PRELIMINARY – DO NOT DISTRIBUTE OR CITE WITHOUT PERMISSION

THE DEMAND FOR SKILLS IN THE UK

Andy Dickerson and Damon Morris

Department of Economics and CVER

University of Sheffield

Draft: June 2017

Abstract

We present estimates of changes in skills utilisation and in the returns to skills in the UK using new measures of skills derived from a detailed matching between the US O*NET system and UK SOC. A decomposition analysis suggests strongly increasing utilisation of both analytical skills and interpersonal skills, and declining use of physical skills over the period 2002-2016. Most of these changes in skills utilisation are within occupations rather than between occupations, suggesting that the changes are pervasive throughout employment. The wage returns to skills are estimated using a standard Mincerian earnings function. We find positive and significant returns to analytical skills, while the returns to interpersonal skills are insignificantly different from zero over almost all years. Finally, the returns to physical skills are significantly negative over the whole period. Our results are robust to changes in the definitions and measurement of the skills variables, and to the empirical specification of the earnings function.

JEL:

Keywords:

Acknowledgements: CVER/BIS/DfE;

1. Introduction

As indicators of the 'demand' for skills, we report estimates of the changing skills utilisation in employment and in the wage returns to skills in the UK. We employ the matching methodology described in Dickerson (2016), in conjunction with data on employment and wages from ASHE and the LFS, and skills and abilities measures from the extensive US Occupational Information Network (O*NET) system, to document the changing utilisation and returns to skills in the UK for 2002-2016. We first utilise the correspondence tables (ONS, 2012) between SOC2000 and SOC2010 to 'backcast' 4-digit employment levels, employment composition and wages for 2002-2010 on a consistent SOC2010 basis and combine this with SOC2010 data from 2011-2016. We then match the O*NET skills and abilities indices to this 4-digit SOC2010 panel, and use the resulting dataset to describe the changes in skills utilisation and the returns to skills over the period 2002-2016.

We illustrate our analysis with the 'data-people-things' (DPT) taxonomy of skills as described in Dickerson (2016). A decomposition analysis and correlation/regression results are presented for changes in skills utilisation and the wage returns to skills and educational qualifications using a standard Mincerian earnings function specification. Our preliminary results suggest strongly increasing usage of both data/analytical skills and people/interpersonal skills, and declining use of manual/practical skills over the period 2002-2016, consistent with our prior expectations. In terms of the wage returns to these skills, we document increasing (conditional) returns to data/analytical skills, suggesting that the demand for such skills is perhaps increasing even more strongly than the growth in their utilisation. In contrast, the returns to people skills are insignificantly different from zero over almost all years. One possible interpretation of this finding is there has been a corresponding growth in the supply of those skills to the labour market to match with the growth in their utilisation. Finally, the estimated returns to manual/practical skills are significantly negative over the whole period, although this negative return is unchanging over time despite the strong secular decline in the utilisation of these kinds of skills.

The remainder of this paper is structured as follows. The next section reviews the relevant literature. Section 3 describes the construction of the 4-digit SOC2010 skills indices for 2002-2016. Section 4 describes the trends in these skills utilisation indices 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. Earnings function regressions of the returns to skills are then presented together with the time-series pattern in these returns. Section 5 concludes with the implications for education and skills policy in the UK.

2. Literature Review

2.1 Literature on defining and measuring skills

The importance of skills in modern economies is widely acknowledged. Skills are important at both micro level eg for the distribution of earnings, and at the macro level eg 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.

'Skills' are typically measured or proxied in a variety of different ways. Typically, while these measures are relatively simple to obtain from survey data, they are poor proxies for skills. For example, educational qualifications are usually acquired while still in (usually full-time) education i.e. before labour market entry, and qualifications (especially perhaps 'academic' qualifications), typically only have a very loose link with job skills. Similarly, the Standard Occupational Classification (SOC) is hierarchical, uni-dimensional, and static, and thus captures neither the breadth nor the changing nature of skills used in different jobs over time. Moreover, employers increasingly focus on 'generic', 'key' or 'core' skills such as 'communication' and 'flexibility' rather than certificates of qualifications. These 'soft' skills are rather more difficult to measure although some progress has been made in recent years in surveys that focus on the tasks that individuals perform in their jobs.

Table X describes the commonly used measures of skills in research and policy, together with their merits and demerits.

2.2 Literature on returns to skills

2.3 Literature on using O*NET measures of skills

Table X documents some of the ways in which this has been accomplished previously using the DOT, O*NET, plus other job-task surveys with similar structures and/or characteristics to the O*NET surveys. It is common to select a subset of 'relevant' O*NET items corresponding to some pre-defined taxonomy, 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 and formally implemented in the DOT occupational codes as 'Data-People-Things' (although the language now used is Analytic/Cognitive, Interpersonal and Physical or some variant thereof). 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.

3. Data and Methodology

A full description of the matching methodology used to construct our skills indices is provided in Dickerson (2016) and its operationalisation in summarised in Annex A. A brief description is provided in subsection 3.1 below.

3.1 Data

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

1. UK LFS data 2002-2016:
2. UK ASHE/NES data 2002-2016;
3. US O*NET 2002-2016;
4. US Occupational Employment Statistics (OES) 2002-2016;

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 area of (a) knowledge, (b) skills, (c) abilities and (d) work activities, both the 'Importance' and 'Level' of each skill or characteristic being measured is recorded. Most descriptors are comparable between occupations (although tasks are occupation-specific).

O*NET information is gathered from postal and online questionnaires administered by the US Bureau of Labor Statistics. Respondents are only asked to complete a random selection of the questionnaires in order to avoid survey fatigue, and also to provide some background demographics (not released). 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 250 descriptors gathered from around 125,000 returned questionnaires. Information is published at the 'O*NET-SOC' occupation level, which is a modification (i.e. slightly more detailed) 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').

UK LFS and ASHE/NES data provide 4-digit SOC2000 data for 2002-2010, and SOC2010 data for 2011-2016, on the structure and composition of earnings and employment. We use ASHE for wages because of the larger sample sizes available for the detailed 4 digit occupations that we are using (and also its lack of proxy responses).

We then match O*NET to the UKSOC using a CASCOT plus expert coder mapping between the 90,000 US job titles underpinning the O*NET and the c. 30,000 job titles which are listed in the UK SOC. This produces a one-to-many mapping between each 4-digit UKSOC and a number of O*NET 7 digit SOC codes (between 1 and 37). Having established the mapping between UKSOC and O*NET, we can utilise the O*NET measures of skills to construct occupational skills profiles for the UK. In order to aggregate the O*NET measures where the mapping is more than one-to-one, we use OES data to compute employment weighted averages of the O*NET skills measures for each UKSOC code. While OES is based on USSOC and O*NET has its own variant of the USSOC, there are 'crosswalks' between USSOC and O*NET SOC.

3.2 Methodology

Our skills measures are constructed as follows. We compute a vector of skills, $S_{jt}^{(x)}$, where the measure of each skill *x*, $S^{(x)}$, for each UK 4-digit occupation $j = 1, \dots J$ at time t is 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*, 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 the particular UKSOC 4-digit occupation *j*. Essentially, this 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*, using the CASCOT-plus-expert-derived match between O*NET and UK SOC as described in Dickerson (2016). We calculate this for each of the $x = 1, \dots, 35$ measures of skills in O*NET.

We then aggregate the resulting 35 skills measures into three indices closely informed by the "data-people-things" (DPT) taxonomy. We here use the terms analytical, interpersonal, and physical, to mean data, people, and things respectively. As detailed by Dickerson (2016), these are:

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 – simply averaging, PCA, etc. - and with additional choices regarding the inclusion of the Levels as well as Importance measures of each skill (see Dickerson, 2016, for details). In our main results, we simply take the average of the importance measures of skills only, although we examine the sensitivity of our findings to this choice in the extensive robustness analysis.

In order to produce a SOC2010-consistent 4-digit panel for 2002-2016, we have to resolve the change in the standard occupational classifications that has taken place in the UK as well as in the US over our sample period. 'Correspondence tables' for the UK, and 'crosswalks' for the US are available to convert between the different SOC classifications. We use these as described in Annex A to produce a SOC2010-consistent 4-digit panel for 2002-2016 with information on employment composition and structure, wages, together with measures of skills and abilities derived from O*NET.

One further issue is that the O*NET measures of skills in the early part of our sample period (2002-2009) were partially provided by job incumbents rather than job analysts. From O*NET v.15.0 (2010) onwards, the skills measures were exclusively provided by job analysts for all occupations. This changing mix of incumbents and analysis has been previously analysed by O*NET and their conclusion is that the different measures are equally valid. However, this issue does have implications for the measures of skills over time as shown in Annex B. Our solution to this issue is also described in Annex B: in essence, we use the changing mix of incumbents and analysts to impute the 'incumbent-effect' by occupation, which we then subtract from the skills measure to produce a job-analyst consistent measure of skills for the whole period. We also investigate the robustness of our findings to the adjustment method we have chosen.

4. Results

4.1 Skill trends and decomposition

The overall changes in the 3 skills indices are reported in the first column of Table 1. Over the whole period, our (employment-weighted[1]) aggregate index of analytical skills suggests that utilisation of this skill set grew by 10% over the period. The increase in people skills was more than double this (+23%), while utilisation of things skills fell by 14%. The direction of these changes are in line with our expectations. These trends accord with our general understanding of the changing occupational structure of employment and the large literatures on skill-biased technical change, routinisation of jobs, post-industrialisation/growth of services/decline of manufacturing etc.

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 this can be obtained from undertaking a decomposition of the overall change in skills utilisation between 2002 and 2016 in each of our skills measures

[1] i.e. the 4-digit indices are weighted by their employment shares in total employment for each year.

in order to assess the extent to which the aggregate changes are 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, $j = 1, \dots, J$, 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 1 reports the decomposition of the overall change in analytical skills, people skills and things 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-occupations changes for all 3 indices whatever level of occupational classification is used. 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 people skills and things skills are even greater – at almost 90%. This suggests the overall changes in skill utilisation are pervasive and affecting all occupations, rather than being concentrated in certain occupational groups.

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 rather than the individual. Table 2 presents the basic log hourly wage regression results. Column (1) shows that wages are positively correlated with analytical skills, and negatively correlated with people skills and things skills. These correlations are highly significant statistically. Column (2) reports the basic earnings function estimates. 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; and that larger firms tend to pay significantly more. These are all standard findings in the earnings function literature. Column (3) augments our earnings equation with the 3 indices of skills. This suggests that skills and education are correlated and at least some of the returns to education are, in fact, returns to skills (and vice versa). Year dummies are included in column (4) since there are macro and other temporal changes which have impacted on earnings in this period, but these do not change the qualitative findings.

Finally, we examine how these returns to skills are changing over time. Selecting the specification in column (4), we hold the returns to education and to gender and the other control variables fixed, and allow the returns to the 3 skills indices to change over time by interacting them with time dummies. The estimated regression coefficients on the interaction terms are shown in Figure 1. The returns to data skills are strongly increasing over time, whereas, in contrast, the returns to people skills and things skills are relatively static, with the returns to people skills close to zero throughout the period.

4.3 Robustness Checks

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

4.3.1 Data Transformations and Sources

In Table 3 we present the robustness of our findings to the particular method we use to aggregate the 35 skills into our DPT measures.

Panel A reports results which are based on aggregations using the importance of skills only. In the first three columns of Panel A the results are based on average of the importance measures of the relevant skills in column (1), standardised measures in column (2) to zero mean and unit variance within years, and the percentile rank within year of the skills measures in column (3) (which will again be robust to rescaling of the indices). Columns (4), (5) and (6) repeats this except that Principal Components Analysis (PCA) is used to aggregate the skills measures into each category rather than the mean, and the first principal component is used as the skills measures. Standardised and rank based transformations are again also investigated.

In Panel B we incorporate the levels information as well as importance measure of skills. We again compute a mean based measure and a PCA based measure. Rather than a simple mean, however, we follow the approach of Blinder (2007) 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 3 provide evidence that the way in which we aggregate the skills information from 35 measures of skill in the raw data to our 3 summary indices does not have an impact on our findings. In each of the 12 estimates, the coefficients for analytical skills and things skills are consistently statistically significant at the 1% level. Notably, incorporating levels information produces some significant estimates for people skills. The magnitudes of the raw skill measures are not directly comparable as different aggregation methods produce variables on different scales with coefficients differing in their interpretation. The coefficients on the standardized and percentile rank transformations do, however, indicate that the size of the effects of skills is similar across the four aggregation methods.

Returning to our chosen method of aggregation for our main variables of interest (simple means of the importance measures of skills only), we also report a range of other robustness checks in Table 4 to assess the sensitivity of our results to transformations we have made to the data or assumptions we have made in constructing it. In particular, we check the robustness of the results to using the LFS rather than the ASHE as our 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 of incumbent-provided skill information), and finally the choice of weights we use to convert information at the SOC2000 level to SOC2010.

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

Our preferred measure of wages is derived from ASHE for the reasons stated above. As an alternative, we can use log mean wages from LFS. We could also use mean log wages for LFS data where we have individual earnings, since the aggregation of individual log earnings functions yields this as the 'correct' dependent variable (although cell-mean regressions of this kind (eg Blanchflower et al (1996) and Dearden et al (2006)) use log mean wages rather than mean log wages). 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 between LFS and ASHE wages has no 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 things skills slightly, although the main conclusions are unaffected.

The baseline results use the average across the three SOC2000-SOC2010 weighting matrices provided by the ONS to construct the correspondence between SOC2000 and SOC2010. In columns (5) to (7) we use each of the three weighting matrices separately in turn to convert the 2002-2011 data to SOC2010 consistency. The three matrices each produce a very similar magnitude of results for the three skill measures, and there are no statistically significant differences from the baseline.

The changes to data source and construction we have investigated in Table 4 do not substantially alter any of our main findings or conclusions. Our estimates of the returns to skill remain the same in terms of sign and significance at the 5% level and the magnitudes are robust to thee choices of aggregation methods.

4.3.2 Specification and Estimation of the Earnings Function

Table 5 presents another set of robustness checks for our main result, in this case focussing on the robustness of results to our chosen specification and estimation technique.

Our main results are based on variables derived from gender-specific variables combined using an employment share weight. In columns (2) and (3) of Table 5, we compare the results when our variables are based on, respectively, males only and females only. We find that both male and female occupational average wages are influenced by analytical, people, and things skills in the same way – positive effect of analytical skills, negative effect of things skills, and no significant effect of people skills. The magnitudes, in both cases, are larger than the 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 individual is employed full time. 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 things skills. Even when comparing 4-digit occupations within the same detailed 3-digit grouping of occupations, differing skill levels across the occupations still account for some of the differences in wages.

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 (Annex B), we re-estimated the returns to skills for the period 2011-2016 since this period is unaffected by any of the changes in SOC or in the reporting of skills. 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 period 2002-2016, but the substantive results are the same. This suggests that the additional manipulations and adjustments required in order to construct the SOC2010 consistent database for 2002-2010 are not unduly responsible for the results obtained, though one caveat to this is that we find much stronger positive and now significant returns to people skills when focussing on the later period only.

The final specification issue is the use of OLS, when it can be argued we should use weighting since we are using group mean regressions. There is some debate in the literature about the necessity of weighting, but it is important to investigate if it makes a difference to the estimated returns. We follow the approach of Dickens (1990) in weighting our regressions. First we estimate our standard specification by WLS, weighting by the square root of cell size (in this case, employment in the occupation). We then estimate the following regression, where the group-specific residual is regressed on a constant and the inverse of employment in the occupation.

$$\hat{\varepsilon}_j^2 = \alpha + \delta \, {}^{1}/_{N_j} \tag{3}$$

The parameter estimates are then used as estimates of the error variance components in equation (4):

$$Var(\bar{e}_j) = \sigma_\gamma^2 + \sigma_u^2 / N_j \tag{4}$$

These estimates, $\hat{\sigma}_\gamma^2$ and $\hat{\sigma}_u^2$, are used to construct the weight $\frac{1}{\sigma_\gamma^2 + \sigma_u^2/N_j}$. The earnings function is then re-estimated using this weight, 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 residuals regressions are identical to 5 decimal places between two iterations. This convergence occurs at the 4th iteration, and it is these results presented in column (7) of Table 5. The full results of this weighting procedure are presented in Table 6, and show that the coefficient estimates (to 3 decimal places) in the earnings function converged immediately at the first iteration. The coefficient estimates in column (7) of Table 5 do not differ significantly from our main results reported in column (1). Our results are therefore not sensitive to whether or not we weight the earnings regressions.

Taken as a whole, our comprehensive set of robustness checks show that our main estimates presented in section 4.2 are highly robust and stable. Despite uncertainties around how (or if) to deal with the issue of job incumbent/job analyst valuations of O*NET skills, the best source of data for wages, how to aggregate raw skills, and how to convert between SOC2000 to SOC2010, we find that our results are not sensitive to these choices in the construction of the dataset.

5. Summary and Conclusions

We derived 3 skill indices at the 4-digit level for 2002-2016 and examined the evolution of the utilisation of these skills over time as the occupational composition and skills content of jobs changed. Strong secular growth in analytical and people/interpersonal skills and declining usage of manual/practical skills is consistent with other literatures which have documented the skill content of jobs. We then estimated earnings functions which controlled for education, 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 was contrasted with zero returns to people/interpersonal skills, while the returns to manual/practical skills is significantly below zero for the whole of the period. Our findings demonstrate the importance of work-related skills for individual earnings over and above their educational qualifications (which are typically obtained before they join the labour market), and in particular, the demand for higher levels of analytical skills (mirrored by the findings on cognitive abilities) in the workforce.

References

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

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, D. Oswald, A. and Sanfey, P (1996), "Wages, profits, and rent-sharing" *The Quarterly Journal of Economics,* 111, pp.227-251

Blinder, A (2007), "How many U.S. jobs might be offshorable?", Princeton University Center for Economic Policy |Studies Working Paper no. 142, Princeton University

BLS (Bureau of Labor Statistics) (2015) Occupational Employment Statistics (OES) Survey http:/stat.bls.gov/oes/home.htm last accessed 11 November 2015.

CASCOT (no date) (Computer Assisted Structured Coding Tool) http://www2.warwick.ac.uk/fac/soc/ier/software/cascot/ last accessed 12 January 2016.

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

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

Dickerson, Andy (2016), "Measuring Skills: Update on CVER Project 3.1: The changing patterns of skills demand/utilisation in the UK", CVER mimeo, January 2016.

Dickerson, Andy, 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.

Dustmann, Christian, Johannes Ludsteck and Uta Schoenberg (2009), "Revisiting the German wage structure", *Quarterly Journal of Economics* 114(May), 843–882.

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.

Goos, Maarten, Alan Manning and Anna Salomons (2009), "Job polarization in Europe", *American Economic Review Papers and Proceedings*, 99 (2), pp.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.

Handel, Michael J (2007), A new survey of Workplace Skills, Technology, and Management Practices (STAMP): Background and descriptive statistics. Paper presented at Workshop on Research Evidence Related to Future Skill Demands, National Academies, Washington, DC.

Handel, Michael J (2008), What do people do at work? A profile of U.S. jobs from the Survey of Workplace Skills, Technology, and Management Practices (STAMP). Paper presented at the Labor Seminar, Wharton School of Management, University of Pennsylvania.

Handel, Michael J (2010), "The O*NET content model: strengths and limitations", mimeo, June 2010.

Howell, D R and E 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, D R and E 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, pp.35-59.

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 12 January 2016.

ONS (2012), *Relationship between: Standard Occupational Classification 2010 (SOC2010) and Standard Occupational Classification 2000 (SOC2000)*, Classification and Harmonisation Unit, User Guide 2010: 22
available at: https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010, last accessed: XX XXX 20XX.

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.

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

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.

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

Annex A: Matching O*NET data to UK SOC

When matching O*NET skills data to the UK SOC2010, we need employment figures to weight the different O*NET occupations which feed into a given SOC code. As O*NET does not collect information on employment, we also need to utilise the OES[2] (Occupational Employment Statistics). Two issues arise from this. Firstly, the OES and O*NET use different occupational classifications. Secondly, neither the O*NET SOC or US-SOC are based on a consistent classification for the period 2002-2016. US-SOC used in the OES changes in 2010, and the O*NET SOC changes three times during the 2002-2016 period for which we construct our database of skills and labour market information. Table A1 summarises these changes in occupational classification in our data sources.

Table A1: Years for which SOC 2010 needs converting to older classifications

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

Stage One: Convert OES data to a common classification

Our source of US employment data is the OES. As US-SOC and O*NET-SOC 2010 are both more closely aligned then previous versions of both classifications, we would like the full 2002-2016 employment data to be based on SOC-2010. In order to convert 2002-2009 employment we use the online crosswalks[3]. We assume that the employment from a given SOC2000 code is equally distributed among each new SOC2010 it maps to. This simple approach, necessary because of the lack of any weights for a more accurate mapping, gives us a full 2002-2016 panel of US SOC2010 employment.

[2] OES data are collected in May of each year and are downloaded from: http://www.bls.gov/oes/tables.htm
[3] US SOC2000 to SOC2010 crosswalk table is downloaded from: http://www.bls.gov/soc/soccrosswalks.htm

Stage Two: Convert O*NET data to a common classification

For the years in which O*NET SOC does not use the 2010 classification, the employment in terms of O*NET SOC2010 needs to be converted to the relevant classification in order to correspond with the classification used to collect the skills data to be merged in. In some years the O*NET database is updated more than once. As we are constructing an annual panel, we only need one database for each year. In the O*NET version column of Table A1, the version in bold indicates the version which we use to construct our database – the decision is arbitrary and results are insensitive to the particular version chosen from a given year.

The mapping of employment between different O*NET SOCs is performed in the same manner as converting OES employment from SOC2000 to SOC2010. As with OES, no weights are provided which would enable a completely accurate mapping from one classification to another. We use the crosswalk tables provided by O*NET[4] and equally spread employment from origin occupations to recipient occupations. We refer to the weights constructed for this purpose as "down weights" – for converting more recent SOCs to older ones.

In the case of one-to-one mapping the down weight is equal to one. If one code is matched to *n* new codes then the down weight for each recipient code is *1/n*. SOC2010 has 840 detailed occupations, O*NET SOC2010 has 1110 detailed occupations including the 840 SOCs. Of these 1110, there are 974 data level occupations (for which data is collected). In order to produce a consistent series, once the skills data is matched to the employment data, it needs to be converted back to O*NET SOC2010 for the years 2002-2010. This is achieved by using employment to construct "up weights". To construct the up weight, we calculate the share of employment in the older classification code out of the total employment which will be mapped to a particular newer classification code. Weighting in this way restores the initial O*NET SOC2010 employment figures.

Table A2 shows a simplified example to illustrate the procedure, assuming three occupations in both the O*NET SOC2010 and O*NET SOC2009 classifications. The "down weight" (for converting the newer to the older classification) is *1/n* where *n* is the number of recipient occupations in the older classification (SOC2009) onto which each given code in the newer classification (SOC2010) matches. The employment totals in column 2 are multiplied by these weights in column (4) to give the corresponding employment totals for the older classification's occupational codes. The versions of O*NET consistent with the SOC2009 classification would at this stage be merged with the employment data, requiring conversion back to SOC2010.

The "up weight" for converting the older to the newer classification by taking the contribution of each O*NET SOC2010 code to the respective O*NET2009 code employment and dividing by total SOC code employment. Multiplying the total SOC2009 employment given in column (6) by the up weight in column (7) yields employment totals which match the original O*NET SOC2010 employment (i.e. column (9) matches column (2)) and therefore appropriately weights the merged O*NET data to SOC2010 consistency. Having applied this

[4] O*NET SOC crosswalk tables are downloaded from: https://www.onetcenter.org/taxonomy.html

procedure, at this stage we have a database of skills and employment for the U.S. which is consistent with O*NET SOC2010 for the full 2002-2016 period.

Table A2: Simplified example of the "down-weight/up-weight" procedure

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
O*NET 2010 Code	O*NET 2010 Employment	O*NET 2009 Code	"Down Weight"	O*NET 2010 contribution to O*NET 2009 Employment	Total by SOC 2009 Code	"Up Weight"	(6)*(7)	
1	20	1	1	20	35	20/35	20	20
2	30	1	0.5	15		15/35	15	30
2		2	0.5	15	15	1	15	
3	50	3	1	50	50	1	50	50

Stage Four: Map O*NET SOC2010 onto UK SOC2010

Using the CASCOT plus expert derived matching matrix between UK SOC and O*NET, merge the UK SOC2010 codes onto the O*NET SOC2010 codes. Collapse the data by UK SOC2010 code and weight the O*NET data by US employment. When non-data level O*NET occupation observations are dropped, this leaves 362 of the 369 four 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 or OES data.

Stage Five: Convert UK data to a common classification

In stage 5 we put together a database containing all other variables required for our analysis from UK ASHE and LFS data. As in stage 1, we need to create a consistent classification so our SOC2010 level skill database for the UK can be mapped onto other UK data for 2002-2016. We do not have to make assumptions about the mapping of employment from one classification to another as we did with the OES and O*NET classifications. The UK SOC2000 data is converted to SOC2010 using correspondence tables[5] produced by the ONS.

These tables make use of dual-coded datasets – individual level datasets where occupation is recorded according to both SOC2000 and SOC2010. These dual-coded datasets are then used to estimate the composition of SOC2010 codes in terms of SOC2000 occupations.

These three dual-coded datasets are:

I. LFS Jan-March 2007
II. 2001 Census
III. LFS December 1996 - February 1997

In the example in Table A3 below, SOC2010 group 2412 Barristers and Judges is associated with two SOC2000 groups – 2411 solicitors and lawyers, judges and coroners and 2419 legal

[5] Available here- https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010

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.

Table A3: Extract from the ONS SOC2000-SOC2010 Correspondence Table

MALE							FEMALE
LFSJM07	Census01	LFS96_97	SOC2000	Unit Group Title	LFS96_97	Census01	LFSJM07
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.	_	_	_

These percentages differ according to the dataset used and in some cases, depending on the dataset used, there is no figure available. As in the case for females in SOC2000 group 2419, there are instances where no figure is available regardless of dataset. 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. This fourth table is the one we use in our main analysis.

SOC2010-consistent data is produced using the respective dataset to create weights. Each SOC2010 occupation employment is derived by taking the sum of each constituent SOC2000 code employment, weighted by the proportion of individuals in that SOC2000 code who are also in the respective SOC2010 occupation.

These employment weights are then used to compute weighted values of all other variables in the occupational level LFS/ASHE data. Overall employment for each SOC2010 group is then calculated simply by adding the separate male and female employment figures, and the relative male/female employment is used to compute weighted overall values of wages and the other variables.

Stage Six: Merge the data

At this point we have a database on wages, employment, education, and other labour market and personal characteristics aggregated to 4-digit SOC2010 occupation level, which we constructed in stage 5. We also have a database of skills defined at the same level. In stage 6 we merge these two sources of data to construct our final dataset.

Annex B: Reconciling job incumbents and job analysts measures of skills

One issue with using the raw skills data available from O*NET is the use of different methods of obtaining valuations of skills. Almost 80% of the skills data are provided by job 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 analysts, and some by job incumbents.

Figure B1: Mean of Skill Importance Measure over Time

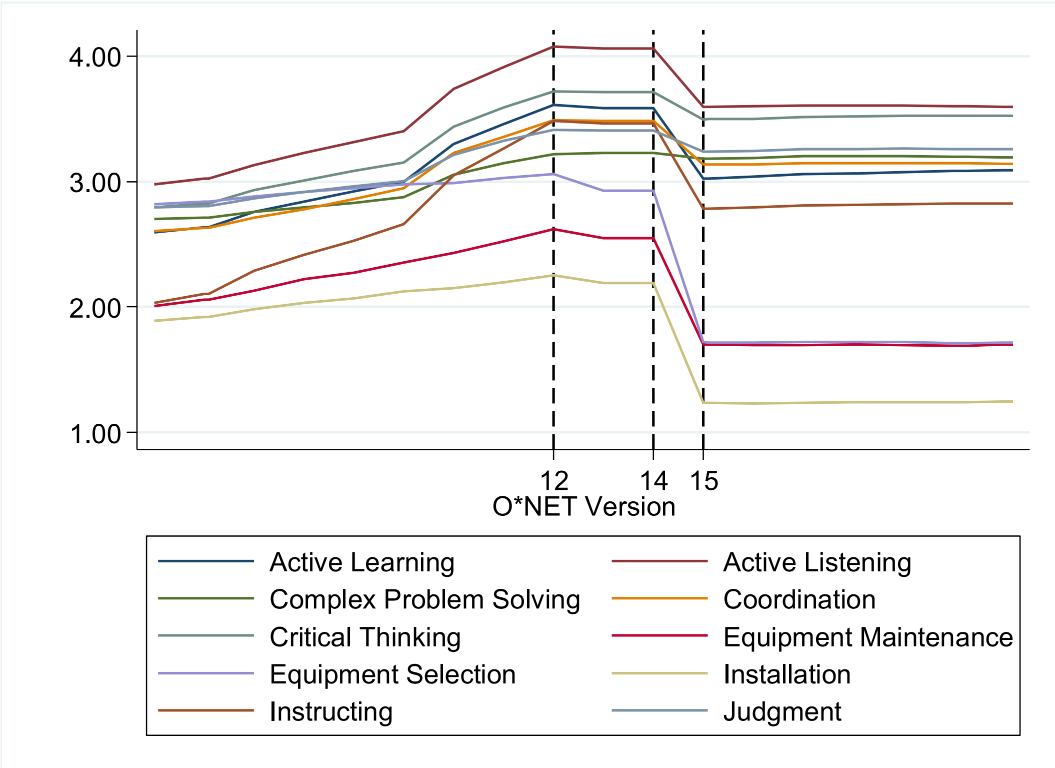


Figure B1 shows the impact of this on the raw O*NET data, plotting the mean importance measure in each O*NET version of a (random) selection of 10 of the 35 skill domains. A pattern amongst these skills is clear. Prior to version 15.0, the mean importance of each of these skills trended up, plateaued around versions 12.0 to 14.0, followed by a sharp fall between versions 14.0 and 15.0, then much more modest upward trends or a flat series thereafter.

Figure B2 illustrates how these inconsistencies and discontinuities can likely be explained by the issue of analysts and job incumbents. Firstly, the sharp drop in the mean value of skills observed in Figure 1 can be compared with the sudden shift from more than 80% of occupations being valued by incumbents in version 14.0, to all occupations being valued by analysts from version 15.0 onwards. If job incumbents systematically value the importance of a given skill to the job they do more than an analyst would, then this abrupt switch in version 15.0 explains the discontinuity in Figure B1.

Figure B2: Proportion of Occupations Valued by Job Incumbents and Analysts

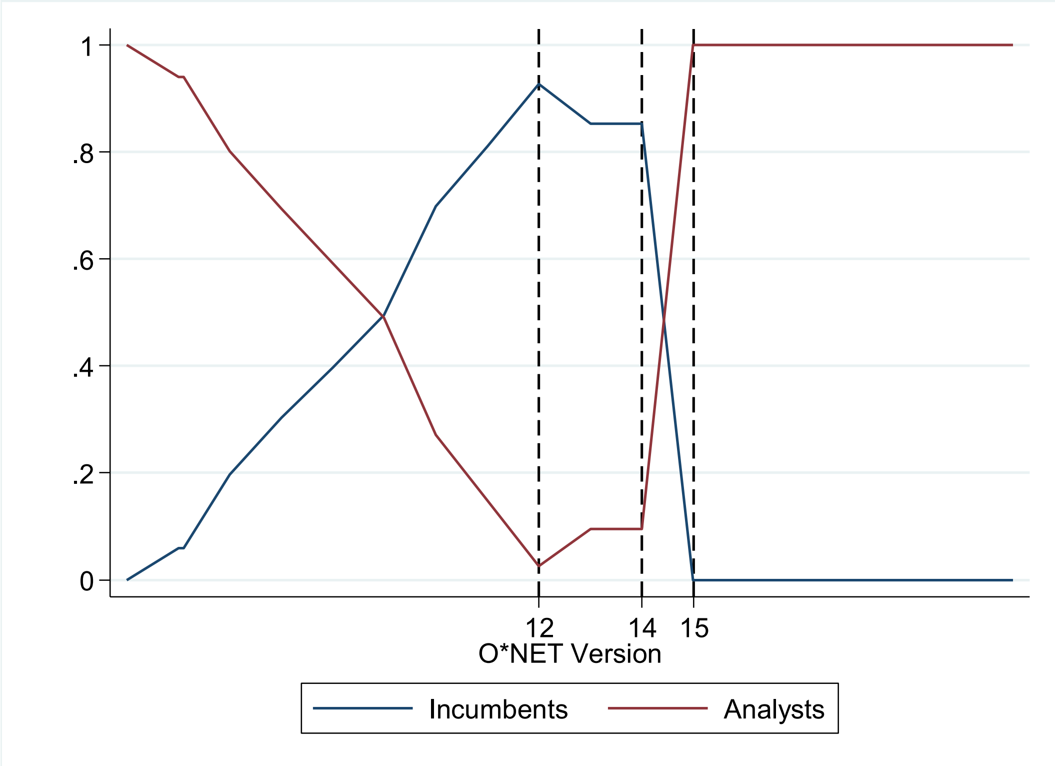


Figure B2 also explains the trend in skills observed prior to version 15.0 observed in Figure B1. The sharp rise in the mean importance of skills can in part be explained by switching observed in Figure B2 prior to version 12.0, whereby increasingly more occupations switch from analyst valuations to job incumbents. If job incumbents over-value skills relative to analysts then some of the increase in skills prior to version 12.0 are due not to genuine increases in the importance of these skills, but due to an increasing proportion of occupations having their skills valued by a source which systematically over-values relative to the other source. The plateau between versions 12.0 and 14.0 in Figure B1 also have a parallel in Figure B2, as during the same period the incumbent and analyst proportions are relatively stable. This further suggests that much of the trend we observe prior to version 12.0 in Figure B1 is an artefact of the changing source composition of the O*NET database prior to version 15.0.

To address the issue that skills are valued by multiple sources, we take the approach of converting incumbent-valued observations into values which are consistent with the analyst valuations. This rescaling approach involves estimating the extent to which incumbents under or over-value skills dimensions at a particular time relative to analysts. We then obtain analyst-consistent observations by removing the incumbent effect.

Our approach is to estimate the parameters of the following model, where we pool each of the O*NET versions into a single dataset:

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

In this model, the O*NET measure O 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 O was valued by an incumbent and zero otherwise, an interaction of the incumbency and time dummies, and an idiosyncratic error. We estimate this model, using occupation fixed-effects, separately for each of the 35 skill domains and for both importance and levels, giving a total of 70 regression models. Due to issues around the changing of the O*NET standard occupational classification, we estimate this model using the pooled O*NET data after the earlier O*NET versions have been reweighted to O*NET SOC-2010. This allows us to define the occupation fixed-effect on a consistent basis over our full sample period using the occupation SOC code.

$\hat{\beta}_1 + (\hat{\beta}_2 * \tau_t)$ is the estimated "incumbent effect" – the systematic over or under-valuation, in period t, made by job incumbents of the skill importance/level in the respective occupation. Subtracting this incumbent effect from the observed O*NET measure where the measure is provided by job incumbents gives us our rescaled skill measure, consistent with an analyst valuation of skill.

It is worth noting that because we rescale the data after the reweighting to O*NET SOC-2010, the I_{kt} variable need not necessarily be a binary variable. If, for example an O*NET SOC-2010 occupation maps onto two O*NET SOC-2009 occupations, one of which has incumbent-provided skills measures and the other analyst-provided skills measures, then the corresponding SOC-2010 occupation skill measure will be made up of a composite of both incumbent and analyst valuations of skill. This is very rarely the case. Of a total 3,268 non-zero observations of the incumbent dummy variable in the reweighted data, only 19 are not equal to one. These values (0.15% of the total observations for the 2002-2009 pooled O*NET data) between 0 and 1 are recoded to 1, so the interpretation of a value of 1 in the binary variable is now that the skills for the occupation are at least partially valued by job incumbents.

We use these incumbent-effect modified skills measures in our analysis, although we also investigate the robustness of our estimates to alternative treatments of this switch between incumbent and analyst measures of skills.

Table X: Measures of skills

Method	Advantages	Disadvantages
1. Qualifications and/or educational attainment	Objective Long-term trends available	Qualifications only have a loose link with job skills and thereby economic performance. Not all skills will be utilised in the labour market due to mismatch. And education may be a signal of ability rather than as a source of skills supply. Acquisition and depreciation of skills continues after education is completed. Learning at work important for acquisition of new skills and for updating existing skills. Hence the relationship between education and skills, and thereby economic performance, is complex – certainly measuring skills by education qualifications alone will be insufficient. International comparisons of attainment also difficult.
2. Education length	Objective Long-term trends available Internationally comparable	Variable quality of education – 1 year in country A is not the same as 1 year in country B Many of the criticisms of the use of qualifications in measuring skills can be similarly applied to the length of education. i.e. there is only a loose link between education and job skills.
3. Occupation	Easily available from Labour Force Surveys and/or censuses Internationally comparable (sometimes)	Occupational classifications have a better link with job skills, but even so, the hierarchy of occupations in the SOC for example is contestable, uncertain and changing. Moreover, over time, skills change within occupations.
4. Tests	Objective International comparisons possible	Formal assessments of skills through tests can only ever measure a limited range of skills (literacy and numeracy are typical) and are comparatively rate because of the costs of administering such testing. There has been criticism of the international comparability of universal testing even when it has been treated very carefully by

		researchers.
5. Self-assessment	Wide range of skills	Subjective, and so used very rarely. However, 5th sweep of NCDS records such measures. Major problem is that skill self-assessment is associated with self-esteem.
6. Job requirements	Wide range of skills Intimately connected with job	Job requirement measures increasingly being used. Obviously, job skill could differ from person skill (mismatch), and is subjective and will only measure skills of those in employment. But can use existing commercial job analysis data, as well as bespoke surveys. Examples include: O*NET (Occupational Information Network) in the US; German BIBB/IAB- and BIBB/BAuA Surveys on Qualifications and Working Conditions in Germany; UK Skills Surveys. These are surveys which ask individuals about the generic tasks and skills they use in their jobs and use those to infer the skills that they have. Of course, mismatch and underutilisation are still a problem, but they have permitted a much richer description of individuals’ skills, including soft/generic skills simply not captured by the other measures.

Source: Dickerson *et al* (2012), based on Green (2006).

Table X: 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 (2013)	Cognitive tasks Interpersonal tasks Physical job tasks (aka 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	Analytic activities	O*NET v. 13 (June 2008)	Analytic: average of 2	Jobs that require more

Reference	Taxonomy	Data	Measures/Methods	Notes/Findings
Spletzer (AER 2009)	Interpersonal activities Physical activities	41 work activities	O*NET activities Interpersonal: average of 2 O*NET activities Physical: 1 O*NET activity	analytical activity pay significantly higher wages, while those that require more interpersonal and physical activity pay lower wages.
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.

Source: Updated from Dickerson *et al* (2012).

Table 1: Decomposition of changing skill utilisation 2002 to 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 2: 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				
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 (11)		YES	YES	YES
Year dummies (14)				YES
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 3: Robustness Checks 1 – Aggregation Method

	Panel A: Importance Measures Only					
	Means			PCA		
	(1)	(2)	(3)	(4)	(5)	(6)
	Raw	Std.	Pctile	Raw	Std.	Pctile
Analytic skills	0.172***	0.077***	0.001***	0.027***	0.095***	0.003***
	(0.014)	(0.006)	(0.000)	(0.002)	(0.006)	(0.000)
Interpersonal skills	0.004	0.007	0.000	-0.004	-0.008	-0.000
	(0.010)	(0.005)	(0.000)	(0.003)	(0.006)	(0.000)
Physical skills	-0.058***	-0.037***	-0.000***	-0.014***	-0.034***	-0.001***
	(0.007)	(0.004)	(0.000)	(0.002)	(0.004)	(0.000)
N	4944	4944	4944	4944	4944	4944

	Panel B: Importance and Levels Measures					
	Cobb-Douglas Weighted Mean			PCA		
	(7)	(8)	(9)	(10)	(11)	(12)
	Raw	Std.	Pctile	Raw	Std.	Pctile
Analytic skills	0.007***	0.072***	0.002***	0.017***	0.086***	0.003***
	(0.001)	(0.007)	(0.000)	(0.001)	(0.007)	(0.000)
Interpersonal skills	0.002*	0.020***	0.001***	0.002	0.008	0.001**
	(0.001)	(0.006)	(0.000)	(0.002)	(0.007)	(0.000)
Physical skills	-0.008***	-0.043***	-0.002***	-0.010***	-0.033***	-0.001***
	(0.001)	(0.004)	(0.000)	(0.001)	(0.004)	(0.000)
N	4934	4934	4934	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. All regressions in this table are estimated using the same specification as column (4) in 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 information.

Table 4: Robustness Checks 2 – Alternative Transformations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Raw Skills	LFS Log Mean	LFS Mean Log	LFS 96-97	Census 01	LFS JM 07
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)
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. The column (1) regression coefficients repeat the specification reported in column (4) of Table 2. Column (2) uses raw skills data, not corrected for incumbent/analyst valuation. 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 SOC-2000 to SOC-2010 with weights using the respective 3 dual-coded datasets.

Table 5: Robustness Checks 3 – Earnings Function Specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Male	Female	Full Time	1-Digit SOC	2011-16	WLS
Analytic skills	0.172***	0.290***	0.375***	0.236***	0.119***	0.308***	0.160***
	(0.014)	(0.017)	(0.016)	(0.015)	(0.014)	(0.031)	(0.013)
Interpersonal	0.004	-0.018	-0.027**	-0.004	0.004	0.066***	0.009
	(0.010)	(0.013)	(0.013)	(0.011)	(0.010)	(0.025)	(0.010)
Physical skills	-0.058***	-0.095***	-0.141***	-0.072***	-0.042***	-0.055***	-0.061***
	(0.007)	(0.008)	(0.008)	(0.007)	(0.007)	(0.013)	(0.006)
N	4944	4362	3774	4647	4944	1918	4944

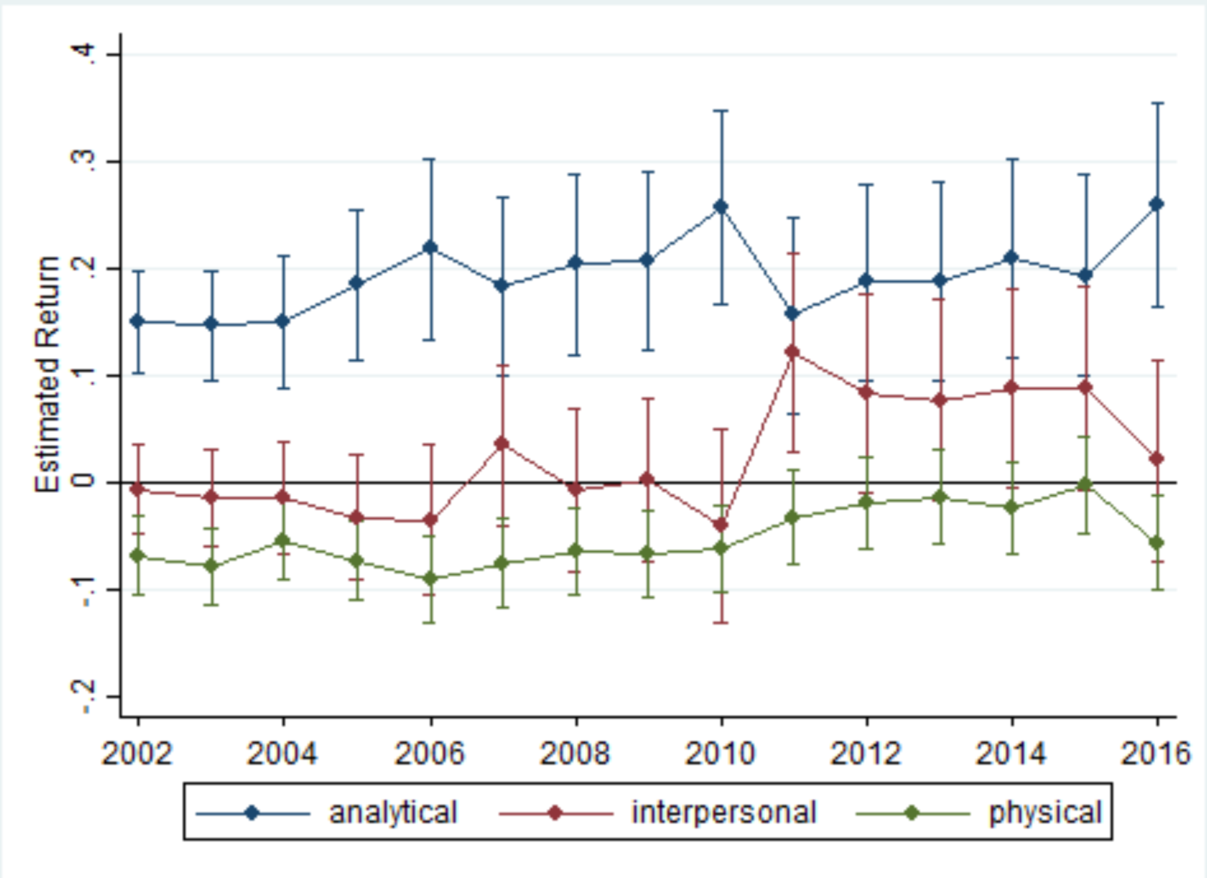
Notes:
1. The dependent variable is log average real hourly wages.
2. Standard errors in parentheses: * p<0.10; ** p<0.05; *** p<0.01.
3. The column (1) regression coefficients repeat the specification reported in column (4) of Table 2. Column (2) constructs the outcome and independent variables from 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 observations only. Columns (5) includes 1-digit SOC occupation fixed effects. Column (6) estimates only for 2011 to 2016. Column (7) uses WLS rather than OLS.

Table 6: WLS Estimates – Full Results

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	WLS		Iterative WLS		
			Iteration 1	Iteration 2	Iteration 3	Iteration 4
Analytic skills	0.172***	0.142***	0.160***	0.160***	0.160***	0.160***
	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Interpersonal	0.004	0.017*	0.009	0.009	0.009	0.009
	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
Physical skills	-0.058***	-0.072***	-0.061***	-0.061***	-0.061***	-0.061***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
	ε^2	ε^2	ε^2	ε^2	ε^2	ε^2
1/N	0.186***	0.222***	0.204***	0.203***	0.203***	0.203***
	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
Constant	0.019***	0.018***	0.018***	0.018***	0.018***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	4944	4944	4944	4944	4944	4944

Notes:

Figure 1: Trends in the returns to skills 2002-2016



Note:

1. These are regression coefficients using the specification in Table 2, column (4), supplemented by interactions of each of the 3 skills indices with year dummies.