 2010, 2014), there are only very imperfect measures of the skills available and in use in employment. In the UK, skills are usually proxied by individuals' qualifications or by the occupational classification of the jobs they do. While these are both reasonably simple to record in surveys and censuses, they are both poor proxies for skills. Qualifications are normally obtained prior to labour market entry, and any knowledge or abilities acquired while in education may have long since become obsolete, forgotten or atrophied, or may be irrelevant to the current employment. More fundamentally, qualifications are not, de facto, skills[1]. The standard occupational classification (SOC) is also extremely imperfect as a measure of a worker's skills. SOC is uni-dimensional and static, and so captures neither the variety of skills used in different jobs, nor the changing nature of skills over time. This paper comprehensively addresses these fundamental weaknesses in the measurement of skills and skills utilisation in employment in the UK for the first time.

In contrast to the paucity of measures of skills for the UK, the US O*NET system (https://www.onetcenter.org/) provides almost 250 descriptors of skills, abilities, work activities, training, work context and job characteristics for each of around 1,000 different occupations. It comprises both job-orientated and worker-orientated characteristics at occupation-specific and cross-occupation levels. O*NET is the main source of occupational competency information in the US. It superseded the DOT in 2001 having been almost 20 years in development (Peterson et al, 1999; Tippins and Hilton, 2010). It is constantly being revised and updated on a rolling basis. Information is gathered from self-reported assessments by job incumbents based on standardised questionnaire surveys as well as from professional assessments by job evaluation analysts[2]. Ideally, there would be an O*NET-type system for the UK. But in the absence of such a system, we match O*NET to the UK in order to provide the same level of detail in terms of the range of skills descriptors that are available. Thus we develop a database of comprehensive and detailed multi-dimensional occupational skills profiles for the UK which describe the utilisation of skills used in the workplace. These occupational skills profiles have many potential uses. For example, they can enable a much richer and deeper understanding of the changing patterns of the demand for skills to be developed. They can also be used to assess the changing value/returns to skills in

[1] Moreover, employers increasingly focus on characteristics of potential employees other than their educational qualifications when recruiting. These characteristics include 'soft' skills and abilities such as creativity, problem solving, teamwork and communication, as well as traits and attitudes such as honesty, integrity and self-motivation. These are undoubtedly more difficult to measure than qualifications or occupation (although some progress has been made in small-scale surveys that focus on the tasks that individuals perform in their jobs e.g. Felstead et al, 2007; 2013).

[2] Further information on O*NET is provided in Appendix A.

employment. Finally, they could be used to help inform individuals, and those who advise them, on the skills that are useful in employment today.

More specifically, we construct a systematic and detailed match between the occupational and job taxonomy in O*NET and the UK SOC. We can then use the information in O*NET to produce a set of descriptors of the skills used in occupations in the UK. Methodologically, we are essentially assuming that, on average, the skills of e.g. a plumber in the UK are similar to the skills of a plumber in the US, and we provide corroborating evidence in support of this assumption using directly comparable US-UK data drawn from PIAAC (Programme for the International Assessment of Adult Competencies - OECD 2016) in Section 3 below.

We then utilise our occupational skills profiles to assess the changing demand for skills in the UK. We construct three indices of skills: analytical/cognitive skills; interpersonal skills; and physical/manual skills. We combine these with individual data on wages and employment from the Annual Surveys of Hours and Earnings (ASHE) and the Labour Force Survey (LFS) to produce a 4-digit SOC occupational-level panel dataset for 2002-2016. We use this dataset to examine the change in skills utilisation in employment over the period, and to estimate the wage returns to these skills. We argue that these two measures together provide a comprehensive picture of the changing demand for skills in the UK.

Our results indicate strongly increasing use of both analytical skills and interpersonal skills, and declining use of physical skills over the period 2002-2016. A decomposition analysis reveals that most of the change in skills utilisation for all three measures is within occupations, rather than between occupations. This indicates that the changes in skills utilisation are pervasive throughout employment. The wage returns to skills are estimated using a Mincerian-type earnings function. The returns to analytical skills are positive and increasing over time, suggesting that the demand for such skills is increasing even more strongly than the growth in their utilisation. While the returns to interpersonal skills are lower than to analytical skills, they are also increasing over time, and are significantly positive post-2010. Finally, the returns to physical skills are significantly negative over the whole period, although are approximately constant despite the strong secular decline in physical skills utilisation in employment over the sample period.

These findings are robust to the definitions and measurement of the skills variables, and to the empirical specification of the earnings function. The results suggest that the UK labour market is strongly increasing its demand for analytical and interpersonal skills.

The remainder of this paper is structured as follows. The next section briefly reviews some previous studies which have used O*NET and similar systems to measure and assess skills. Section 3 briefly outlines the methodology we have developed to construct our occupational skills profiles. Section 4 describes the trends in skills utilisation over time and presents a decomposition of the change in each skill index over the whole period into its between-occupation and within-occupation changes. Estimates of the returns to skills are then presented together with the changing patterns in these returns. Section 5 concludes with a discussion of some potential implications for education and skills policy in the UK.

2. Measuring skills

The importance of skills in modern economies is widely acknowledged. Skills are important at both micro level e.g. for the distribution of earnings, and at the macro level e.g. for explanations of productivity and growth. Despite the fundamental importance of skills in economic policy discourse, procedures for measuring skills are comparatively under-developed in almost all countries. Skills are multi-dimensional, intangible and often unobservable. Each of the different conceptualisations of skills and their proxies that are commonly employed in research and policy analysis can be argued to have a number of serious weaknesses and limitations (Green, 2006, 2013).

The most commonly employed proxy for the skills of an individual is their qualifications or educational attainment. This measure has the advantage of being objective, and long-term trends can be assessed if qualification standards remain unchanged. However, qualifications only have a loose link with job skills and thereby individual and economy-wide economic performance. Not all educationally-derived skills will be utilised in the labour market (due to mismatch/overqualification), and the acquisition and depreciation of skills continues after education is completed. Moreover, education may be a signal of ability rather than a source of skills supply (Spence, 1973). Learning at work is important for the acquisition of new skills and for updating existing skills. Hence the relationship between education and skills, and thereby both individual and macroeconomic performance, is complex. Certainly, measuring skills by education qualifications alone is not sufficient. International comparisons of skills using educational qualification attainment are also problematic because qualifications are not comparable across countries. Measuring the length of time in education (e.g. Barro and Lee, 2013) is not a solution to this problem since there is variability in the quality of education provision between countries and over time.

A second commonly employed proxy for skills is occupation. While this measure can be readily obtained from surveys and censuses, the uni-dimensional hierarchy of occupations in occupational classifications is contestable, uncertain and changing. In addition, while detailed, disaggregated occupational classifications are provided by national statistical agencies (e.g. SOC2010, 2010; BLS, 2018), data are often only available at more aggregated levels which can make comparisons rather crude. Moreover, over time, skills change within and between occupations, and these changes are not reflected in the SOC which is static, and only periodically updated (every 10 years in the UK for example).

Formal tests of skills can be made, and international comparisons are possible, if expensive. However, formal assessments of skills through tests can only ever measure a limited range of skills (literacy and numeracy are typical). They are comparatively rare and typically have small sample sizes because of the costs of administering such testing. Examples include the International Adult Literacy Surveys (IALS) (OECD, 2000) and the Programme for the International Assessment of Adult Competencies (PIAAC) (OECD, 2016). There has also been criticism of the international comparability of universal testing even when it has been treated very carefully by researchers. An alternative is self-assessment of skills. While this is subjective, and so used very rarely, the 5th sweep of the UK National Child Development Survey (NCDS) records such measures (McIntosh and Vignoles, 2001). The major problem using self-assessment to measure skills is that skill self-assessment is associated with self-esteem.

Finally, there is the job requirements approach. These are surveys which ask individuals about the generic tasks and skills they use in their jobs and use their responses to infer the skills that they have. Of course, mismatch and underutilisation are still a potential problem, but they have permitted a much richer description of individuals' skills, including soft/generic skills which are simply not captured by the standard measures of skills. They also permit a wide range of skills to be assessed. Obviously, job skills could differ from person skills (because of mismatch), and skills are only measured for those in employment. But this method can make use of commercial job analysis data (which is arguably objective), as well as bespoke (subjective) surveys of individuals. Examples include: O*NET in the US; BIBB/IAB and BIBB/BAuA Surveys on Qualifications and Working Conditions in Germany (BIBB, no date; Rohrbach-Schmidt and Tiemann, 2016); and the Skills Surveys in the UK (Dickerson and Green, 2004; Felstead et al, 2007; 2013).

Table 1 summarises a selection of the papers which have utilised measures of skills derived from data collected using the job requirements approach. These papers use the DOT, or O*NET, or other job-task surveys with similar structures and/or characteristics to the O*NET. It is common to select a subset of 'relevant' O*NET items corresponding to some pre-defined taxonomy of skills, although this selection can sometimes seem somewhat arbitrary. As can be seen, a three-way classification of skills/attributes has proven popular, following the development of Fine's Functional Job Analysis (FJA) theory in the 1950s (Fine, 1955; Fine and Cronshaw, 1999) and formally implemented in the DOT occupational codes as 'Data-People-Things' (although the taxonomy today is typically: Analytic/Cognitive, Interpersonal and Physical/Manual skills, or some variant thereof). There is little standardisation of the measures that are chosen even when the language/description of the skills taxonomy is very similar. However, a focus on cognitive and non-cognitive routine and non-routine tasks (and the substitution of – especially – computing technology for routine tasks as emphasised by David Autor and co-authors) is also popular. Amalgamation/aggregation methods include averaging a very small number of descriptors from the O*NET system, through to factor analysis across a very broad range of (possibly heterogeneous) indicators. One common characteristic of all the O*NET based studies listed in Table 1, and others that have also used O*NET data, is that they all use only a single O*NET vintage even when there is a time dimension to the analysis undertaken. An important advantage of O*NET is that it is being continuously updated to reflect changing skills and skills utilisation between and within occupations, and the different O*NET vintages incorporate the updated measures of skills. In the analysis presented below, different vintages of O*NET are matched as appropriate to each year of the data that we use.

3. Data and Methodology

3.1 Data

We combine 4 different sources of data to construct a UK SOC2010-consistent 4-digit[3] occupational panel dataset for 2002-2016 comprising detailed measures of wages, employment composition, qualifications and skills[4]. The data sources are:

1. UK Labour Force Survey (LFS) data, 2002-2016;
2. UK Annual Survey of Hours and Earnings (ASHE) data, 2002-2016;
3. US O*NET, 2002-2016 (v4.0 to v21.0);
4. US Occupational Employment Statistics (OES), 2002-2016.

UK LFS microdata and ASHE/NES occupation-level public release tables are used to provide data on the structure and composition of earnings and employment at the 4-digit (unit group) occupation level. For 2002-2010, the data provided are classified on SOC2000 while for 2011-2016, SOC2010 is used. We use ASHE data for occupational wages because of the larger sample sizes available for the detailed 4-digit occupations that we are using. ASHE is based on a 1% random sample of all employees in employment, and so is a considerably larger sample than available from the LFS. The average coefficient of variation ($CV = \sigma/\bar{x}$) for mean hourly occupational wages calculated from ASHE is approximately one tenth of the magnitude of the comparable LFS statistic. ASHE data on earnings is also provided by employers rather than the employees themselves. As well as its much more limited scale, the LFS also has a large number of proxy responses, and these may be particularly problematic when recording earnings[5]. However, ASHE has only very limited information on personal characteristics and, in particular,

---

[3] UK SOC2010 has nine major groups, 25 sub-major groups, 90 minor groups and 369 unit groups, and uses a 4 digit system for classification. The first digit represents the major group, the second digit represents the sub-major group, the third digit represents the minor group and the final digit represents the unit group. Our analysis is at the most disaggregated unit group – 4-digit – level.

[4] Further details on the data and methodology are provided in Appendix B.

[5] The LFS also suffers from a declining response rate, and which is now amongst the lowest in the EU (ONS, 2014).

it has no information on qualifications. Thus LFS microdata, aggregated to the 4-digit occupation level, is used to provide occupation-level data on educational qualifications.

In order to produce data on a consistent occupational classification, we use the ONS-supplied 'correspondence tables' to convert the SOC2000 data for 2002-2010 to SOC2010 (ONS, 2012). The ONS weights for this mapping are derived from dual-coded individual level datasets in which detailed occupation is recorded according to both SOC2000 and SOC2010. These dual-coded datasets are then used to estimate the employment composition of SOC2010 codes in terms of SOC2000 occupations. There are three dual-coded datasets: LFS January-March 2007 (LFSJM07); 2001 Census (Census01); and LFS December 1996-February 1997 (LFS96-97). The weights differ according to the dataset used and, in some cases (where occupational employment is minimal), there are no figures available. Each dual-coded dataset is used in turn to produce SOC2010-consistent occupational level data for the 2002-2010 period. Our main results reported below are for the average weights calculated across the three dual-coded datasets, although we investigate the sensitivity of the results to that decision in our robustness tests.

The Occupational Information Network, O*NET, system provides measures of skills, abilities, work activities, training, and job characteristics for almost 1,000 different US occupations. It is the main source of occupational competency information in the US. Information is gathered from self-reported assessments by job incumbents based on standardised questionnaire surveys together with professional assessments by job evaluation analysts. For the four areas of (a) knowledge, (b) skills, (c) abilities and (d) work activities, both the 'Importance' and 'Level' of each characteristic being measured is recorded. Most descriptors are comparable between occupations, although 'tasks' are occupation-specific. In the analysis that follows, we only utilise the O*NET measures of skills.

O*NET information is gathered from postal and online questionnaires administered by the US Bureau of Labor Statistics (BLS). Respondents are only asked to complete a random selection of the questionnaires in order to avoid survey fatigue. They also indicate from a wide range of occupation-specific tasks those that apply to their particular job. O*NET publishes occupation averages, rather than the individual micro-data. However, these averages are based on large samples - an average of 31,000 responses for each of the 239 descriptors gathered from around 125,000 returned questionnaires - and response rates are reported to be high - in excess of 75% for employers and 65% for employees (Handel, 2016a; US Department of Labor, 2005). Information is published at the 'O*NET-SOC' occupation level, which is a slightly more detailed version of the US SOC. There are currently 1,110 occupations in O*NET SOC2010 (cf 840 in US SOC2010), although data are only collected on 974 of these occupations (these are termed the 'data level occupations'). Further information on O*NET is provided in Appendix A.

3.2 Methodology

A brief description of the matching methodology used to construct our skills indices is provided in this subsection; full details are in Appendix B. Our skills measures are constructed as follows. We first identify the match between the O*NET occupations and UK SOC 4-digit occupations. Matching is undertaken using CASCOT (Computer Assisted Structured Coding Tool). CASCOT utilises the underlying job titles and SOC structure in the O*NET SOC and UK SOC to identify the O*NET occupations that are most closely aligned to each UK 4-digit SOC. Given there are 369 4-digit UK SOC occupations and 1,110 O*NET occupations, this necessarily a one-to-many match, but we also permit the same O*NET occupation to be matched to more than one UK 4-digit occupation. Full details of the matching are described in Appendix B.

We then compute a vector of skills, $S_{jt}^{(x)}$, for each UK 4-digit occupation $j = 1, \ldots, J$ at time t, defined as:

$$S_{jt}^{(x)} = \sum_{\substack{k=1 \\ k \in \{S_j\}}}^{K_j} O_{kt}^{(x)} \frac{n_{kt}}{\sum_k n_{kt}} \tag{1}$$

where $O_{kt}^{(x)}$ is the measure of skill x for O*NET occupation k at time t, n_{kt} is employment in occupation k as derived from OES, and $\sum_k n_{kt}$ is total employment across all occupations k. The summation is over the set $k \in \{S_j\}$ of the K_j O*NET occupations that are matched to each particular UK 4-digit occupation j. Thus for each skill, $S_{jt}^{(x)}$ is the OES employment-weighted average of the O*NET measure of skill for the set of O*NET occupations that matches to each 4-digit UK occupation j. We calculate these indices separately for each of the

$x = 1, \ldots, 35$ measures of skills in O*NET and for each year $t = 2002, \ldots, 2016$ using successive vintages of the O*NET data.

We aggregate the resulting 35 skills measures into three indices closely informed by the 'data-people-things' taxonomy originally utilised in DOT, although we use the terms 'analytical skills', 'interpersonal skills', and 'physical skills' respectively. This taxonomy is defined as:

Analytical skills (21 items):

Reading Comprehension, Writing, Mathematics, Science, Critical Thinking, Active Learning, Learning Strategies, Monitoring, Coordination, Negotiation, Complex Problem Solving, Operations Analysis, Technology, Design, Programming, Troubleshooting, Judgment and Decision Making, Systems Analysis, Systems Evaluation, Time Management, Management of Financial Resources, Management of Material Resources

Interpersonal skills (7 items):

Active Listening, Speaking, Social Perceptiveness, Persuasion, Instructing, Service Orientation, Management of Personnel Resources

Physical skills (7 items):

Equipment Selection, Installation, Operation Monitoring, Operation and Control, Equipment Maintenance, Repairing, Quality Control Analysis

There are a number of ways in which these items can be aggregated to provide a single index of skills (e.g. simple averaging across the component skills indices, or using Principal Components Analysis). There are additional choices regarding the inclusion or otherwise of the Levels as well as Importance measures of each skill. In our main results, we simply take the average of the importance measures of the skills only, although we examine the sensitivity of our findings to this choice in our extensive robustness analysis.

In order to produce a SOC2010-consistent 4-digit panel for 2002-2016, it is necessary to resolve the changes in the occupational classification that have taken place in the US as well as in the UK over our sample period. Equivalent to the correspondence tables for the UK SOC changes, there are 'crosswalks' for the changes in the US SOC and O*NET SOC to enable conversions between the different SOC classifications. We use these to produce a UK

SOC2010-consistent 4-digit panel for 2002-2016 with information on employment composition and structure, wages, together with the measures of skills derived from O*NET.

One final issue is that the O*NET measures of skills in the early part of our sample period (2002-2009) were wholly or partially provided by job incumbents rather than job analysts. However, from O*NET version 15.0 (2010) onwards, the skills measures were exclusively provided by job analysts for all occupations. The differences between incumbents' and analysts' ratings of O*NET skills importance measures have been analysed by Mumford et al (1999) and, more comprehensively, by Tsacoumis and Van Iddekinge (2006)[6]. Both conclude that, while job incumbents tend to provide higher skills ratings than job analysts on average (Mumford et al, 1999) report a mean difference of 0.58 standard deviations for example), there is a very close correspondence in the skills ratings both within and between occupations by the two types of respondents. The average correlation of the skills importance ratings between job incumbents and job analysts was very high within SOC in particular (Tsacoumis and Van Iddekinge, 2006, report an average correlation of r = 0.85). As Tsacoumis and Van Iddekinge (2006) conclude, "the results of this study revealed minimal differences between the two systems of obtaining skill information." (p.17).

While this suggests that any cross-sectional analysis using O*NET skills measures will be little affected by whether job incumbents or job analysts provide the skills ratings, the switch from incumbents to analysts does have potential implications for any comparisons of skills over time. As the share of job analysts' ratings increases, the mean skills rating will fall, and to the extent this affects occupations differentially over time, this may also impact on the measures of UK occupational skills that we construct. One solution is to use the changing mix of job incumbents and job analysts between occupations to impute the 'job incumbent-effect' by occupation, which we can then subtract from the skills measure to produce a job-analyst consistent measure of skills for the whole period[7]. Details of this adjustment are shown in

[6] The importance and level scales are closely correlated, and hence both Mumford et al (1999) and Tsacoumis and Van Iddekinge (2006) focus on skills importance measures only. This close relationship between the 'importance' and 'levels' indices – perhaps because of the difficulty in distinguishing between them – has been noted by others e.g. Handel (2016a).

[7] This is not to suggest that we regard the job analysts' ratings as 'correct' or somehow better than job incumbents' ratings - we could equally construct a job analyst effect and add this to produce a job-incumbent consistent measure of skills for the whole period.

Appendix C. We also investigate the robustness of our findings to the adjustment method we have employed.

3.3 Correspondence between UK and US skills using PIAAC

The validity of the methodology adopted for measuring UK skills depends on there being a close correspondence between the skills utilised in similar occupations in the US and the UK so that the mapping of skills between O*NET occupations and UK SOC occupations is valid. In order to be confident that this is indeed the case, in this subsection we compare skills in US and UK occupations in a dataset where the two are directly comparable. The Survey of Adult Skills is part of the OECD Programme for the International Assessment of Adult Competencies (PIAAC) (OECD, 2016; Hanushek et al, 2015). PIAAC is an internationally comparable survey that assesses the proficiency of adults in 'numeracy', 'literacy' and 'problem-solving in technology-rich environments'. The first round of PIAAC data was collected between August 2011 and March 2012 in most participating countries, including the US and the UK (for England and Northern Ireland only). Approximately 9,000 adults participated in the UK and 5,000 in the US.

PIAAC is ideal for our purposes because it collects data on the same skills, using the same methodology and questions, and is coded to a common occupational classification (ISCO-08). This enables direct comparison of occupational skills in the UK and the US[8]. Numeracy skill in PIAAC is defined as: "the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life." (OECD, 2012, p.32). Literacy is defined as: "understanding, evaluating, using and engaging with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential." (OECD 2012, p.20). Finally, problem-solving is defined as "using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks. … the abilities to solve problems for personal, work and civic purposes by setting up appropriate

[8] One limitation is that the skills in PIAAC are primarily cognitive skills – numeracy, literacy and problem solving – and so the UK-US comparison provides a useful indicator of the validity of mapping cognitive skills in O*NET to the UK SOC but is perhaps less suitable for assessing the mapping of non-cognitive skills.

goals and plans, and accessing and making use of information through computers and computer networks." (OECD, 2012, p.47).

Given the sample sizes, we aggregate the PIAAC data to the 2-digit occupation level, yielding 39 occupational groups, and compute the averages of the skills assessments in each occupational group in both countries. Figure 1 compares UK and US numeracy skills. With the exception of a single outlier (ISCO-08 92: Agricultural, forestry and fishery labourers), there is a very close correspondence between the level of numeracy skills in each country. The points lie close to the 45^{o} line denoting equal UK and US occupational skill levels. The linear fit is close to being parallel to the 45^{o} line, suggesting that numeracy skills in the UK are above those in the US by approximately the same amount in each occupation[9]. The exception is agricultural workers who have significantly lower numeracy skills in the UK than in the US.

A similar pattern is evident in Figure 2 for literacy skills which are again marginally higher in the UK than the US, although in occupations where workers exhibit the highest level of literacy skills, skill levels are very comparable between the UK and the US. Finally, Figure 3 compares problem-solving skills in the UK and the US. While there is more variation here, as might be expected given the nature of the problem-solving assessment, the correlation between UK and US occupational skill levels is still strongly positive and there is a clear linear association between the skill levels.

We compute two measures of the differences in skills by 2-digit ISCO occupation between the UK and the US. Let $\hat{y}_j^S$ be the level of skill S in occupation j in the UK and y_j^S be the corresponding level in the US. We calculate the mean absolute percentage error (MAPE) as:

$$MAPE^S = \frac{100}{J}\sum_{j=1}^{J} | \frac{\hat{y}_j^S - y_j^S}{y_j^S} |$$

[9] We also estimated a weighted regression with weights proportional to the average US-UK cell counts for each occupation (i.e. proportional to average employment). As can be seen, weighting makes very little difference to the fitted relationship.

and the Root Mean Square Error (RMSE) as:

$$RMSE^S = \sqrt{\frac{\sum_{j=1}^{J}(\hat{y}_j^S - y_j^S)^2}{J}}$$

which can also be 'normalised' by dividing by the mean of y_j^S. Table 2 presents the MAPE and RMSE measures for numeracy, literacy, and problem-solving skills. These differences are small on average.

Figures 1 to 3 together with the statistics in Table 2 suggest that mapping the skills for US occupations to UK occupations has high validity for both the levels (subject to a mean shift) and the rankings of numeracy, literacy and problem-solving skills at the 2-digit level. This is not the first attempt to compare O*NET descriptors between countries. Taylor et al (2008) report good levels of correspondence (in means and rank orderings) for a limited range of O*NET measures for US, New Zealand, Hong Kong and Chinese workers for example.

4. Results

In this section we utilise the occupational skills profiles we have constructed to assess the changing demand for skills in the UK. First, we examine the change in skills utilisation in employment over the period 2002-2016 both in aggregate, and also decomposed into within and between occupation components. Any net increase or decrease in skills utilisation will necessarily reflect changes in both demand and supply of course. Thus we also estimate the wage premium paid to skills. Observing both changes in quantities and changes in 'prices' enable an assessment of the changing demand for skills in the UK.

4.1 Skill trends and decomposition

The overall changes between 2002 and 2016 in analytic, interpersonal and physical skills are reported in the first column of Table 3. Over the whole period, the employment-weighted[10] aggregate index of analytical skills suggests that utilisation of this skill set grew by

[10] i.e. the skills indices are weighted by their employment shares in total employment for each year.

10% over the period. The increase in interpersonal skills was more than double this (+23%), while utilisation of physical skills fell by 14%. These trends accord with our general understanding of the changing occupational structure of employment and the growth of services and the decline of manufacturing (e.g. Oesch, 2013).

At the aggregate level, these trends are a consequence of a combination of both changing skills within (broader) occupations, and changes in the occupational structure of employment. Some evidence on where the changes are primarily situated can be obtained from undertaking a decomposition of the overall change in skills utilisation between 2002 and 2016 in each of the three skill measures. Specifically, we examine the extent to which the aggregate changes in each index of skills is a consequence of within-occupation or between-occupation changes. The change in average skill utilisation over time, ΔS, can be decomposed as follows:

$$\Delta S = \sum_{j=1}^{J} \Delta e_j \bar{S}_j + \sum_{j=1}^{J} \Delta S_j \bar{e}_j \quad (2)$$

where *j* indexes occupations, $j = 1, \ldots, J$, an overscore denotes an average over time, $e_j = \frac{E_j}{E}$ is the share of total employment in occupation *j*, and S_j is the level of skill utilisation in occupation *j*. The first term on the right-hand side of equation (2) is the between-occupation change in skill utilisation, while the second term is the within-occupation change.

Table 3 reports the decomposition of the overall change in analytical skills, interpersonal and physical skills over the period 2002 to 2016 using 1-digit, 2-digit, 3-digit and 4-digit occupational classifications. As can be seen, the within-occupation changes in skills dominate the between-occupation changes for all three indices whatever level of occupational disaggregation is employed. Around 20-25% of the increase in analytical skills utilisation is between occupations, while the remaining 75-80% is within occupations. The within-occupation changes for interpersonal skills and physical skills are even greater. This decomposition suggests that the overall changes in skill utilisation are pervasive throughout employment and are affecting all occupations, rather than being concentrated in certain occupational groups. Thus, over the period 2002 to 2016, the UK labour market has seen a substantial increase in the utilisation in employment of analytic and, especially, interpersonal skills, and a decline in the use of physical skills in employment.

4.2 Returns to skills

We next turn to examine the returns to skills. We use a simple Mincerian log earnings function specification to estimate the conditional (wage) returns to skills and to compute the changing returns over time. This is similar in spirit to Ingram and Neumann (2006) for example, although here the unit of observation is the 4-digit occupation. Table 4 presents the basic log hourly wage regression results[11]. Column (1) shows that wages are positively correlated with analytical skills, and negatively correlated with interpersonal and physical skills. These correlations are highly significant statistically. Column (2) reports the basic earnings function estimates without the skills indices. This demonstrates that higher qualifications are associated with higher earnings in general; wages increase with age at a decreasing rate, and that the age-earnings profile is inverse-U-shaped; occupations with higher proportions of women and public sector workers pay less on average; and that larger firms tend to pay significantly more. These are all consistent with standard findings in the earnings function literature. Column (3) augments our earnings equation with the three indices of skills. This suggests that skills and education are positively correlated in general, such that at least some of the returns to education are, in fact, returns to skills. Year dummies are included in column (4) since there are macro and other temporal changes which have impacted on earnings in this period, including the 2008 financial crisis which has significantly affected the level and growth of average real earnings in the UK (e.g. Gregg et al, 2014). These do not change the qualitative findings. The results in column (4) suggest that there are positive and statistically significant return to analytical skills, and negative and statistically significant returns to physical skills.

Over the period of analysis, there have been considerable changes in the UK labour market at both micro and macro levels. Both the composition of the labour force and of employment have changed significantly over the period, and these changes are also reflected in the utilisation of skills as reported in subsection 4.1 above. In order to allow for this, we estimate a fully interacted variant of Table 4, column (4), in which the regression coefficients

[11] Clearly, no causal interpretation of the coefficients is being attempted here (Card, 2000). Rather we present the estimated wage premia associated with analytic, interpersonal and physical skills over time, conditional on the educational qualifications and other characteristics of the occupational group (Ingram and Neumann, 2006).

are allowed to differ by year (equivalent to estimating a series of annual cross-section regressions). The returns to our three measures of skills are illustrated in Figure 4, where we have standardised (mean 0, variance 1) the skills indices in order that comparisons between them can more easily be made.

The dashed lines connect the year-by-year point estimates of the wage returns to analytic, interpersonal and physical skills. As can be clearly seen, the returns to analytic skills are strongly trended upwards over time. An alternative specification which interacts a linear time trend with the index of analytical skills is superimposed (together with its 95% confidence interval). The coefficient on the time trend for analytic skills 0.017 (0.003***), such that an occupation with a one standard deviation higher level of analytic skills will be associated with almost 2% higher wage growth relative to an occupation with an average level of analytic skills. Clearly, over the sample period, the returns to analytic skills have been not only positive and statistically significant but have been increasing strongly. It is important to note that this increase in returns has occurred while the utilisation of analytical skills has also been increasing as shown in the previous subsection.

The returns to interpersonal skills were clearly close to zero in the early part of the sample period, but have also been increasing over time. A linear time trend has a statistically significant slope coefficient (the coefficient on the trend is 0.008 (0.002***)), and the returns are statistically significantly positive post-2010. Again, this increasing return has occurred at the same time as the utilisation of interpersonal skills has been increasing sharply as documented in subsection 4.1 above. We therefore conclude that the demand for both analytical and interpersonal skills is strongly increasing over the period of analysis.

Finally, the returns to physical skills are negative throughout the period but are fairly constant over time. In this case, the slope of the time trend is insignificantly different from zero (the coefficient on the trend is -0.002 (0.002)). Recall that the utilisation of these skills has been falling sharply over the period. This suggests declining demand for these skills in employment over time, although this has been coupled with a corresponding reduction in supply.

4.3 Robustness Checks

In order to investigate the robustness of our findings to the various decisions made in constructing the dataset, as well choices regarding our specification and econometric approach, we undertake a number of sensitivity checks of our main results.

4.3.1 Data Transformations and Sources

In Table 5 we present the robustness of our findings to the particular method we use to aggregate the 35 skills into our analytic, interpersonal and physical skills indices. Panel A reports results which are based on aggregations using the importance of skills only. In the first column of Panel A, the estimates are based on the mean of the importance measures of the component skills as in Table 4. In column (2), the skills indices are standardised to zero mean and unit variance within years, which allows for any aggregate rescaling between years due to incumbent-analyst changes. Columns (3) and (4) repeat these specifications except that Principal Components Analysis (PCA) is used to aggregate the skills measures into each category and the scores produced by the first principal component are used as the skill measures. 54% of the variance in analytic skills is explained by the first principal component. For interpersonal and physical skills, the respective figures are 76% and 71%.

In Panel B we incorporate the skill levels as well as the skills importance measures. As noted above, the levels and importance measures tend to be highly correlated (Handel, 2016a). We again compute a mean based measure and a PCA based measure. Rather than a simple mean of all importance and levels measures, we follow the approach of Blinder (2009) and calculate a weighted average, using Cobb-Douglas weights of 2/3 and 1/3 respectively for the importance and levels measures.

The results presented in Table 5 provide evidence that the way in which we aggregate the skills information from 35 skills measures in the raw data to our three summary indices does not have an impact on our findings. In each of the 8 sets of estimates, the coefficients for analytical skills and physical skills are consistently statistically significant positive and negatively at the 1% level. Notably, incorporating skills level information produces significantly positive coefficient estimates for interpersonal skills for the Cobb-Douglas weighted means specification in columns (5) and (6), although not in the PCA specification in columns (7) and

(8). The magnitudes of the raw skill measures are not directly comparable as different aggregation methods produce variables on different scales. However, the coefficients on the standardised transformations are comparable, and they indicate that the magnitude of the effects of skills on earnings is similar across the four different aggregation methods. *Ceteris paribus*, occupations in which employees utilise a 1 standard deviation higher level of analytical skills are associated with 7-10% higher hourly wages. For physical skills, a 1 standard deviation higher level of physical skills is associated with 3-4% lower hourly wages.

Returning to our chosen method of aggregation for our main variables of interest (i.e. means of only the importance measures of skills), we report a range of other robustness checks in Table 6. These investigate the sensitivity of our results to options we have chosen when constructing the database. In particular, we check the robustness of the results to using the LFS rather than the ASHE as the source of data on wages, using the occupational mean of log wages rather than the log of occupational mean wages, to using the 'raw' skill measures in O*NET (i.e. without any adjustment for the change from job incumbents' measures to job analysts' measures of skills in the early part of the sample period), and finally the choice of correspondence table that we use to convert information at the SOC2000 level to SOC2010.

Column (1) In Table 6 repeats our baseline results from Table 2, column (4). Comparing these results to those of column (2), it is clear that estimates of the return to skill are not sensitive to whether we attempt to adjust for the mix of job incumbents and job analysts providing skills information in the O*NET data.

Our preferred measure of wages is derived from ASHE for the reasons stated in subsection 3.1 above. As an alternative, we can use log mean hourly wages derived from LFS data. We could also use mean log wages for LFS data since we have individual earnings in these data, and the aggregation of individual log earnings functions to occupational earnings functions yields this as the 'correct' dependent variable[12]. Our results are not substantially affected by how wages are aggregated to the occupation level, or the dataset we source the wage information from. Column (3) is directly comparable to column (1) as both of these use log mean wages, and we find no significant difference between the two, indicating the choice

[12] However, cell-mean regressions of this kind (e.g. Blanchflower et al, 1996, and Dearden et al, 2006) frequently use log mean wages rather than mean log wages as the dependent variable.

between LFS and ASHE wages would appear to have no substantive bearing on our results. Comparing columns (3) and (4), using mean log wages rather than log mean wages does attenuate the magnitude of the returns to analytical skills and physical skills slightly, although the main conclusions are unaffected.

The baseline results presented above use the average across the three different SOC2000-SOC2010 correspondence tables provided by the ONS to convert between SOC2000 and SOC2010. In columns (5), (6) and (7), we utilise each of the three weighting matrices separately to transform the 2002-2010 SOC2000 data to SOC2010. The three correspondence tables produce very similar magnitudes of estimated returns for the skills measures as when using the mean of the three tables, and there are no statistically significant differences from the baseline.

The changes to data source and construction of correspondences we have investigated in Table 6 do not alter any of our main findings or conclusions. Our estimates of the returns to skill remain the same in terms of sign and statistical significance, and the magnitudes are robust to these choices.

4.3.2 Specification and Estimation of the Earnings Function

Table 7 presents a set of robustness checks for our main results, in this case focussing on the robustness to our chosen specification for the earnings function. Our main results, repeated once again in the first column of Table 7, are estimated using gender-specific variables combined using employment share weights. In columns (2) and (3) of Table 7, we compare the results when all variables are based on, respectively, males only and females only. We find that both male and female occupational average log wages are associated with analytical, interpersonal and physical skills in the same way – positively correlated with analytical skills and negatively correlated with physical skills. For females, there is also a weakly negative association with interpersonal skills. The magnitudes of these effects, in both cases, are larger than the aggregate baseline results. The difference in results will in part reflect the fact that when splitting by gender we lose occupations with very small or no employment. This is particularly the case for females, where we lose around one fifth of our sample observations. In column (4), we restrict the individual observations used to construct our occupation level variables to those where the employees are full time only. Relative to our baseline results in column (1), we find larger effects of skills on earnings.

In column (5) we report estimates of our standard specification with the addition of 1-digit SOC dummies. As we would expect, this decreases the magnitudes of the estimated returns to skills, but the general conclusions remain unchanged. We also experimented with 2-digit and 3-digit occupation dummies and still found positive and significant coefficients for analytical skills and negative and significant coefficients for physical skills.

Given the multiple changes in SOC classifications in the UK and the US (both O*NET and SOC) prior to 2010, together with the changing incumbent-analyst ratio in reporting the measures of skills (Appendix C), we re-estimated the returns to skills for the period 2011-2016 only since this period is unaffected by any of the changes in SOC or in the reporting of skills in O*NET. The results of this exercise are shown in column (6). The average returns to analytical skills for this subperiod are rather higher than the average for the whole 2002-2016 sample period, exactly as suggested by Figure 4. We also now find much stronger positive and now statistically significant returns to interpersonal skills when focussing on the later period only. Again, this is consistent with Figure 4 in which the returns to interpersonal skills are positive and increasing after 2010. Together, these findings suggest that the additional manipulations required in order to construct SOC2010-consistent data for 2002-2010 are not unduly responsible for the results obtained.

The final specification issue is the use of OLS when it could be argued we should use weighted least squares since we are estimating group mean regressions. There is some debate in the literature about the necessity of weighting, but here we simply investigate if it makes a difference to the estimated returns. We follow the approach of Dickens (1990) in weighting our regressions. In group mean regressions we cannot simply weight by the square root of the cell size. This is because individuals within groups (in this case occupations) are likely to share unobserved characteristics, in which case the regression error term will consist of an individual error component plus a shared group component. The variance of the group-mean regression residuals is, in this case, given by equation (3):

$$Var(\bar{e}_j) = \sigma_\gamma^2 + \frac{\sigma_u^2}{N_j} \tag{3}$$

The error variance for group j is given by the shared group component, σ_γ^2, plus the individual component, σ_u^2/N_j. If the two variances are equal, and the group sizes are large then there will be little variation in the overall variances and heteroscedasticity will be minimal. In this case, weighting by the square root of group size will introduce substantial heteroscedasticity if there are large differences in group size. If, however, σ_γ^2 is zero or small relative to σ_u^2, then large variation in N_j will result in considerable heteroscedasticity. In this case, the regressions should be weighted.

Table 8 reports the results of this exercise. First, we estimate our standard specification by WLS, weighting by the square root of group size (in this case, employment in the occupation). The results are reported in Table 8, column (2). We then test for heteroscedasticity by regressing the squared residuals on employment in the occupation. The coefficient on employment in this regression is statistically significant, suggesting the presence of a group component in the error term. We then estimate the following regression, where the group-specific residual is regressed on a constant and the inverse of employment in the occupation.

$$\hat{\varepsilon}_j^2 = \alpha + \delta\left(\frac{1}{N_j}\right) \tag{4}$$

The parameters α and δ are estimates of the corresponding error variance components in equation (3). These estimates, $\hat{\sigma}_\gamma^2$ and $\hat{\sigma}_u^2$, are used to construct the weights $1/(\sigma_\gamma^2 + \sigma_u^2/N_j)$. The earnings function is then re-estimated using these weights, from which the new error component variances can be constructed to again re-estimate the earnings function. This iterative process continues until both coefficients in the residual regression (i.e. equation (4)) are identical (to 5 decimal places) between two iterations. This convergence occurs at the 4th iteration (convergence occurs to 3 decimal places after the first iteration), and it is these results presented in column (3) of Table 8.

The coefficient estimates in column (3) do not differ significantly from our main results reported in column (1). Our results are therefore not sensitive to whether we use weights the earnings regressions.

Taken together, our comprehensive set of robustness checks show that our main estimates presented in section 4.2 are highly robust and stable. Despite the uncertainties around how to deal with the issue of changing between job incumbents and job analysts' measurement of of O*NET skills, the choice of data source for wages, how to aggregate the skills measures, and how to convert between SOC2000 to SOC2010, we find that our results for the estimated returns to skills are not sensitive to these various decisions made in the construction of the occupational skills profiles.

5. Summary and Conclusions

This paper exploits the O*NET system for measuring and assessing skills in combination with a detailed and systematic matching and mapping exercise to construct occupational skills profiles for the UK. These provide a much more detailed depiction of skills than is available through more conventional measures of skills such as educational qualification or occupation classification. Given the inherent weaknesses in these standard proxies for skills, the methodology developed in this paper has the potential to substantially advance our understanding and knowledge of the nature of skills demand in the UK.

We computed three broad indices of skills at the 4-digit occupational level for the UK for 2002-2016 and examined the evolution of the utilisation of these indices as the occupational composition and skills content of jobs has changed over time. Strong secular growth in the utilisation of analytical and interpersonal skills and declining usage of physical skills is consistent with other literatures which have documented the changing skill content of jobs. We then estimated earnings functions which control for education qualifications, gender, firm size and other established determinants of differences in earnings, in order to investigate the conditional returns to these skills. High and statistically significant and increasing returns over time to analytical skills, particularly for full-time workers and for women, was contrasted with lower, but still increasing returns to interpersonal skills, especially since 2010. Thus, while there is no evidence in the UK for any 'reversal' in the returns to more cognitive skills (Beaudry et al, 2016), the latter finding is consistent with Deming (2017) for the US and Edin et al (2017) for Sweden who also document growing importance of 'social' or 'interaction' skills. Finally, the returns to physical skills was found to be significantly below zero for the whole of the period.

The findings demonstrate the increasing importance of work-related skills and attributes for individuals' earnings, over and above their educational qualifications and, in particular, for higher levels of analytical skills and interpersonal skills in the workplace. Our interpretation of the increased utilisation coupled with increasing returns to analytic and interpersonal skills is that the UK is experiencing significantly increased demand for these skills in the labour market. These findings have clear implications for future education policy.

There are a number of possible additional uses for the occupational skills profiles developed in this paper. These include providing information to IAG (Information, Advice and Guidance) practitioners and careers advisors on the types of skills that are necessary for, and useful in employment, and in estimating future skills demand by linking the skills measures to occupational employment projections (e.g. UKCES, 2016).

References

Abraham, Katherine G and James R Spletzer (2009), "New evidence on the returns to job skills", *American Economic Review Papers and Proceedings*, 99(2), 52-57.

Autor, David H, Frank Levy and Richard J Murnane (2003), "The Skill Content of Recent Technological Change: An Empirical Investigation", *Quarterly Journal of Economics*, 118(4), 1279-1333.

Autor, David H, Lawrence Katz and Mekissa S Kearney (2008), "Trends in US wage inequality: Revisiting the revisionists", *Review of Economics and Statistics*, 90(2), 300-323.

Autor, David H and Michael J Handel (2013), "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages", *Journal of Labor Economics*, 31(2), S59-S97.

Barro, Robert and Jong-Wha Lee (2013), "A New Data Set of Educational Attainment in the World, 1950-2010", *Journal of Development Economics*, 104, 184-198.

Beaudry, Paul, David A Green and Benjamin M Sand (2016), "The Great Reversal in the Demand for Skill and Cognitive Tasks", *Journal of Labor Economics*, 34(10), S199-S247.

Berman, Eli, John Bound and Zvi Griliches (1994), "Changes in the Demand for Skilled Labor within US Manufacturing: Evidence from the Annual Survey of Manufacturers", *Quarterly Journal of Economics*, 109(2), 367-397.

Berman, Eli, John Bound and Stephen Machin (1998), "Implications of Skill-Biased Technical Change: International Evidence", *Quarterly Journal of Economics*, 113(4), 1245-1279.

BIBB (Federal Institute for Vocational Education and Training) (no date), *BIBB/BAuA Employment Survey of the Working Population on Qualification and Working Conditions in Germany*, https://www.bibb.de/en/2815.php last accessed 09 February 2019.

Black, Sandra E and Alexandra Spitz-Oener (2010), "Explaining Women's Success: Technological Change and the Skill Content of Women's Work", *Review of Economics and Statistics*, 92(1), 187-194.

Blanchflower, David G, Andrew J Oswald and Peter Sanfey (1996), "Wages, profits, and rent-sharing" *Quarterly Journal of Economics,* 111(1), 227-251.

Blinder, Alan S (2009), "How many U.S. jobs might be offshorable?", *World Economics*, 10(2), 41-78.

BLS (Bureau of Labor Statistics) (2017), *Occupational Employment Statistics (OES) Survey*, https://www.bls.gov/oes/ last accessed 09 February 2019.

BLS (Bureau of Labor Statistics) (2018), *Standard Occupational Classification 2018*, https://www.bls.gov/soc/ last accessed 09 February 2019.

Card, David (2000), "The causal effect of education on earnings", in Ashenfelter, Orley and David Card (eds), *The Handbook of Labor Economics*, Elsevier Scientific Publishing, Amsterdam.

CASCOT (no date) (Computer Assisted Structured Coding Tool) http://www2.warwick.ac.uk/fac/soc/ier/software/cascot/ last accessed 09 February 2019.

Dearden, Lorraine, Howard Reed and John Van Reenen (2006), "The impact of training on productivity and wages: Evidence from British panel data", *Oxford Bulletin of Economics and Statistics,* 68(4), 397-421.

DEE (Department for Education and Employment) (2000), *Skills for All*, Report of the National Skills Task Force.

Deming, David J (2017), "The Growing Importance of Social skills in the Labour Market", *Quarterly Journal of Economics*, 132(4), 1593-1640.

DBIS (Department for Business, Innovation and Skills) (2009), *Skills for Growth – The National Skills Strategy*, London: DBIS Department for Business, Innovation and Skills.

Dickens, William T (1990), "Error components in grouped data: is it ever worth weighting?", *Review of Economics and Statistics,* 72(2), 328-333.

Dickerson, Andrew and Francis Green (2004), "The growth and valuation of computing and other generic skills", *Oxford Economic Papers*, 56(3), 371-406.

Dickerson, Andrew, Rob Wilson, Genna Kik and Debra Dhillon (2012), *Developing Occupational Skills Profiles for the UK: A Feasibility Study*, Evidence Report 44, February 2012, UK Commission for Employment and Skills.

Edin, Per-Anders, Peter Fredriksson, Martin Nybom and Björn Öckert (2017), "The Rising Return to Non-Cognitive Skill", IZA DP No. 10914.

Felstead, Alan, Duncan Gallie, Francis Green and Ying Zhou (2007), *Skills at Work 1986-2006*, ESRC Centre on Skills, Knowledge and Organisational Performance, SKOPE, Universities of Oxford and Cardiff.

Felstead, Alan, Duncan Gallie, Francis Green and Hande Inanc (2013), *Skills At Work In Britain: First Findings from the Skills and Employment Survey 2012,* London: Centre for Learning and Life Chances in Knowledge Economies and Societies, Institute of Education.

Fine, Sidney (1955), “A structure of worker functions”, *Personnel and Guidance Journal*, 34, 66-73.

Fine, Sidney and Steven Cronshaw (1999), *Functional job analysis: A foundation for human resources management,* Mahwah, NJ: Lawrence Erlbaum Associates.

Goos, Maarten and Alan Manning (2007), “Lousy and Lovely Jobs: The Rising Polarisation of Work in Britain”, *Review of Economics and Statistics*, 89(1), 118-113.

Goos, Maarten, Alan Manning and Anna Salomons (2009), “Job polarization in Europe”, *American Economic Review Papers and Proceedings*, 99(2), 58-63.

Goos, Maarten, Alan Manning and Anna Salomons (2014), "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring”, *American Economic Review* 104(8), 2509-2526.

Green, Francis (2006), *Demanding Work: The Paradox of Job Quality in the Affluent Society*, Princeton: Princeton University Press.

Green, Francis (2013), *Skills and Skilled Work: An Economic and Social Analysis*, Oxford: Oxford University Press.

Gregg, Paul, Stephen Machin and Mariña Fernández-Salgado (2014), “Real Wages and Unemployment in the Big Squeeze”, *Economic Journal*, 124(576), 408–432.

Handel, Michael J (2016a), “The O*NET content model: strengths and limitations”, *Journal for Labour Market Research*, 49(2), 157-176.

Handel, Michael J (2016b), “What do people do at work?”, *Journal for Labour Market Research*, 49(2), 177-197.

Hanushek, Eric, Guido Schwerdt, Simon Wiederhold and Ludger Woessmann (2015), “Returns to skills around the world: Evidence from PIAAC”, *European Economic Review*, 73(C), 103-130.

Howell, David R and Edward N Wolff (1991), “Trends in the Growth and Distribution of Skills in the US Workplace, 1965-1985”, *Industrial and Labour Relations Review*, 44(3), 486-502.

Howell, David R and Edward N Wolff (1992), “Technical Change and the Demand for Skills for US Industries”, *Cambridge Journal of Economics*, 16, 127-146.

Ingram, Beth and George R Neumann (2006), “The returns to skill”, *Labour Economics*, 13, 35-59.

Katz, Lawrence and Kevin M Murphy (1992), "Changes in Relative Wages 1963-1987: Supply and Demand Factors", *Quarterly Journal of Economics*, 107(1), 35-78.

Leitch, Lord Sandy (2006), *Prosperity For All in the Global Economy – World Class Skills*, London: HM Treasury.

LMI for All (2013), *Developing a Careers LMI Database: Phase 2A Report,* Career Database Project Team, Warwick Institute for Employment Research, UK Commission for Employment and Skills, July 2013.

LMI for All (2015), *Developing a Careers LMI Database: Final Report*, Career Database Project Team, Warwick Institute for Employment Research, UK Commission for Employment and Skills, July 2015.

LMI for All (2016), http://www.lmiforall.org.uk/ last accessed 09 February 2019.

McIntosh, Steven and Anna Vignoles (2001), "Measuring and assessing the impact of basic skills on labour market outcomes", *Oxford Economic Papers*, 53(3), 453–481.

Mumford, Michael, Norman Peterson and Ruth Childs (1999), "Basic and cross-functional skills", pp.46-69 in Peterson, Norman G, Michael D Mumford, Walter C Borman, P. Richard Jeanneret and Edwin A Fleishman (1999), *An occupational information system for the 21st century: The development of O*NET*, Washington, DC: American Psychological Association.

OECD (2000), *Literacy in the Information Age: Final report of the International Adult Literacy Survey*, Paris: Organisation for Economic Co-operation and Development and Statistics Canada.

OECD (2012), *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of adult skills*, Paris: OECD Publishing.

OECD (2016), *Skills Matter: Further Results from the Survey of Adult Skills,* OECD Skills Studies, Paris: OECD Publishing.

Oesch, Daniel (2013), *Occupational Change in Europe: How technology and education transform the job structure*, Oxford: Oxford University Press.

ONS (2012), *Relationship between: Standard Occupational Classification 2010 (SOC2010) and Standard Occupational Classification 2000 (SOC2000)*, Classification and Harmonisation Unit, User Guide 2010, available at: https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010, last accessed: 09 February 2019.

ONS (2014), *Review of the Labour Force Survey*, National Statistics Quality Review: Series (2), Report No. 1, Office for National Statistics.

Peterson, Norman G, Michael D Mumford, Walter C Borman, P. Richard Jeanneret and Edwin A Fleishman (1999), *An occupational information system for the 21st century: The development of O*NET*, Washington, DC: American Psychological Association.

Rohrbach-Schmidt, Daniela and Michael Tiemann (2016), “Educational (Mis)match and skill utilization in Germany: Assessing the role of worker and job characteristics”, *Journal for Labour Market Research*, 49(2), 99-119.

SOC2010 (2010), *Volume 1: Structure and Descriptions of Unit Groups*, Basingstoke: Palgrave MacMillan.

Spence, Michael (1973), "Job Market Signaling", *Quarterly Journal of Economics*, 87(3), 355-374.

Spitz-Oener, Alexandra (2006), “Technical change, job tasks and rising educational demands: Looking outside the wage structure”, *Journal of Labor Economics*, 24(2), 235–70.

Taylor, Paul J, Wen-Dong Li, Kan Shi, and Walter C Borman (2008), “The transportability of job information across countries”, *Personnel Psychology*, 61(1), 69-111.

Tippins, Nancy and Margaret L Hilton (eds.) (2010), *A Database for a Changing Economy: Review of the Occupational Information Network (O*NET)*, National Research Council, National Academies Press, Washington DC.

Tsacoumis, Suzanne and Chad H Van Iddekinge (2006), “A Comparison of Incumbent and Analyst ratings of O*NET skills”, Human Resources Research Organisation, available at: https://www.onetcenter.org/reports/SkillsComp.html, last accessed 09 February 2019.

UKCES (UK Commission for Employment and Skills) (2009), *Ambition 2020: World Class Skills and Jobs for the UK*, London: UK Commission for Employment and Skills.

UKCES (UK Commission for Employment and Skills) (2010), *Skills for jobs: today and tomorrow, Volume 2: Evidence Report*, London: UK Commission for Employment and Skills.

UKCES (UK Commission for Employment and Skills) (2014), *The Future of Work: Jobs and Skills in 2030*, London: UK Commission for Employment and Skills.

UKCES (UK Commission for Employment and Skills) (2016), *Working Futures 2014-2024*, London: UK Commission for Employment and Skills.

US Department of Labor (1991), *Dictionary of Occupational Titles: Revised fourth edition*, Washington, DC: Employment and Training Administration.

US Department of Labor (2005), *O*NET Data Collection Program*, Office of Management and Budget Clearance Package Supporting Statement and Data Collection Instruments, Washington D.C. Employment and Training Administration.

Table 1: Summarising Skills, Tasks and Work Activities: Examples from the Literature

Reference	Taxonomy	Data	Measures/Methods	Notes/Findings
Autor, Levy and Murnane (*QJE* 2003)	Non-routine analytic tasks Non-routine interactive tasks Routine cognitive tasks Routine manual tasks Non-routine manual tasks (omitted from most analysis)	DOT (US Dictionary of Occupational Titles) 1977 and 1991	(i) Single DOT variable for each task measure (ii) Principal components for 4 selected DOT variables for each task measure	• Computers have substituted routine tasks and complemented non-routine tasks. • This shift in job tasks can help explain the increased returns to college education. • Within-occupation change is a significant component of the change in task demand.
Howell and Wolff (*ILRR* 1991 and *CJE* 1992)	Cognitive skills Interactive/People skills Motor skills	DOT 1977	Cognitive skills: factor analysis over 46 DOT variables Interactive skills: single DOT variable Motor skills: factor analysis over 3 DOT variables	• Suggests education is a poor measure of workforce skills. • Technical change helps to explain increasing cognitive skill requirements and changing occupational distribution of employment.
Autor and Handel (*JLE* 2013)	Cognitive tasks Interpersonal tasks Physical job tasks ('data- people-things' as used in DOT)	Princeton Data Improvement Initiative (PDII) O*NET v.14 40 items from a number of domains (work activities, skills, knowledge, work context)	Additive multi-item scales - O*NET items collated into 10 measures (minimum 2 items, maximum 8 items)	• Job tasks vary within occupations (by race, gender and English language proficiency) as well as between occupations. • Tasks at both individual and occupational level are important predictors of hourly wages.
Abraham and Spletzer (*AER* 2009)	Analytic activities Interpersonal activities Physical activities	O*NET v. 13 (June 2008) 41 work activities	Analytic: average of 2 O*NET activities Interpersonal: average of 2 O*NET activities Physical: 1 O*NET activity	• Jobs that require more analytical activity pay significantly higher wages, while those that require more interpersonal and physical activity pay lower wages.

Reference	Taxonomy	Data	Measures/Methods	Notes/Findings
Black and Spitz-Oener (*REStats* 2010), Spitz-Oener (*JLE* 2006)	Non-routine analytic tasks Non-routine interactive tasks Routine cognitive tasks Routine manual tasks Non-routine manual tasks (i.e. based on ALM, 2003)	West Germany Qualification and Career Survey 1979-99	Task measure is the proportion of job activities in each task group	• Substantial relative decline in routine task input for women driven by technological change has significantly contributed toward the narrowing of the gender pay gap.
Goos, Manning and Salomons (*AER* 2009 and *AER* 2014)	Abstract tasks (intense in non-routine cognitive skills) Routine tasks (intense in cognitive and non-cognitive routine skills) Service tasks (intense in non-routine, non-cognitive skills)	O*NET v. 11 (2006) 96 items selected from a range of domains	(i) Abstract=first principal component of 72 O*NET items; Routine=first principal component of 16 O*NET items; Service=first principal component of 8 O*NET items (ii) Principal components of all items together – identifies 2 components corresponding to the 'Routine', and the 'Abstract and Service' dimensions	• Evidence of job polarization across Europe. • Technologies are becoming more intensive in non-routine tasks at the expense of routine tasks. • Evidence for off-shoring and inequality driving polarisation is much weaker.

Table 2: Correspondence between UK and US measures of skills using PIAAC

	Skills		
	Numeracy	Literacy	Problem Solving
MAPE	6.11%	4.09%	3.54%
RMSE	1.59	1.59	3.11
RMSE (normalised)	0.01	0.01	0.01

Table 3: Decomposition of changing skill utilisation 2002-2016

	Aggregate change in skills 2002-16	Decomposition of changing skills utilisation		
		Between occupations	Within occupations	Total Change
1-digit SOC2010 *(9 categories)*		%	%	
Analytic skills	+10%	24	76	100%
Interpersonal skills	+23%	11	89	100%
Physical skills	−14%	10	90	100%
2-digit SOC2010 *(25 categories)*		%	%	
Analytic skills	+10%	25	75	100%
Interpersonal skills	+23%	12	88	100%
Physical skills	−14%	14	86	100%
3-digit SOC2010 *(90 categories)*		%	%	
Analytic skills	+10%	26	74	100%
Interpersonal skills	+23%	15	85	100%
Physical skills	−14%	17	83	100%
4-digit SOC2010 *(369 categories)*		%	%	
Analytic skills	+10%	18	82	100%
Interpersonal skills	+23%	11	89	100%
Physical skills	−14%	24	76	100%

Note:

1. Decomposition of the overall change in skill utilisation between 2002 and 2016 into between-occupation and within-occupation changes. See text, equation (2), for details.

Table 4: Returns to Skills 2002-2016

Dependent Variable: Log Average Hourly Real Wages	(1)	(2)	(3)	(4)
Analytic skills	0.839***		0.191***	0.172***
	(0.013)		(0.013)	(0.014)
Interpersonal skills	-0.225***		-0.032***	0.004
	(0.011)		(0.009)	(0.010)
Physical skills	-0.150***		-0.057***	-0.058***
	(0.007)		(0.007)	(0.007)
Highest Qual NQF 4+		1.130***	0.869***	0.899***
		(0.025)	(0.029)	(0.029)
Highest Qual NQF 3		0.643***	0.462***	0.489***
		(0.032)	(0.034)	(0.033)
Highest Qual NQF 2		0.584***	0.422***	0.438***
		(0.043)	(0.044)	(0.044)
Highest Qual below NQF 2		0.236***	0.195***	0.206***
		(0.051)	(0.050)	(0.049)
Highest Qual Apprenticeship		0.542***	0.584***	0.599***
		(0.049)	(0.049)	(0.048)
Female		-0.312***	-0.289***	-0.298***
		(0.011)	(0.012)	(0.012)
Age		0.133***	0.121***	0.117***
		(0.004)	(0.004)	(0.004)
Age Squared		-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)
Firm Size 25-49		0.016	0.041	0.034
		(0.038)	(0.037)	(0.037)
Firm Size 50-499		0.066***	0.082***	0.093***
		(0.018)	(0.018)	(0.018)
Firm Size 500+		0.340***	0.341***	0.357***
		(0.021)	(0.021)	(0.021)
Public Sector		-0.154***	-0.130***	-0.160***
		(0.011)	(0.012)	(0.012)
Constant	1.435***	-1.222***	-1.071***	-1.310***
	(0.028)	(0.116)	(0.117)	(0.116)
Region dummies		✓	✓	✓
Year dummies				✓
N	5156	5172	4944	4944

Notes:

1. The dependent variable is log mean real hourly wages.
2. Standard errors in parentheses: * p<0.10; ** p<0.05; *** p<0.01.
3. Base category for highest qualification is other qualifications or no qualifications. Base category for firm size is less than 25 employees.

Table 5: Robustness to aggregation methods

	Panel A: Importance Measures Only				Panel B: Importance and Levels Measures			
	Mean		PCA		C-D weighted mean		PCA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Raw	Std.	Raw	Std.	Raw	Std.	Raw	Std.
Analytic skills	0.172***	0.077***	0.027***	0.095***	0.007***	0.072***	0.017***	0.086***
	(0.014)	(0.006)	(0.002)	(0.006)	(0.001)	(0.007)	(0.001)	(0.007)
Interpersonal skills	0.004	0.007	-0.004	-0.008	0.002*	0.020***	0.002	0.008
	(0.010)	(0.005)	(0.003)	(0.006)	(0.001)	(0.006)	(0.002)	(0.007)
Physical skills	-0.058***	-0.037***	-0.014***	-0.034***	-0.008***	-0.043***	-0.010***	-0.033***
	(0.007)	(0.004)	(0.002)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)
Education controls	✓	✓	✓	✓	✓	✓	✓	✓
Other controls	✓	✓	✓	✓	✓	✓	✓	✓
Region dummies	✓	✓	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓	✓	✓
N	4944	4944	4944	4944	4934	4934	4944	4944

Notes:

1. The dependent variable is log average real hourly wages.
2. Standard errors in parentheses: * p<0.10; ** p<0.05; *** p<0.01.
3. All regressions in this table are estimated using the same specification as in column (4) of Table 2.
4. Panel A reports results for skill aggregations which only use the importance measure of the 35 source skills in the aggregation. In Panel B the aggregations are based on both importance and levels measures.

Table 6: Robustness to alternative data sources and transformations

	(1) Baseline	(2) Raw Skills	(3) LFS Log Mean	(4) LFS Mean Log	(5) LFS96_97	(6) Census01	(7) LFSJM07
Analytic skills	0.172***	0.173***	0.170***	0.153***	0.167***	0.165***	0.175***
	(0.014)	(0.014)	(0.013)	(0.012)	(0.014)	(0.014)	(0.014)
Interpersonal	0.004	-0.013	-0.015	-0.016*	0.003	0.008	0.002
	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)
Physical skills	-0.058***	-0.065***	-0.065***	-0.051***	-0.056***	-0.056***	-0.056***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Education controls	✓	✓	✓	✓	✓	✓	✓
Other controls	✓	✓	✓	✓	✓	✓	✓
Region dummies	✓	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓	✓
N	4944	4944	5060	5060	4887	4920	4930

Notes:
1. The dependent variable is log average real hourly wages.
2. Standard errors in parentheses: * p<0.10; ** p<0.05; *** p<0.01.
3. Column (1) regression coefficients repeat the specification reported in column (4) of Table 2. Column (2) uses raw skills data, not corrected for changes in the incumbent and analyst evaluations. Column (3) uses the log of occupational mean wages as an alternative dependent variable and column (4) uses the occupational mean of log wages, using LFS data in both cases. Columns (5) to (7) re-estimates with data which is converted from SOC2000 to SOC2010 with weights using each of the 3 dual-coded datasets separately. See text for details.

Table 7: Robustness to earnings function specification

	(1) Baseline	(2) Male	(3) Female	(4) Full Time	(5) 1-Digit SOC	(6) 2011-16
Analytic skills	0.172***	0.290***	0.375***	0.236***	0.119***	0.308***
	(0.014)	(0.017)	(0.016)	(0.015)	(0.014)	(0.031)
Interpersonal	0.004	-0.018	-0.027**	-0.004	0.004	0.066***
	(0.010)	(0.013)	(0.013)	(0.011)	(0.010)	(0.025)
Physical skills	-0.058***	-0.095***	-0.141***	-0.072***	-0.042***	-0.055***
	(0.007)	(0.008)	(0.008)	(0.007)	(0.007)	(0.013)
Education controls	✓	✓	✓	✓	✓	✓
Other controls	✓	✓	✓	✓	✓	✓
Region dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
N	4944	4362	3774	4647	4944	1918

Notes:
1. The dependent variable is log average real hourly wages.
2. Standard errors in parentheses: * $p<0.10$; ** $p<0.05$; *** $p<0.01$.
3. The column (1) regression coefficients repeat the specification reported in column (4) of Table 2. Column (2) constructs the outcome and independent variables from male observations only. Column (3) constructs the outcome and independent variables from female observations only. Column (4) constructs the outcome and independent variables from full-time workers observations only. Columns (5) includes 1-digit SOC occupation fixed effects. Column (6) estimates only for 2011 to 2016. See text for details.

Table 8: Weighted least squares estimates

	(1) OLS	(2) WLS	(3) Dickens Iterative WLS
Analytic skills	0.172***	0.163***	0.175***
	(0.014)	(0.013)	(0.014)
Interpersonal	0.004	0.011	0.003
	(0.010)	(0.010)	(0.010)
Physical skills	-0.058***	-0.084***	-0.066***
	(0.007)	(0.007)	(0.007)
Education controls	✓	✓	✓
Other controls	✓	✓	✓
Region dummies	✓	✓	✓
Year dummies	✓	✓	✓
	ε^2	ε^2	ε^2
1/N	0.020***	0.022***	0.021***
	(0.005)	(0.006)	(0.005)
Constant	0.025***	0.025***	0.024***
	(0.001)	(0.001)	(0.001)
N	4944	4944	4944

Notes:
1. The dependent variable is log average real hourly wages.
2. Standard errors in parentheses: * p<0.10; ** p<0.05; *** p<0.01.
3. The lower panel regresses the squared residuals from the regression in the upper panel on a constant and the inverse of employment in the occupation. See text for details.

Figure 1: Comparing UK and US numeracy skills using PIAAC

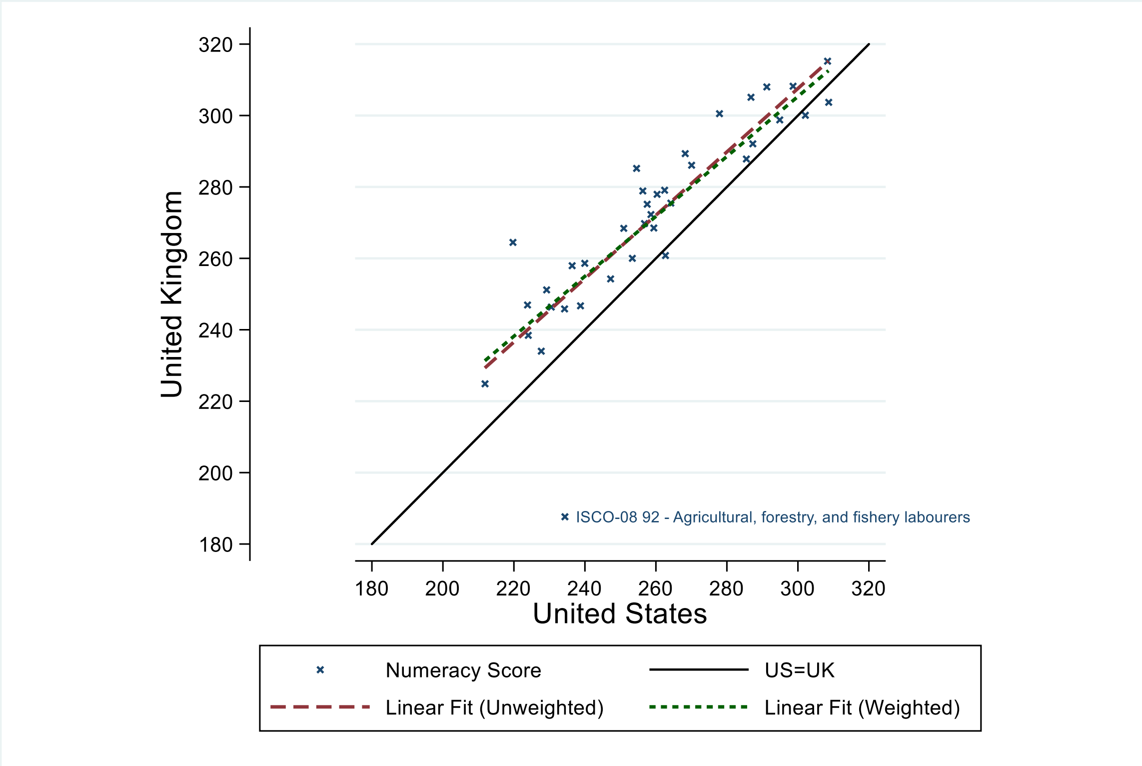


Figure 2: Comparing UK and US literacy skills using PIAAC

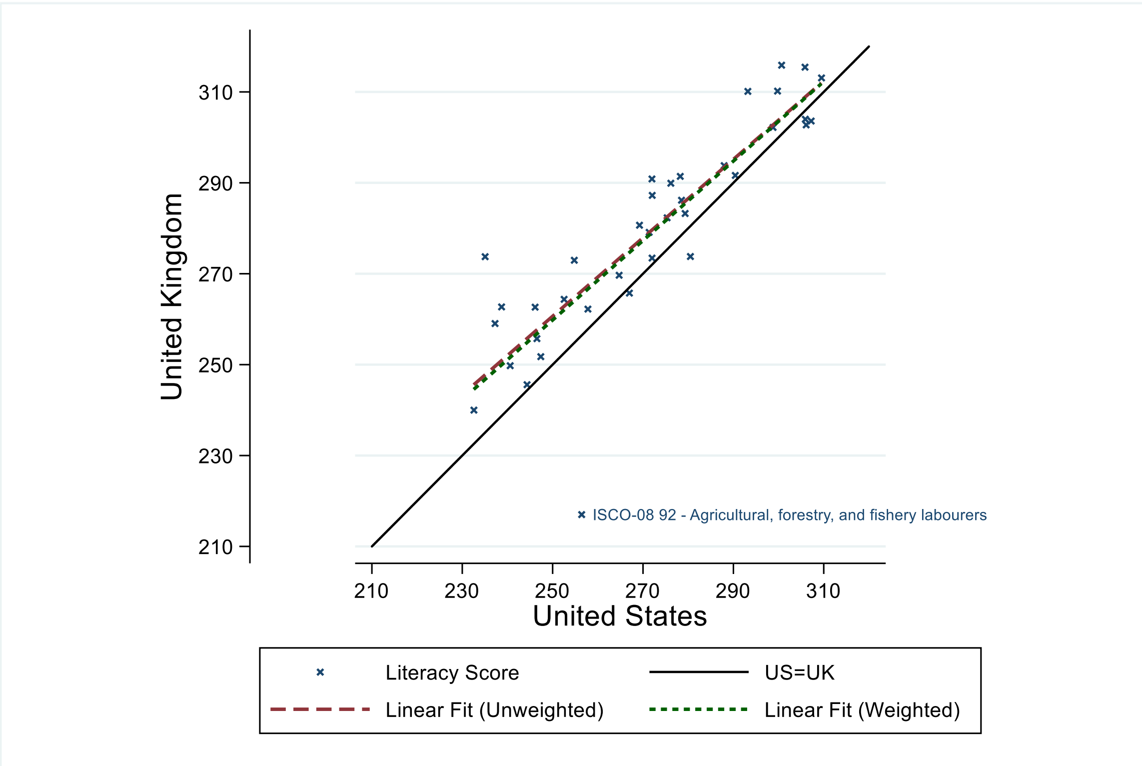
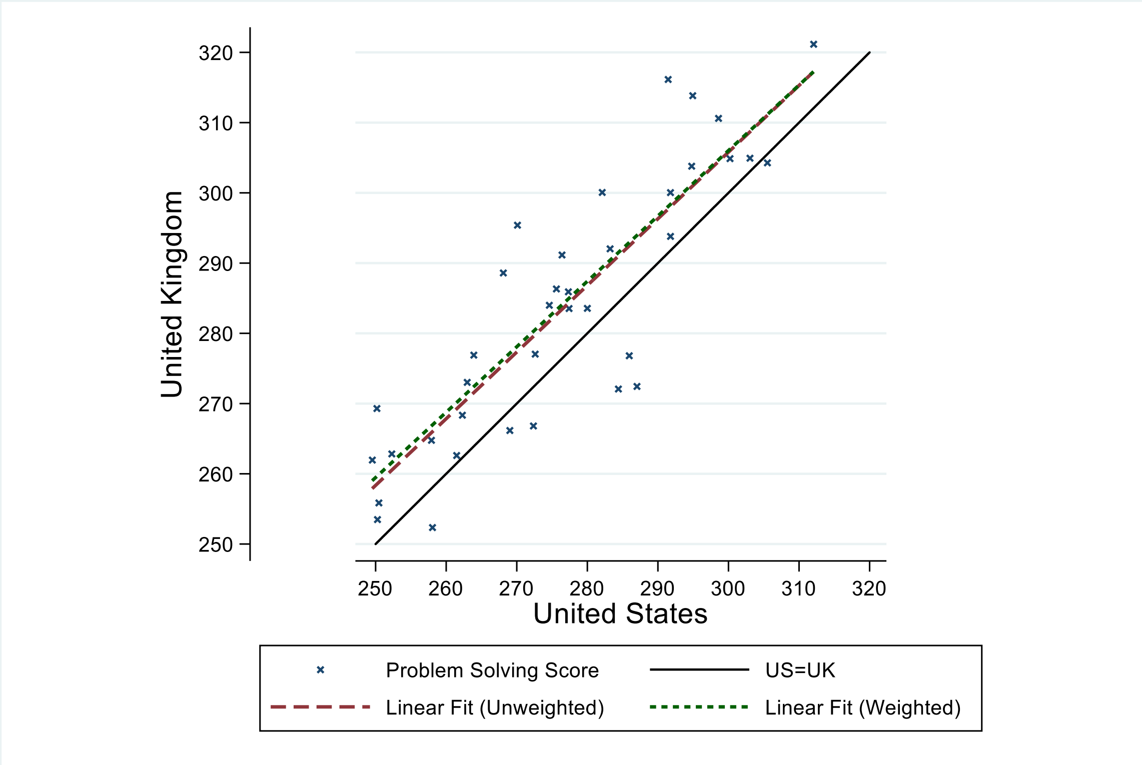


Figure 3: Comparing UK and US problem solving skills using PIAAC

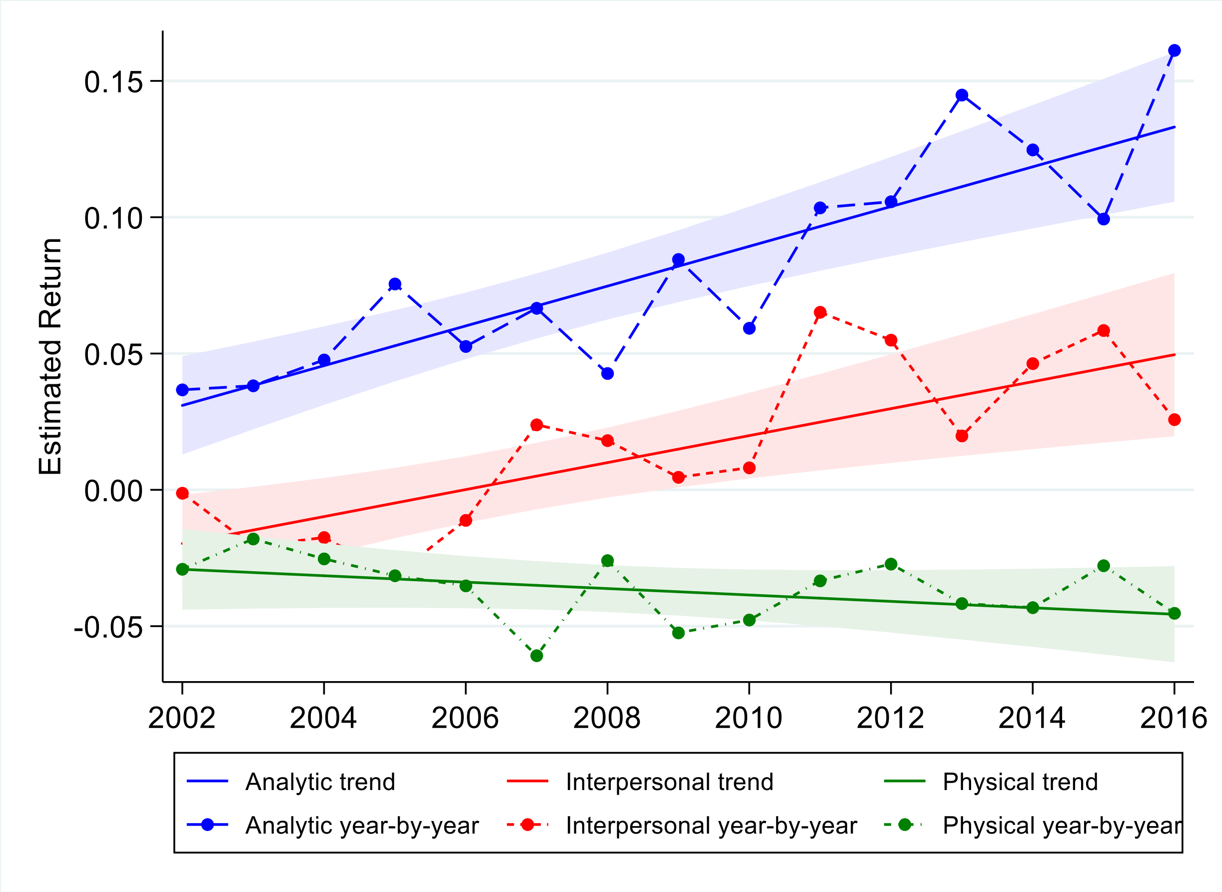


Notes:

1. Occupations ISCO-08 94: Food Preparation Assistants and ISCO-08 95: Street and Related Sales and Service Workers are omitted due to having no UK skill data for these two occupations.

2. The weighted regressions account for the differences in occupation size. See text for details.

Figure 4: Trends in the Returns to Skills 2002-2016



Notes:

1. These are regression coefficients using the specification in Table 2, column (4), supplemented by interactions of each of the three skills indices with: (i) a linear time trend (solid lines) and (ii) year dummies (connected by dashed lines). To enable comparisons between the skills measures, the three skills indices are standardized (mean 0, variance 1).

2. 95% confidence intervals for the linear trends are shaded.

Appendix A: The US Occupational Information Network (O*NET) database

The Occupational Information Network (O*NET) system (https://www.onetcenter.org/) is administered by the Bureau of Labor Statistics (BLS) and is the largest source of occupational competency data in the US. It was almost 20 years in development as a replacement to the DOT (Dictionary of Occupational Titles) system which was first published in 1939 (US Department of Labor, 1991). O*NET contains information on a wide range of occupational descriptors with a total of 239 measures of skills, abilities, training, educational and experience requirements, work contexts and other such domains.

Figure A1 shows the O*NET content model which describes the data structure. O*NET comprises worker-orientated and job-orientated characteristics at both an occupation-specific level and across occupations. Information is collated into six broad areas which include qualifications required, indicators of practical and technical skills, a wide range of 'soft skills' such as communication skills, stamina etc, and details of the tasks involved in the job. Most descriptors are comparable between occupations, although tasks are occupation-specific. O*NET information is gathered from job incumbents through postal and online questionnaires administered by the BLS, and from professional assessments by job evaluation analysts. Survey respondents are only asked to complete a random selection of the questionnaires in order to avoid survey fatigue. In addition, all respondents provide some background demographic information (which is not released) and are also asked to indicate from a wide range of occupation-specific tasks those that apply to their particular job. O*NET publishes occupation averages, rather than the individual micro-data. However, these averages are based on large samples - an average of 31,000 responses for each descriptor gathered from around 125,000 returned questionnaires. Information is published at the 'O*NET-SOC' occupation level, which is a slightly more detailed version of the US SOC. O*NET data collection began in 2001 and is being continually updated, with approximately 100 occupations updated each year (larger occupations are more regularly updated that the smaller occupations). Indeed, one of the main strengths of the O*NET design is that it is being constantly updated so that changes in skills utilisation *within* occupations can be discerned. There are currently 1,110 O*NET-SOC2010 occupations, of which 974 are 'data-level' (i.e. occupations for which data are separately collected). A comprehensive description and review of the O*NET system can be found in Peterson et al (1999) and Tippins and Hilton (2010).

The skills items contained in O*NET are comparable between occupations. Each of the 35 descriptors of skills (as well as the descriptors in the knowledge, abilities, and work activities areas) is given an 'importance score' and a 'levels score', and we retain information on both scores in constructing our UK skills database. Respondents are first asked how important the skill is to the performance of their job, and then what level of the skill is required to perform the job. Respondents rank importance on an ordinal five point Likert scale where a value of 1 means 'not important' and a value of 5 means the skill is 'extremely important'. Skill level is measured similarly, but on a seven point Likert scale. In order to help respondents to give an accurate and comparable measure of the required level of skill, the levels information is accompanied by 'scale level anchors' - short descriptions or examples at a number of points on the scale specific to the skill to indicate what a given value means in terms of the respective skill. An example is presented in Figure A2 for the 'reading comprehension' skill. Here, a level of 2 is described as 'read step-by-step instructions for completing a form' and level 6 is described as 'read a scientific journal article describing surgical procedures'. Respondents do not give a level rating if they rated the importance of the skill at 1 ('not important'). In our processing of the data, we give a level score of 0 to occupations with an importance score of 1 (i.e. not important).

While our focus is primarily on skills, our methodology for matching O*NET to UK occupations can equally be applied to other elements of the O*NET content model. Provided the particular set of descriptors are ordinal numerical variables which can be applied uniformly and are comparable across occupations, our methodology can be used to map the O*NET descriptors to the UK SOC.

Appendix B: Overview of the matching and mapping methodology

There are five inter-related stages involved in constructing the 4-digit occupational skills profiles that are used in the analysis presented in this paper:

(1) Constructing O*NET employment;

(2) Converting O*NET data to a common classification;

(3) Establish a matching matrix to convert O*NET SOC to the UK SOC;

(4) Converting UK SOC2000 data to UK SOC2010;

(5) Combining the data.

These stages are each described in turn in the following subsections.

B1: Constructing O*NET employment

When mapping O*NET skills data to the UK SOC2010, US occupational employment data are needed to weight the different O*NET occupations which map onto any given UK SOC code. As O*NET does not contain information on employment, we therefore utilise the BLS Occupational Employment Statistics (OES) to obtain employment by occupation for the US[13].

Two issues immediately arise. Firstly, the OES and O*NET use (slightly) different occupational classifications. OES is based on the US SOC, whereas O*NET makes use of its own unique classification which is a slightly extended version of US SOC. Secondly, neither the O*NET SOC or US SOC are based on a single consistent classification for the whole period 2002-2016. The US SOC changes in 2010 from SOC2000 to SOC2010, and the O*NET SOC changes three times during the same time period. This is also an issue in the UK data, where both ASHE and LFS data are based on UK SOC2000 until 2011 when a transition was made to UK SOC2010. Table B1 summarises these changes in the three occupational classifications used in our database construction.

[13] OES data (BLS, 2017) are collected in May of each year and the annual tables are downloadable from https://www.bls.gov/oes/tables.htm

The first stage is to base US occupational employment on a common classification. As well as giving us consistent occupations for the full period over which we are constructing our skills database, using SOC2010 makes the mapping of employment onto the O*NET SOC simpler as SOC2010 and O*NET SOC2010 correspond more closely than previous versions of the two classifications. We therefore base the full 2002-2016 data on SOC2010 to simplify the process of deriving O*NET employment.

In order to convert 2002-2009 employment to SOC2010, we use the online crosswalks provided by the BLS[14]. Ideally we would have weights for each SOC2000 occupation representing the proportion of employment on that SOC code which maps onto a given recipient SOC2010 occupation (one SOC2000 occupation can map onto more than one new SOC2010 occupation). Such a set of weights would allow us to immediately convert the distribution of employment across SOC2000 occupations onto the new SOC2010 classification.

However, we do not have this information. The crosswalk tables simply state, for each SOC2010 occupation, the corresponding SOC2000 occupation(s) which map to it. We therefore assume that, where a SOC2000 occupation maps onto more than one SOC2010 occupation, employment is equally distributed amongst each new SOC2010 occupation it maps onto. This approach produces a full 2002-2016 panel of US SOC2010 employment[15].

O*NET employment is compiled by converting employment in each year from US SOC2010 to the O*NET SOC2010. US SOC2010 consists of 840 occupations compared to the 1,110 occupations covered by the O*NET SOC. Mapping from the 840 SOC codes to the 1110 O*NET SOC codes is relatively straightforward as the 6-digit codes map onto corresponding 6-digit codes in O*NET SOC which are then further disaggregated into more detailed 8 digit occupations. For example, 11-3071.00 (Transportation, Storage, and Distribution Managers) in the US SOC2010 maps onto three different 8-digit occupations in O*NET which have the same 6-digit code as in the US SOC. These O*NET occupations are 11-3071.01 (Transportation Managers), 11-3071.02 (Storage and Distribution Managers), and 11-3071.03 (Logistics Managers). The other US SOC2010 and O*NET SOC2010 codes correspond in a similar manner,

[14] The crosswalk for SOC2000 to SOC2010 is at https://www.bls.gov/soc/soccrosswalks.htm

[15] In 2002/2003 there are a few occupations for which there are no employment data. In subsequent years there are employment figures for all occupations (with the exception of the military which is not covered).

so the mapping between the two classifications is a one-to-many merge of the former onto the latter.

As with the conversion of SOC2000 to SOC2010 we do not know the exact composition of a given SOC2010 code in terms of employment in the more detailed O*NET SOC occupations, so we again assume that SOC2010 employment is spread evenly across the recipient O*NET SOC2010 occupations. Applying this procedure for each year 2002-2016 produces a database of O*NET employment consistent with O*NET SOC2010.

B2: Converting O*NET data to a common classification

The second stage involves combining the O*NET employment data with the O*NET information on skills. As Table B1 illustrates, however, the O*NET skills data for 2002-2010 is collected and defined for earlier versions of the O*NET occupational classifications. This needs to be reconciled so that we have a database of both employment and skills which is consistent with O*NET SOC2010 over the full 2002-2016 period. This entails converting O*NET SOC2000, O*NET SOC2006, and O*NET SOC2009 skills to O*NET SOC2010. This requires employment weights and so a two-step approach is required. In the first stage, O*NET SOC2010 employment in each year from 2002-2010 needs to be converted to employment in the corresponding O*NET occupational classification for that year. These employment totals can then be used, in the second step, as weights to map the skills data onto O*NET SOC2010.

The mapping of employment from O*NET SOC2010 to its earlier counterparts is performed in the same manner used to convert employment from SOC2000 to SOC2010. As with OES, no weights are provided which would enable a fully specified mapping from one classification to another. We use the crosswalk tables provided by O*NET and again equally allocate employment from origin occupations to recipient occupations[16]. Once the skills data is matched to employment totals, it is converted back to O*NET SOC2010 for the years 2002-2010.

[16] The O*NET SOC crosswalk tables for converting between classifications are available at https://www.onetcenter.org/taxonomy.html

Having applied this procedure, the result is a panel dataset of skills and employment for the US which is consistent with O*NET SOC2010 for the full 2002-2016 period. As shown in Table B1, in some years the O*NET database is updated more than once. As we are constructing an annual dataset, we only need one database for each year. In the O*NET version column of Table B1, the version in bold indicates the version which we use to construct our database. The decision over which version to use in a given year is essentially arbitrary, but results are insensitive to the particular version chosen from a given year.

B3: Establish a matching matrix to convert O*NET SOC to the UK SOC

With the skills data all converted to a common classification, it can then be merged to the UK SOC. This requires the development of a correspondence table, matching O*NET SOC2010 occupations to UK SOC2010 occupations. In order to match between O*NET and the UK SOC, we use a Computer Assisted Structured Coding Tool (CASCOT), https://warwick.ac.uk/fac/soc/ier/software/cascot/) developed by the Warwick Institute for Employment Research (IER). CASCOT enables us to produce a systematic and automated mapping between job titles in the US (of which there are 59,634 as of O*NET version 20) and job titles in the UK (27,739 in SOC2010), with a matrix of scores between 0 and 100 reflecting the closeness of the match.

Each job title belongs to a distinct occupation, and thus the job-job matching also produces an occupation-occupation match. We can therefore measure both the coverage of the mapping, and also the quality of the match between the US and UK occupational classifications. The strength of the match is given a score by CASCOT. As an example, Figure B1 displays the CASCOT output in response to the job title ‘Economist’. It returns a score of 99 for SOC2010 unit group 2425 (actuaries, economists, and statisticians) and this is its top recommendation as it is the match with the highest score. There are other possibilities, and these are listed beneath the entry for SOC2010 2425. These additional choices mean it is possible to select accordingly if the top recommendation does not look appropriate. In an earlier feasibility study, Dickerson et al (2012) used CASCOT in its automated mode to produce a many-to-many occupational level match between 4-digit unit groups from UK SOC2010 and O*NET SOC2009.

In the automated mode, the match is to the index entry which generates the best score. Dickerson et al (2012) also made a number of recommendations for refining the process in order to improve the quality of the matching. For example: the SOC2010 dictionary and rules could be amended to recognise US spellings (e.g. O*NET job “Tire molder” which scores 0, vs “Tyre Moulder” which would score 96); and/or low scoring matches could be ignored (or downgraded).

Subsequent analysis, reported in LMI for All (2013, 2015), investigated the matching process further and made a number of modifications. While the initial matching undertaken by Dickerson et al (2012) was an automated procedure and thus ‘objective’ and replicable given the vintage of the CASCOT software and data, it was apparent that ‘expert’ intervention could improve the quality of the match. While this introduces an element of subjectivity, the expert CASCOT coder removes ambiguities and obvious errors (although the scores are then redundant because of reallocations). The expert coder produces a one-to-many matching matrix between 4-digit UK SOC2010 and O*NET SOC2010 and it is this correspondence matrix that we exploit further below (LMI for All, 2016).

The number of O*NET occupations which are assigned to each UK SOC2010 unit group is shown in Figure B2. 85 of the 369 (23%) unit groups are one-to-one matches and a further 74 have two O*NET occupations for each unit group; 70% have five or fewer matches, although there is one UK unit group which is matched with 35 different O*NET occupations. The same O*NET occupation can be matched to more than one UK SOC2010 unit group, and in total, 1,644 O*NET codes are matched, so each of the 1,110 O*NET occupation is used on average 1.5 times.

Using the CASCOT plus expert derived matching matrix between UK SOC and O*NET, we map the O*NET SOC2010 codes onto their corresponding UK SOC2010 codes. We then use O*NET SOC2010 employment to create UK occupation skills as weighted averages of O*NET occupation skills. When non-data level O*NET occupation observations are dropped, this leaves 362 of the 369 4-digit occupations in the UK SOC2010 to which skills data can be mapped. Three of the seven occupations for which we cannot obtain skills are military occupations, for which there are no O*NET nor OES data.

B4: Convert UK SOC2000 data to UK SOC2010

Having converted all of the O*NET skills data to O*NET SOC2010 and established a matching matrix to convert O*NET SOC2010 to UK SOC2010, we construct an occupation level panel dataset containing other relevant variables such as qualifications, wages, employment, and personal characteristics.

As with converting US SOC2000 employment to US SOC2010, we need to create a consistent classification so our SOC2010 skill database for the UK can be mapped onto data for 2002-2016. Our sources of data are the quarterly Labour Force Survey (LFS) microdata and Annual Survey of Hours and Earnings (ASHE) occupation level public release tables. Both of these data sources classify occupation by SOC2000 from 2002 to 2010, and so data for these years must be converted to SOC2010.

Unlike the US SOC, we do not need to make assumptions about the mapping of employment from one classification to another. The UK SOC2000 to SOC2010 correspondence tables additionally include weights produced by the ONS, reflecting the percentage of employed individuals in a given SOC2000 occupation which correspond to a particular SOC2010 code[17].

The ONS weights are derived from dual-coded individual-level datasets where detailed occupation is recorded according to both SOC2000 and SOC2010. These dual-coded datasets are then used to estimate the employment composition of SOC2010 codes in terms of SOC2000 occupations. These three dual-coded datasets are: (i) the LFS January-March 2007 quarter (LFSJM07); (ii) the 2001 Census (Census01), and (iii) the LFS December 1996-February 1997 quarter (LFS96-97)[18]. The weights differ according to the dataset used and, in some cases, there is no data available. Each dual-coded dataset is used in turn to produce SOC2010-consistent occupational level data for the 2002-2009 period, plus a fourth correspondence table which is calculated as the average across the three dual-coded datasets.

---

[17] These are provided by the Office for National Statistics (ONS) at https://www.ons.gov.uk/%20methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010

[18] The LFS began as a seasonal quarterly survey but switched to calendar quarterly for consistency with the European Labour Force Surveys.

In the example in Figure B3 below, SOC2010 group 2412 Barristers and Judges is associated with two SOC2000 groups, namely 2411: solicitors and lawyers, judges and coroners and 2419: legal professionals not elsewhere classified. For each of the dual-coded datasets the percentage by gender of those employed in the respective SOC2000 occupation that are also classified in SOC2010 group 2412 is reported. For example, using LFSJM07, 15.7% of males and 5.6% of females employed in SOC2000 group 2411 are also employed in SOC2010 group 2412.

Overall employment for each SOC2010 group is calculated simply by adding the separate male and female employment figures. In addition to estimating employment, these weights are used to convert occupational mean variables in the LFS/ASHE data from SOC2000 to SOC2010. Weighted overall values of wages and other variables are computed as the mean of the female and male figures for each occupation weighted by the gender shares of employment.

B5: Combining the data.

At this point we have a database containing information on wages, employment, education, and other labour market and personal characteristics aggregated to 4-digit SOC2010 occupation level. We also have a database of skills defined at the same level. The final step is to merge the skills data with these other personal characteristics and labour market information.

When non-data level O*NET occupation observations are dropped, this leaves 362 of the 369 4-digit occupations in the UK SOC2010 covered. The seven occupations not matched are: officers in the armed forces, NCOs and other ranks, officers of NGOs, finance officers, senior care workers, care escorts, and window cleaners. There are also a small number of occupations which are missing O*NET data in some, but not all, years. This is due to employment not being recorded for all occupations in the 2002 and 2003 OES, and the changing occupation classifications in the US data. In total, 233 observations are missing O*NET data out of a total of 5,535 observations (15 years of data with 369 occupations in each year).

Appendix C: Reconciling job incumbent and job analysis skills ratings

One issue with using the skills data available from O*NET is the use of different respondents for the assessment of the skills used in employment. Approximately 80% of the skills data are provided by job evaluation analysts, and from 2010 onwards this is exclusively the case. In the O*NET versions for which data were collected in the period 2002-2009, however, there is a mix where some occupations were valued by job analysts, and some by job incumbents.

As demonstrated by Mumford et al (1999) and Tsacoumis and Van Iddekinge (2006), there is a systematic difference between the skills importance ratings provided by job incumbents and those provided by job analysts. In general, job incumbents tend to provide higher skill ratings than job analysts. While the assessment of skills *rankings* of job incumbents and job analysts are very similar, this difference in means may impact on the derived measure of skills, and hence on our analysis.

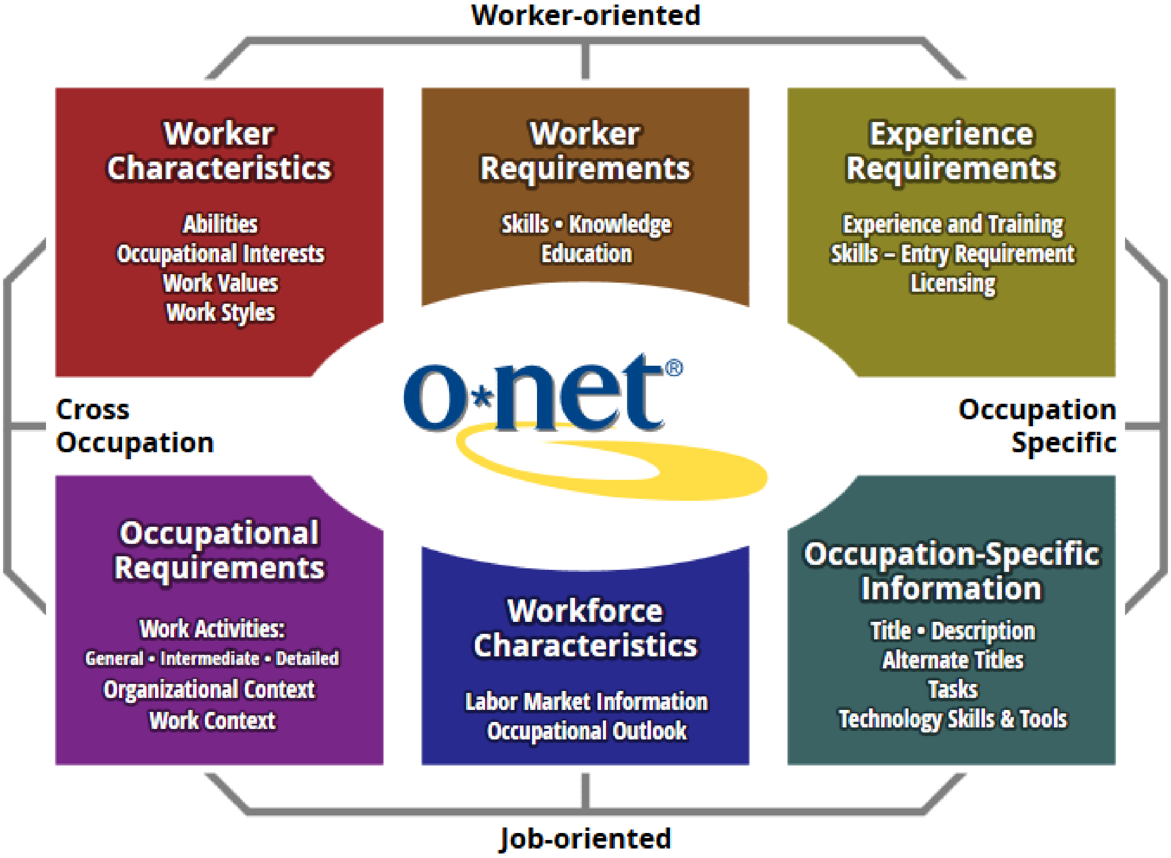
Figure C1 illustrates the switch between incumbents and analysts between O*NET versions/over time. Up to O*NET version 12, job incumbents' assessments were increasingly replacing the job analysts' ratings. However, from O*NET version 15 onwards, there was a switch, and 100% of the skills ratings were provided by job analysts thence forth. The impact of this change was that mean skills levels fell in O*NET version 15.0.

To address this issue, we use a regression-based approach to adjust incumbent-rated observations into values which are consistent with the analyst ratings. This rescaling approach involves estimating the extent to which job incumbents under- or over-assess skills at a particular time relative to job analysts. We then obtain job analyst-consistent observations by removing the incumbent effect. Our approach is to estimate the parameters in equation (C1), separately for importance and levels measures of skills, using pooled data from 2002-2009 O*NET versions:

$$O_{kt}^{(x)} = \alpha + \tau_t + \mu_k + \beta_1 I_{kt} + \beta_2 (I \times \tau)_{kt} + \varepsilon_{kt} \tag{C1}$$

The O*NET measure of skill x in occupation k at time t is a function of a vector of time dummies indicating the O*NET version, an occupation fixed-effect μ, a dummy, I, equal to one if skill x was assessed by an incumbent and zero otherwise, an interaction of the incumbency and time dummies, and an idiosyncratic error. $\beta_1 + \beta_2\tau_t$ is the estimated 'incumbent effect' in period t representing the systematic over- or under-assessment made by job incumbents of the skill in occupation k, relative to job analysts. Our rescaled importance and levels ratings for each skill is obtained by subtracting this incumbent effect from the observed O*NET skill. We use these incumbent-effect-modified skills measures in our analysis, although we investigate the robustness of our estimates to this treatment of the switch between incumbents' and analysts' measures of skills.

Figure A1: The O*NET content model



Source: https://www.onetcenter.org/content.html

Figure A2: Example from skills questionnaire

1. Reading Comprehension — **Understanding written sentences and paragraphs in work-related documents.**

A. How <u>important</u> is READING COMPREHENSION to the performance of *your current job*?

Not Important*	Somewhat Important	Important	Very Important	Extremely Important
1	2	3	4	5

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

B. What <u>level</u> of READING COMPREHENSION is needed to perform *your current job*?

1	2	3	4	5	6	7
	Read step-by-step instructions for completing a form		Read a memo from management describing new personnel policies		Read a scientific journal article describing surgical procedures	Highest Level

Source: https://www.onetcenter.org/questionnaires.html

Table B1: Changes in occupational classification for O*NET, US SOC (OES) and UK SOC

Year	O*NET versions	O*NET SOC	US SOC	UK SOC
2002	**4.0**	2000	2000	2000
2003	**5.0**, 5.1			
2004	**6.0**, 7.0			
2005	**8.0**, 9.0			
2006	10.0, **11.0**	2006		
2007	**12.0**			
2008	**13.0**			
2009	**14.0**	2009		
2010	**15.0**		2010	
2011	15.1. **16.0**	2010		2010
2012	**17.0**			
2013	**18.0**			
2014	18.1, **19.0**			
2015	**20.0**, 20.1			
2016	20.2, 20.3, **21.0**, 21.1			

Note: O*NET versions in bold indicate the version chosen for each year

Figure B1: CASCOT illustration – classification of 'economist'

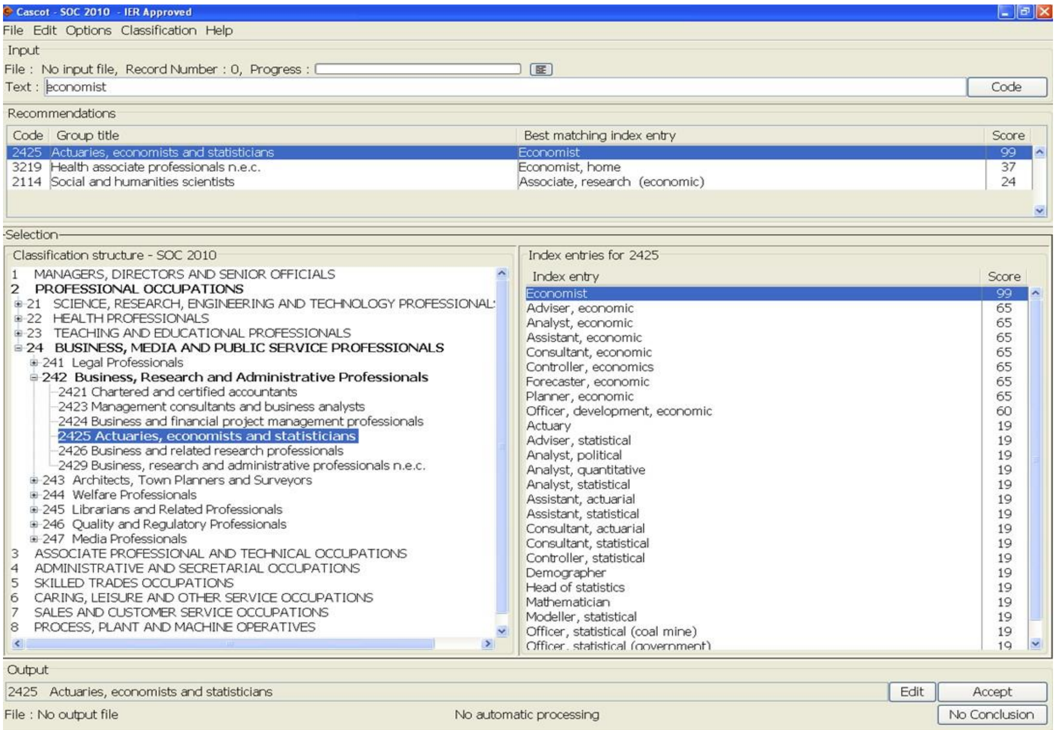


Figure B2: Distribution of the number of O*NET SOC2010 to UK SOC2010 matches

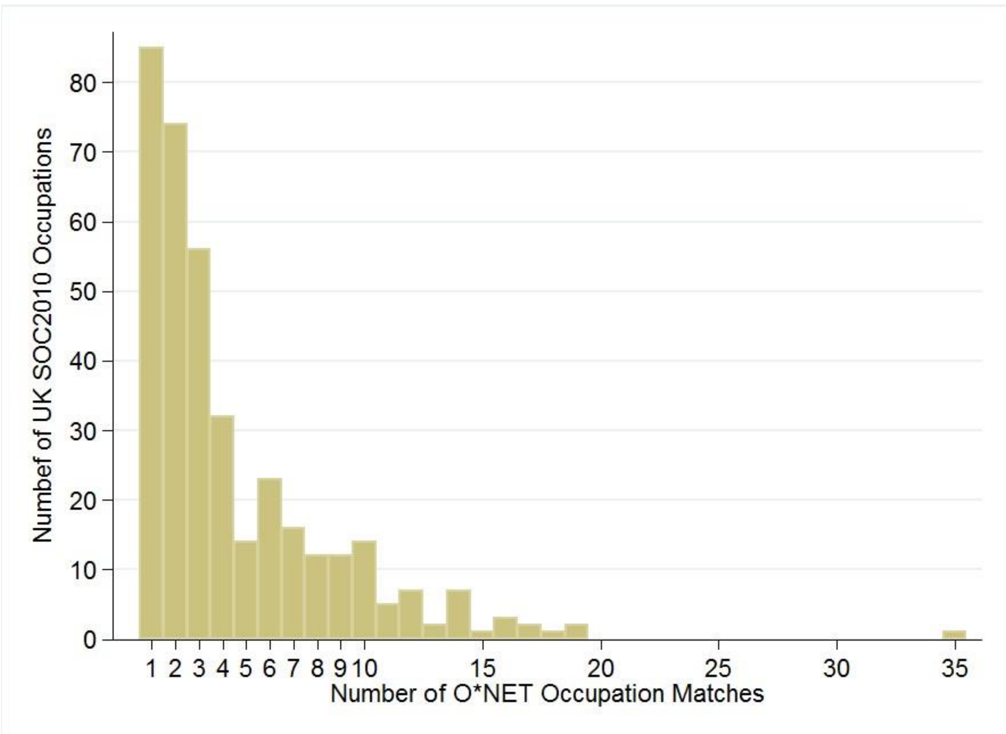


Figure B3: Extract from the ONS SOC2000 to SOC2010 correspondence table

MALE							FEMALE
LFS JM07	**Census01**	**LFS 96_97**	**SOC2000**	**Unit Group Title**	**LFS 96_97**	**Census01**	**LFS JM07**
2412 Barristers and judges †							
15.7	17.0	13.8	2411	Solicitors and lawyers, judges and coroners	5.6	11.7	14.4
5.0	–	–	2419	Legal professionals n.e.c.	–	–	–

Figure C1: Proportion of skills ratings by job incumbents and analysts