Developing Occupational Skills Profiles for the UK: A Feasibility Study

Evidence Report 44
February 2012

Intelligence > Investment > Impact

Developing Occupational Skills Profiles for the UK: A Feasibility Study

Andy Dickerson
University of Sheffield

Rob Wilson
University of Warwick

Genna Kik
Senior Manager

Debra Dhillon
Manager

UK Commission for Employment and Skills

February 2012

Foreword

The UK Commission for Employment and Skills is a social partnership, led by Commissioners from large and small employers, trade unions and the voluntary sector. Our mission is to raise skill levels to help drive enterprise, create more and better jobs and promote economic growth. Our strategic objectives are to:

- Provide outstanding labour market intelligence which helps businesses and people make the best choices for them;
- Work with businesses to develop the best market solutions which leverage greater investment in skills;
- Maximise the impact of employment and skills policies and employer behaviour to support jobs and growth and secure an internationally competitive skills base.

These strategic objectives are supported by a research programme that provides a robust evidence base for our insights and actions and which draws on good practice and the most innovative thinking. The research programme is underpinned by a number of core principles including the importance of: ensuring '**relevance'** to our most pressing strategic priorities; '**salience'** and effectively translating and sharing the key insights we find; **international benchmarking** and drawing insights from good practice abroad; **high quality** analysis which is leading edge, robust and action orientated; being **responsive** to immediate needs as well as taking a longer term perspective. We also work closely with key partners to ensure a **co-ordinated** approach to research.

This current study, which was undertaken by Andy Dickerson at the University of Sheffield and Rob Wilson at the University of Warwick, documents the development and feasibility of a new methodology to improve occupational skills profiles at the Standard Occupational Classification (SOC) unit group (4-digit) level for the UK.

By matching US occupations to UK occupations at a very detailed level, the authors have been able to exploit information contained in the US Occupational Information Network (O*NET) system for US occupations, to generate occupational skills profiles for the UK.

The work offers the potential to do further work and create a much more detailed depiction of skills utilisation, and changes in utilisation, than is currently available in the UK. Consequentially, the findings so far will be of interest to agencies and organisations which have an interest in skills, and are concerned about their importance and their impact for individual labour market outcomes as well as macro-economic performance.

Sharing the findings of our research and engaging with our audience is important to further develop the evidence on which we base our work. Evidence Reports are our chief means of reporting our detailed analytical work. Each Evidence Report is accompanied by an executive summary. All of our outputs can be accessed on the UK Commission's website at www.ukces.org.uk

But these outputs are only the beginning of the process and we will be continually looking for mechanisms to share our findings, debate the issues they raise and extend their reach and impact.

We hope you find this report useful and informative. If you would like to provide any feedback or comments, or have any queries please e-mail info@ukces.org.uk, quoting the report title or series number.

Lesley Giles
Deputy Director
UK Commission for Employment and Skills

Table of Contents

List of Figures and Tables

Executive Summary

Skills are a major policy priority both nationally and internationally. Yet we only have very imperfect measures of the skills available and in use in employment in the UK today. This report explores the feasibility of the development of a new and comprehensive set of detailed, multi-dimensional occupational skills profiles for the UK which describe the skills required by employers and used by individuals in the modern workplace. These occupational skills profiles can have a myriad of potential uses and users, including providing a much richer and deeper understanding of the changing patterns of the demand for skills in the UK, and informing individuals and those who advise them on the skills that are useful in employment today.

Skills can be measured in a variety of different ways. The two most commonly employed measures are: (i) the qualifications that individuals have previously acquired; and (ii) the occupational classification of the jobs that they do. These both have the considerable virtue of being relatively simple to measure, but are poor proxies for the actual skills required by employers and used by individuals. Indeed, when asked about skills and skills needs, employers tend to focus on aspects of individuals and jobs other than their qualifications or occupations. These other aspects have been variously termed generic, key or core skills and attributes. Examples include communication, problem solving, numeracy and literacy skills. They are rather more difficult to measure precisely, although considerable progress has been made in recent years in the UK Skills Surveys using questionnaires focussed on the nature of the tasks that individuals perform in their jobs[1]. However, the Skills Surveys are unable to provide a very detailed or comprehensive picture, mainly due to their limited scope and small sample size.

In contrast to the comparative lack of such information for the UK, the US-based Occupational Information Network (O*NET) system provides almost 250 measures of skills, abilities, work activities, training, work context and job characteristics for each of around 1,000 different US occupations (based on a modified version of the US Standard Occupational Classification), with information gathered from both job incumbents through standardised survey questionnaires, as well as assessments by professional job analysts. This information is also linked to information on current employment levels, rates of pay and future employment prospects. The O*NET system replaced the long-established Dictionary of Occupational Titles (DOT), with the first complete version of O*NET becoming available in 2008.

[1] The Skills Surveys are a representative series of surveys focused on individuals living in Britain (and for the latest survey, the UK). The surveys gather information on the skills used at work via survey questions directed at workers themselves, using several different measures of work skills, some of which have been used in previous surveys. The surveys have taken place in 1997, 2001 and 2006 - see Felstead et al (2007) for details.

Ideally we would like to have an O*NET-type system for the UK which could provide a broad set of descriptors of the skills that people utilise in their jobs. We could then use the trends in skills to inform public policy about the skills that are utilised and rewarded in employment in the UK today. This would provide a much richer description than our current measures of skills such as qualifications or simply the fixed and uni-dimensional hierarchy of the Standard Occupational Classification (SOC). However, the costs of developing such as system would be considerable, both financially and in terms of time. As an alternative therefore, this project investigates the feasibility of developing a mapping between the US SOC and UK SOC taxonomies in order to exploit the information that is already collected for the US O*NET system.

The primary objective of this project is therefore to construct a systematic 'mapping' between the occupational classification utilised in O*NET and the UK Standard Occupational Classification (SOC). We then develop a methodology for assigning the detailed content of the O*NET system to the UK SOC in order that profiles of the skills and other job characteristics used in all UK occupations in the UK SOC can be developed.

The project proceeds in four main stages. The first stage (Chapter 3) is to match the occupational taxonomy in O*NET (which is a slightly extended version of US SOC) to that of the UK SOC. This has been accomplished by using CASCOT (Computer Assisted Structured COding Tool) which we use to match almost 57,000 US job titles to around 28,000 UK job titles. CASCOT produces a systematic mapping between US job titles and UK job titles, with scores between 0 and 100 which reflect the quality of the match. Each job belongs to a specific SOC, and hence the job–to–job matching can be aggregated to produce a corresponding O*NET SOC–to–UK SOC matrix of matching scores. In the second stage (Chapter 4) of the project, we use these scores together with the relative employment in the O*NET occupation to produce a matrix of weights which enables us to match between O*NET occupations to 4-digit (unit group) UK occupations. In the third stage (Chapter 5), we select a range of dimensions or descriptors of job skills, abilities and the other occupational characteristics on the O*NET system, and use the weights to assign these measures to each 4-digit UK occupation. We illustrate the methodology with two examples. The first is the classic 'data', 'people' and 'things' classification of skills as used by Autor *et al* (2003) amongst others, which we construct from the 35 skills descriptors in O*NET using factor analysis. The three categories of data (cognitive skills), people (interpersonal skills) and things (physical job tasks) reflect the quite different sets of skills that individuals may use in their jobs. The second example is a measure of STEM (science, technology, engineering, and mathematics) skills which we generate by simple averaging of 8 relevant descriptors selected from the abilities and skills domains. These include: Deductive Reasoning, Information Ordering, Mathematical Reasoning, and

Number Facility (from the abilities domain) and Mathematics, Science, Technology Design, and Programming (from the skills domain). The final stage of the project (Chapter 6) undertakes some evaluation and validation of the methodology in order to assess the quality and robustness of the resulting occupational skills profiles.

Our findings suggest that it is indeed possible to create such a mapping between US and UK occupational taxonomies, and thus to be able to assign the job tasks, skills and other content of the US O*NET system to the matched UK occupations. This mapping appears to be quite robust to the methodological approach employed. When used to generate occupational skills profiles for data-people-things and for STEM skills at the 4-digit (unit group) level of SOC2010, the resulting occupational profiles appear to be sensible and reasonable and conform to our prior expectations. Moreover, when we use the mapping to derive measures of required qualifications and training time and compare these with similar measures taken from the 2006 Skill Survey (Felstead et al, 2007), the correspondence between the two different sources are very high – at least at the SOC Major Group level – giving us further confidence in the validity and robustness of the methodology we have developed.

Exploiting the mapping that we develop between O*NET SOC and UK SOC enables the multi-dimensional O*NET system to be used to generate a comprehensive database of occupational skills profiles for the UK, providing a much more detailed depiction of skills utilisation, and changes in utilisation, than is currently available for the UK. This is crucial if we are to really develop an understanding of skill utilisation and changing skill needs in the UK. The profiles are likely to be of considerable interest to agencies which have an interest in skills and their importance as well as their impact for individual labour market outcomes, and also for macro-economic performance.

Additional potential uses of the methodology developed in this report include:

- An assessment of trends in skills demand (as recorded by their changing utilisation in employment), and estimating future skills demand.
- Supplying useful information to Information, Advice and Guidance (IAG) practitioners – and also to individuals – on the types of skills that are necessary for, and useful in employment today, and are likely to be of importance and value in the future in terms of labour market outcomes.
- Estimating the value of skills in employment.
- Extending the information available to the Migration Advisory Committee (MAC) on the measurement of skills, and on the specific skills that are in shortage

In summary, this report is a **technical paper** describing the complex matching procedure that we have undertaken between the O*NET and UK SOC. It examines the feasibility of matching US and UK occupations, and includes an evaluation of the quality of the matching and some assessment of the sensitivity of the resulting profiles to the various assumptions that are necessarily made at different stages of the process. We also provide some examples of the occupational skills profiles that can be constructed, and compare these profiles with some other extant measures of job skills and activities in order to provide some assessment of the validity of the methodology that we have developed.

1 Introduction

This project examines the feasibility of constructing a detailed set of occupationally-based 'profiles' describing the many different skills that are used in employment in the UK. These occupational skills profiles are intended to be multi-dimensional and therefore would provide a much richer description and measurement of skill demand and skill utilisation than is possible using the existing methods of measuring skills that are commonly employed in UK research and policy. We make use of the most detailed and comprehensive assessment of skills used in employment that exists as provided by the US Occupational Information Network (O*NET) system. O*NET includes self-reported assessments by job incumbents based on questionnaire surveys, as well as professional assessments by job evaluation analysts, across 239 different dimensions[2], including qualifications required, practical and technical skills, and 'soft skills' such as communication skills. For two-thirds of the dimensions, both the level and the intensity of their use are recorded. The O*NET system gathers this information for almost 1,000 separate occupations. By matching US jobs to comparable UK jobs in a systematic and transparent manner, we adapt the O*NET skills descriptors to jobs (i.e. occupations) in the UK.

The primary task in this project is thus to construct a systematic 'mapping' between the US O*NET and UK occupational classifications[3], and then to assign the skills measures and other content of the US O*NET system to the matched UK occupations.

There are four main stages to the project:

- **Stage 1:** Matching between the O*NET occupational classification and the UK SOC;
- **Stage 2:** Assigning job skills and abilities provided within O*NET to UK occupations based on weighting and aggregating according to the quality of the match;
- **Stage 3:** Summarising job skills and abilities to produce useful taxonomies for UK occupations;
- **Stage 4:** Assessing and validating the matching and assignment by making comparisons with other measures of job skills and activities.

The remainder of this report is structured as follows. Chapter 2 describes the background and methodology that we use, including a brief outline of the O*NET system. Chapters 3, 4, 5 and 6 report the details from each of the four stages of the project as outlined above.

[2] These 'dimensions' include abilities, skills, work activities, context and styles, tasks, knowledge, education requirements etc.

[3] Note that the O*NET occupational classification differs slightly from the US standard occupational classification in that the former includes additional occupations within some of the US SOC occupations. Full details of the relationship between the two taxonomies are provided in Annex C.

Finally, we conclude in Chapter 7 with some suggestions for further developments of the occupational skills profiles. We also briefly describe some of the potential uses and applications of the occupational skills profiles to the UK Commission for Employment and Skills (UK Commission) and wider research and policy community concerned with skills that are possible using the methodology that we have developed and described in this report.

2 Background and Methodology

Chapter Summary

This Chapter describes the various measures of skills demand and utilisation that are in current usage in research and policy, and briefly describes the US O*NET system that forms the basis of the measures that are developed in this project.

The methodology to be employed, and the four distinctive stages of the project which form the remaining substantive Chapters of the report are then outlined.

2.1 The Measurement of Skills

The importance of skills in modern economies and in economic policy discourse is widely acknowledged. Changing skills are important for example at both micro – individual – level for the distribution of earnings, and at the macro level, for explanations of productivity and growth – especially of the endogenous growth kind. Despite the fundamental importance, both theoretically and practically, of skills to the discourse surrounding the knowledge economy, procedures for measuring skills are comparatively under-developed in the UK. But much more effort has been devoted to this issue in the US, with their O*NET system, which has seen over 50 years of investment and development[4]. This project explores the potential for the O*NET system to be exploited in a UK context.

We can identify (at least) 6 distinct ways of defining and measuring skills in contemporary research. These are summarised in Table 1. There are a number of advantages and disadvantages associated with each of the different conceptualisations of skills that are commonly employed as listed in the table. Skills are multi-dimensional, socially constructed, intangible and often unobservable, and each of the different measures of skills can be argued to have some relative merits and demerits associated with them.

[4] A comprehensive description and review of the O*NET system has been recently published (Tippins and Hilton, 2010).

Table 1: Measurement 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, the 5th interview wave of the National Child Development Survey (NCDS) does record 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: *Based on Green (2006).*

In addition to the inherent difficulties associated with measuring skills, different disciplines have somewhat different conceptions of 'skill'. However, within economics at least, skill is a quite general concept. It can be considered as a characteristic that can be acquired, and that enables individuals to produce valued services in work – i.e. is an element of human capital (and therefore is conceptually the same as health, since both can be invested in, and enhance the stream of revenues that can be earned). It is also conceptually equivalent to behavioural traits – such as honesty and motivation – since these can also be 'acquired' and are productive in employment.

Ideally we would like to have objective, internationally-comparable measures of skills. Of the different measures and conceptualisation of skills in Table 1, the most commonly utilised are the qualifications that individuals have acquired (Row 1) and the occupations of the jobs that they undertake (Row 3). These both have the considerable virtue of being relatively simple to measure and afford some international comparability, particularly when international classification systems are employed such as ISCED which is maintained by UNESCO,[5] and ISCO which is compiled by the ILO[6]. However, qualifications in particular can be regarded as a poor measure of skills used in employment: they are typically gained before individuals enter the labour market, and any skills that are specifically acquired in the process of gaining any particular qualification soon depreciate. Rather, qualifications arguably provide a means of entering particular employments or employment levels.

The skills subsequently gained while in employment by learning-by-doing, through formal and informal on-the-job training, or in any subsequent off-the-job training and then utilised in employment are those that are of primary interest for individuals and employers, and for public policy. Individuals seeking to move jobs, firms seeking new employees, agencies responsible for assisting people back into work, training providers, HR managers and policy makers responsible for identifying skills shortages, trends and future requirements all require measure of the skills that are used, valued and rewarded in employment.. Moreover, when asked about skill needs, employers increasingly focus not on qualifications, but on other aspects which have been variously termed key, generic, soft, or core skills. Examples include: numeracy and literacy; communication skills; team-working; problem solving etc. Qualifications are, therefore, at best, only a poor proxy for the skills that individuals have acquired or utilise in their jobs. They are also a weak measure of the attributes that individuals possess that are rewarded in the labour market.

Occupations arguably provide a more meaningful summary of the skills that individuals are using in employment, particularly where the occupational classification is hierarchical so that higher occupational levels can be associated with higher levels of skills. However,

[5] http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx
[6] http://www.ilo.org/public/english/bureau/stat/isco/index.htm

it still fails to record the actual skills that are being utilised, and nor does it effectively recognise that jobs are typically bundles of skills (Sattinger, 1979, 1993) and thus the skills being utilised in any job cannot be captured by a uni-dimensional indicator such as the SOC code[7]. And, of course, even within occupations, skills can differ – for example, according to sector or organisation size.

More recently, the advantages of the so-called 'job requirements' approach (Table 1, Row 6) have found increasing favour. These measure skills that are being used by individuals in their jobs by their (self-reported) answers to questions regarding the degree (and sometimes intensity) to which their jobs require them to perform particular tasks. Examples include the UK Skills Surveys (Felstead *et al*, 2007). In these surveys, individuals were asked to rank on a 5-point Likert scale (running from 'essential' to 'not at all important') how important a range of 35 tasks were in their jobs. For example, respondents were asked:

> In your job, how important is …
> … paying close attention to detail?
> … dealing with people?
> … instructing, training or teaching people, individually or in groups?
> … making speeches or presentations?
> … persuading or influencing others?
> … selling a product or service?
> … counselling, advising or caring for customers or clients?
> etc

However, given the relatively small scale of the Skills Surveys (the 2001 and 2006 surveys comprise approximately 4,500 and 8,000 individuals in each survey respectively), it is not possible to use the information to provide a comprehensive assessment of the skills utilised in all occupations in the UK at other than a quite aggregated level. Thus the Skills Surveys are unable to capture much of the heterogeneity within and between jobs. Moreover, the range of job skills recorded is limited to the 35 dimensions captured by the particular job task questions listed above, (together with questions on the level and intensity of computer use).

In contrast, the US has long devoted considerable resources to measuring and recording the skills used in employment in America. Starting with the first edition of the Dictionary of Occupational Titles (DOT) published in 1939, this has evolved considerably over time. O*NET has recently replaced the DOT, and the first full version of this new system was

---

[7] The SOC manual does also list the tasks involved in different occupations together with a general indication of the level of qualifications required.

published in June 2008. A brief outline is provided in the next sub-section while further details are in Annex A.

2.2 The O*NET System[8]

The US O*NET system is the main source of occupational competency information in the US. It utilises a modified version of the US SOC to record information across 6 different 'domains' as outlined in Table 2 for around 1,000 different occupations. It has been almost 20 years in development as a replacement to the DOT system, with the first complete version becoming available in June 2008 as noted above. Much of the information in the O*NET 'content model' is gathered from self-reported assessments by job incumbents based on standardised questionnaire surveys, supplemented by professional assessments by job evaluation analysts, and it is this content model that is the focus here. Additional information on pay, and on recent employment trends and future projections from the Bureau of Labor Statistics (BLS) is also included in the O*NET system.

In total, 239 different dimensions or 'descriptors' of skills and job characteristics including: qualifications required; practical and technical skills; a wide range of soft skills such as communication skills, stamina etc; as well as details of the tasks involved in the job (see Table 3). For the four domains of Knowledge, Skills, Abilities and Work activities, both the 'Importance' and 'Level' of each skill or characteristic being measured is recorded. As an example, Figure 1 presents the rating scales for the 'Reading Comprehension' skill. There has been some criticism of the level scale and the 'anchors' that are used to define the scale points – see Handel (2010) for example – and the fact that importance and level responses are typically highly correlated. However, despite this criticism, clearly the number and range of items recorded are extensive. Indeed, part of the process of defining occupational skill profiles for the UK is to reduce the number of dimensions to a more limited set of descriptors in a meaningful and appropriate manner.

[8] Further detail on the O*NET system is provided in Wilson (2009) which is in turn summarised in Annex A of this report. O*NET is also described and evaluated in the comprehensive review recently published by Tippins and Hilton (2010).

Table 2: Summary of O*NET Content Model

		DOMAIN	ELEMENT DESCRIPTION
1		**Worker Characteristics**	
1	A	Abilities	Enduring attributes of the individual that influence performance
1	B	Interests	Preferences for work environments and outcomes
1	C	Work Styles	Personal characteristics that can affect how well someone performs a job
2		**Worker Requirements**	
2	A	Basic Skills	Developed capacities that facilitate learning or the more rapid acquisition of knowledge
2	B	Cross-Functional Skills	Developed capacities that facilitate performance of activities that occur across jobs
2	C	Knowledge	Organized sets of principles and facts applying in general domains
2	D	Education	Prior educational experience required to perform in a job
3		**Experience Requirements**	
3	A	Experience and Training	If someone were being hired to perform this job, how much would be required?
3	B	Basic Skills - Entry Requirement	Entry requirement for developed capacities that facilitate learning or the more rapid acquisition of knowledge
3	C	Cross-Functional Skills - Entry Requirement	Entry requirement for developed capacities that facilitate performance of activities that occur across jobs
3	D	Licensing	Licenses, certificates, or registrations that are awarded to show that a job holder has gained certain skills. This includes requirements for obtaining these credentials, and the organization or agency requiring their possession
4		**Occupational Requirements**	
4	A	Generalized Work Activities	General types of job behaviors occurring on multiple jobs
4	B	Organizational Context	Characteristics of the organization that influence how people do their work
4	C	Work Context	Physical and social factors that influence the nature of work
4	D	Detailed Work Activities	Detailed types of job behaviors occurring on multiple jobs
5		**Occupation-Specific Information**	
5	A	Tasks	Occupation-Specific Tasks
5	B	Tools and Technology	Machines, equipment, tools, software, and information technology workers may use for optimal functioning in a high performance workplace
6		**Workforce Characteristics**	
6	A	Labor Market Information	Labor Market Information
6	B	Occupational Outlook	Occupational Outlook

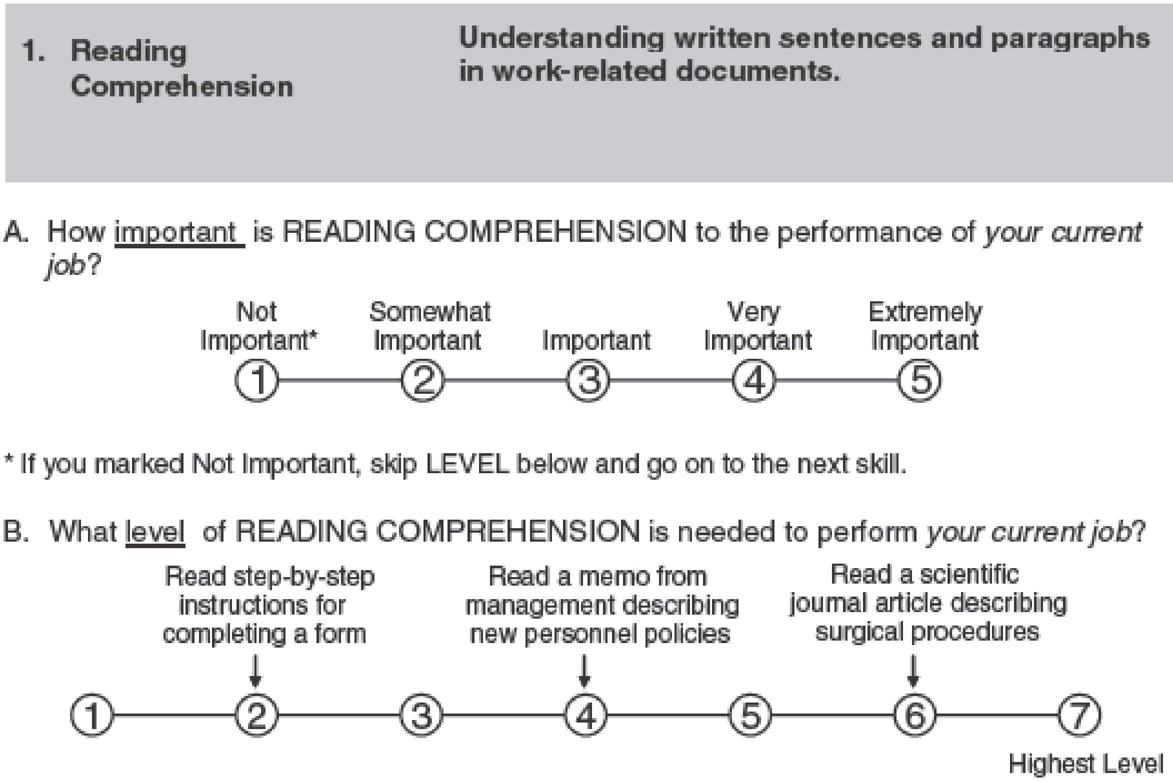
Source: *O*NET Content Model*

Table 3: Description of Main O*NET Questionnaires

	Survey instrument	Main content	No. of items ('descriptors')	Data source	Information recorded
1	**Education and training**	required education, related work experience, training	5	Job incumbents	Levels
2	**Knowledge**	various specific functional and academic areas (e.g., physics, marketing, design, clerical, food production, construction)	33	Job incumbents	Importance and levels
3	**Skills**	reading, writing, math, science, critical thinking, learning, resource management, communication, social relations, technology	35	Analysts	Importance and levels
4	**Abilities**	writing, math, general cognitive abilities, perceptual, sensory-motor, dexterity, physical coordination, speed, strength	52	Analysts	Importance and levels
5	**Work activities**	various activities (e.g., information processing, making decisions, thinking creatively, inspecting equipment, scheduling work)	41	Job incumbents	Importance and levels
6	**Work context**	working conditions (e.g., public speaking, teamwork, conflict resolution, working outdoors, physical strains, exposure to heat, noise, and chemicals, job autonomy)	57	Job incumbents	Levels
7	**Work style**	personal characteristics (e.g., leadership, persistence, cooperation, adaptability)	16	Job incumbents	Importance
	TOTAL		**239**		

Source: *Handel (2010), p.15, and Tippins and Hilton (2010), p.72, p.74.*

Figure 1: O*NET Reading Comprehension Rating Scale



Source: *Tippins and Hilton (2010), p.75.*

The questionnaires are administered by post by the US Bureau of Labour Statistics and, in order to avoid survey fatigue, job incumbents are only asked to complete a random selection of the questionnaires listed in Table 3. 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. Sample sizes are not easy to identify from the information made available since only occupation means rather than the individual micro-data data are publicly released, but one estimate is that there is an average of 31,000 responses for each of the 239 descriptors gathered from around 125,000 returned questionnaires, thus representing an average of just under 40 respondents per descriptor for each of the separate occupations in the O*NET system (Handel, 2010). One of the main strengths of the O*NET design is that it is being updated on a 5-year rolling basis so that changes in skills utilisation *within* occupations will be able to be discerned. A completely new set of descriptors for all occupations will therefore be available by 2013.

The O*NET system provides an enormous amount of information on skills used in occupations. National Occupational Standards (NOS) in the UK, and similar systems at a pan-European level, provide some of the same information[9,10]. However, generally they

[9] The German system – the BIBB/IAB- and BIBB/BAuA Surveys on Qualifications and Working Conditions in Germany is particularly notable and covers much of the same ground as the US O*NET system.

do not form part of an integrated labour market information (LMI) system in the same way that O*NET does, and O*NET goes well beyond NOS by gathering information directly and indirectly from employees themselves, rather than simply listing the skills, knowledge and understanding that employers deem are required to perform competently in any given occupation. Furthermore, O*NET provides a wealth of information which is not available in the UK. This project develops a detailed mapping from US to UK occupational categories and thereby enables us to exploit the information in O*NET for the first time.

2.3 Methodology

Ideally we would like to have an O*NET-type system for the UK – that is a broad set of descriptors of the skills that people utilise in their jobs. We could then use the trends in skills use to inform public policy about the skills that are needed in employment (and rewarded in employment) in the UK today. This would provide a much richer description than our current measures of skills such as qualifications or simply the fixed and uni-dimensional hierarchy of the Standard Occupational Classification (SOC)[11].

In the absence of such a system specifically for the UK, our ambition in this project is to adapt the various measures of skills in the US O*NET content model to the UK SOC in order to provide the same level of detail in terms of both the occupations that can be separately identified and described, and the range of skills descriptors that are available. In essence, we combine the advantages of the occupational measures of skills (Table 1, Row 3) with those of the job-requirement approach (Table 1, Row 6), by exploiting the considerable effort and investment that the US has made in the development of their O*NET system. Job-requirement measures provide us with a better indicator of the demand for skills than qualifications or occupational classifications. Thus we use the skills dimensions that are recorded in the O*NET system (or at least a subset of them) and apply these to occupations in the UK. Assuming that sensible matches can be obtained between occupations, these can then be used to provide a set of descriptors of the skills used in jobs (occupations) the UK.

While there are many differences in occupational classification between the UK and the US, there are also many similarities. The review in Wilson (2009) suggests that, although O*NET has been designed specifically for the US, there is considerable potential for it to be exploited in other countries. Many of the characteristics of jobs are common across countries. Indeed, the O*NET system has already been applied (with only minimal modification) to a number of countries outside the US including Australia, Czech Republic, New Zealand, China and Hong Kong. For example, in work currently being

[10] O*NET descriptors have also been shown to have application outside the US (Taylor *et al,* 2008).
[11] A particular criticism of SOC is that skills are changing within jobs so that simply using the same categories over time is not really capturing changing skills usage and utilisation.

undertaken for Cedefop, EPC at Charles University Prague are working with Warwick IER to explore the potential for exploiting O*NET at a pan-European level (Koucký *et al*, 2010; Wilson, 2010).

Our project progresses in number of distinct phases. These are briefly outlined in the following four sub-sections:

2.3.1 Stage 1: Matching

The first stage is to provide a match between the occupational taxonomy in O*NET to that of the UK SOC. This has been undertaken using specialist software, originally developed by Warwick IER, called CASCOT (Computer Assisted Structured COding Tool)[12]. As part of our project, we have separately developed a classification database or 'dictionary' specifically for O*NET, and this has potential to be developed further in the future. CASCOT enables us to produce a systematic mapping between the jobs in the US and jobs in the UK, with a matrix of scores (between 0 and 100) which reflect the quality of the match. Further details are in Chapter 3 of this report.

2.3.2 Stage 2: Assigning Job Skills and Abilities to UK Occupations

Secondly, we have accessed the O*NET content model database which describes the anatomy of every occupation identified in the O*NET system. The data includes information on the distinctive characteristics of each occupation, including the knowledge, skills and abilities required, and the activities and tasks performed. These are the so-called 'descriptors' which are grouped into a number of 'domains' according to how the data are collated. We are utilising Version 15.0 of the O*NET database which was released in June 2010 to researchers. A key advantage of using Version 15.0 is that the skills domain measures for over 800 occupations were all updated by job analysts. Since the measures of skills was the area where job incumbents were argued to have found most difficulty in assessing their relative levels, this domain is now exclusively covered by professional job analysts' responses. Using Version 15.0 of the database will mean that we have access to their comprehensive and completely updated set of measures on skills. The skills domain is also one of the four domains in which both importance and level of each descriptor activity is recorded.

Given the matching matrix constructed in the first stage, we can then assign the skills and other descriptors in the O*NET system to UK occupations. We weight the O*NET descriptors, with weights dependent on: (i) the quality of the match as reflected in the CASCOT scores; and (ii) on the relative 'importance' of the occupation (based on

[12] We have be able to obtain expert advice from Professor Peter Elias (Warwick IER) who was responsible for developing CASCOT as well as revising the UK SOC into SOC2010, and also from Ritva Ellison (Warwick IER) who has been involved in the development of the CASCOT software. We would like to thank them both for their input into this project.

occupational employment shares). This weighting thus takes into account how close the match is between any particular pair of O*NET and UK SOC occupations, but also the relative employment of the O*NET occupation in question. Thus, close matches with occupations which account for a larger number of workers are therefore given greater weight than close matches with minor occupations (in terms of employment), and both are given greater weight than weaker matches with the same occupations.

2.3.3 Stage 3: Summarising Job Skills and Abilities for UK Occupations

Given there are 239 dimensions or descriptors of skills and job characteristics recorded on the O*NET system, it is clearly important to select and summarise them in some way in order to provide useful information for the UK. Our primary focus in this project will be on just the skills and abilities domains, but this still yields (35+52=) 87 descriptors (each with both level and importance reported for each descriptor) for each O*NET occupation[13]. Statistical methods such as factor analysis (see, for example, Kim and Mueller, 1978) or cluster analysis (see, for example, Everitt, 1993) can be used to combine the descriptors, or particular skills of interest could be selected from the myriad of those available. Examples of such aggregations for skills used previously and which have been found to be useful include that introduced by the Dictionary of Occupational Titles (DOT) which classified skills broadly into 'data', 'people' and 'things'. This 3-way classification has been used repeatedly in the literature e.g. Autor *et al* (2003), and Autor and Handel (2009) (with more descriptive labels of 'cognitive skills', 'interpersonal skills' and 'physical job tasks'). A similar approach is used by Abraham and Spletzer (2009) to define 'analytic skills', 'interpersonal skills' and 'physical skills' from just the 41 work activities measures by selecting a limited subset of descriptors amongst these.

Having selected a set of dimensions to focus upon, and an aggregation method for those dimensions, this will enable us to characterise all UK occupations in terms of a set of skills and other characteristics. These will form our **occupational skills profiles**.

2.3.4 Stage 4: Assessment and Validation

The final stage of the project is a validation exercise to assess the quality and robustness of the occupational skills profiles, and to use the profiles to illustrate some different ways in which such information can extend our understanding of skills, their importance, and their utilisation.

[13] We focus on these two domains only partly for parsimony, but mainly because these are the most useful elements of O*NET for our purposes. Other domains such as those covering prior educational experience, certification./licensing, labour market information and occupational outlook are country–specific (and the information could be sourced form eg LFS data). However, the analysis presented here could readily be applied to some of the other domains that we are not considering in this particular study.

A description of each of these four stages comprises Chapters 3, 4, 5 and 6 respectively of this report.

3 Matching O*NET SOC to UK SOC

Chapter Summary

This Chapter first describes the occupational hierarchies in the O*NET system and in the UK SOC, and then explains how the mapping between the two classifications has been undertaken.

An appropriately modified version of the specialist software programme CASCOT has been used to provide a match between the job/occupational taxonomy in O*NET and that of the UK SOC. We report on the 'completeness' or coverage of the mapping, and also on the quality of the matching between the US and UK occupational classifications.

3.1 The UK and US Standard Occupational Classifications (SOC)

For the UK occupational classification, we have adopted the new SOC2010[14] occupational classification. This classification system has recently replaced SOC2000 (and, for example, will be used to classify occupations for the 2011 Census). As such, it is not yet the standard classification system implemented in the CASCOT software that we are using which still incorporates SOC2000, and so we have obtained a 'pre-release' beta version of the new UK SOC2010 classification structure through Professor Peter Elias at IER. Note however that we also still have the ability to match SOC2000 to examine how skill utilisation has developed historically and, by combining with SOC2000-based employment projections, how they may develop into the future.[15]

For the US, O*NET is currently classified according to a modified (i.e. slightly extended) version of US-SOC2009. This is referred to as the O*NET-SOC2009 taxonomy in what follows. Further details on each occupational classification system are presented in Annex B and Annex C for UK-SOC2010 and US O*NET-SOC2009 respectively.

There are some important differences between the US and UK occupational classification systems and these are worth briefly noting here since they impact upon the matching between the two classification systems.

- Neither classification system is consistent with the revised International Labour Office (ILO) International Classification of Occupations (ISCO) 2008. Hence there is no direct correspondence between the UK SOC2010 and the US SOC classification.

[14] http://www.ons.gov.uk/about-statistics/classifications/current/soc2010/index.html
[15] One potential use of the occupational profiles is in the new *Working Futures IV* projections which will be based on SOC2010, and so this provided anther motivation for using this new classification.

- Both classification systems are skill-based, hierarchical and with 4 distinct levels of aggregation:
- UK SOC2010 has nine major groups, 25 sub-major groups, 90 minor groups and 369 unit groups. All UK-SOC2010 occupations are assigned a 4 digit codes. 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.
- US SOC2009 has 23 major groups, 96 minor groups, 449 broad occupations and 821 detailed occupations. This information is recorded in a 6 digit code. The first and second digits represent the major group; the third digit represents the minor group; the fourth and fifth digits represent the broad occupation; and the sixth digit represents the detailed occupation.
- The O*NET-SOC2009 taxonomy (which is the current version used in O*NET) is an extended version of the US-SOC2009 classification. While O*NET does not gather detailed information for all 821 occupations in US-SOC2009, the additional detailed occupations defined O*NET-SOC2009 results in a total of 1,102 O*NET-SOC titles, of which information is separately recorded on 965 (so-called data-level) occupations in the O*NET system.[16]
- Finally, it is important to note that the primary purpose of an occupational classification is to classify a 'job'. A job is defined as a set of tasks or duties to be carried out by one person and represents a basic element in the employment relationship. Jobs are recognised primarily by their associated job title. There are 27,739 entries in the job title index for UK-SOC2010, while there are 56,636 entries in the job title index for O*NET-SOC2009.

3.2 CASCOT

We use CASCOT to produce the mapping between UK-SOC2010 and O*NET-SOC2009. CASCOT operates by matching input text to be coded against an index of words ('dictionary') to which the relevant codes have been allocated. The codes represent the classification, and for our purposes, this is an occupational classification (UK-SOC2010 or O*NET-SOC2009 as appropriate). A classification is usually described via a structure, an index and a set of rules.

The index is a collection of text descriptions, each associated with a specific category within a classification. The index may be comprehensive: containing the entire range of all possible valid pieces of text. However typically it is representative containing a large number of example pieces of text for each category but not limiting any category only to those entries. Rules provide operations to treat specific text in particular ways – such as

[16] O*NET-SOC2009 is an extended version of US-SOC2009 providing for more detailed occupations frequently reflecting new and emerging occupations – see Annex C for details.

resolving common abbreviations, downgrading certain words (eg man, woman), using alternative or equivalent words etc, and conclusions to make when encountering particular text. The set of rules to be associated with any classification are accumulated with experience, and can be added to over time. Currently, for example, there are over 2,000 rules associated with the SOC2010 classification.

Having assessed the quality of the match between the input text and the index of words, CASCOT produces a score on a 0 to 100 scale. The scoring mechanism as implemented in CASCOT is described in Box 1:

Box 1: Scoring in CASCOT

Cascot is designed to assign a code to a piece of text. Ideally the input text should contain sufficient information to distinguish it from alternative text descriptions which may be coded to other categories within the classification and should not contain superfluous words. This ideal will not always be met. Cascot has been designed to perform a complicated analysis of the words in the text, comparing them to the words in the classification, in order to provide a list of recommendations. When compiling this list of recommendations Cascot also calculates a score from 0 to 100 which approximates the probability that the code recommended for a specific piece of input text is correct.

Frequently the input text may be a word or phrases that is descriptive of an occupation or industry but lacks sufficient information to distinguish it from other categories (i.e. without any further qualifying terms). For example in SOC2000 the text ‘Teacher’ cannot be coded unambiguously to a single category because the word occurs in several categories (including: 2315 Primary and nursery education teaching professionals; 2314 Secondary education teaching professionals; 2312 Further education teaching professionals; 2311 Higher education teaching professionals etc). When this situation is encountered there may be a rule which defines a default category, e.g. in SOC2000 the code 8212 is recommended for the unqualified text ‘Driver’. If there is no default category Cascot will still list recommendations but the score is limited to below 40 to indicate the uncertainty associated with the suggestion. For example in SOC2000 ‘Teacher’ or ‘Engineer’ have no default category.

If the input text is not sufficiently distinctive it may be the case that the top recommendation, the one with the highest score, is not the most appropriate code. This is likely to occur when there are two or more closely competing categories to which a text description could be coded.

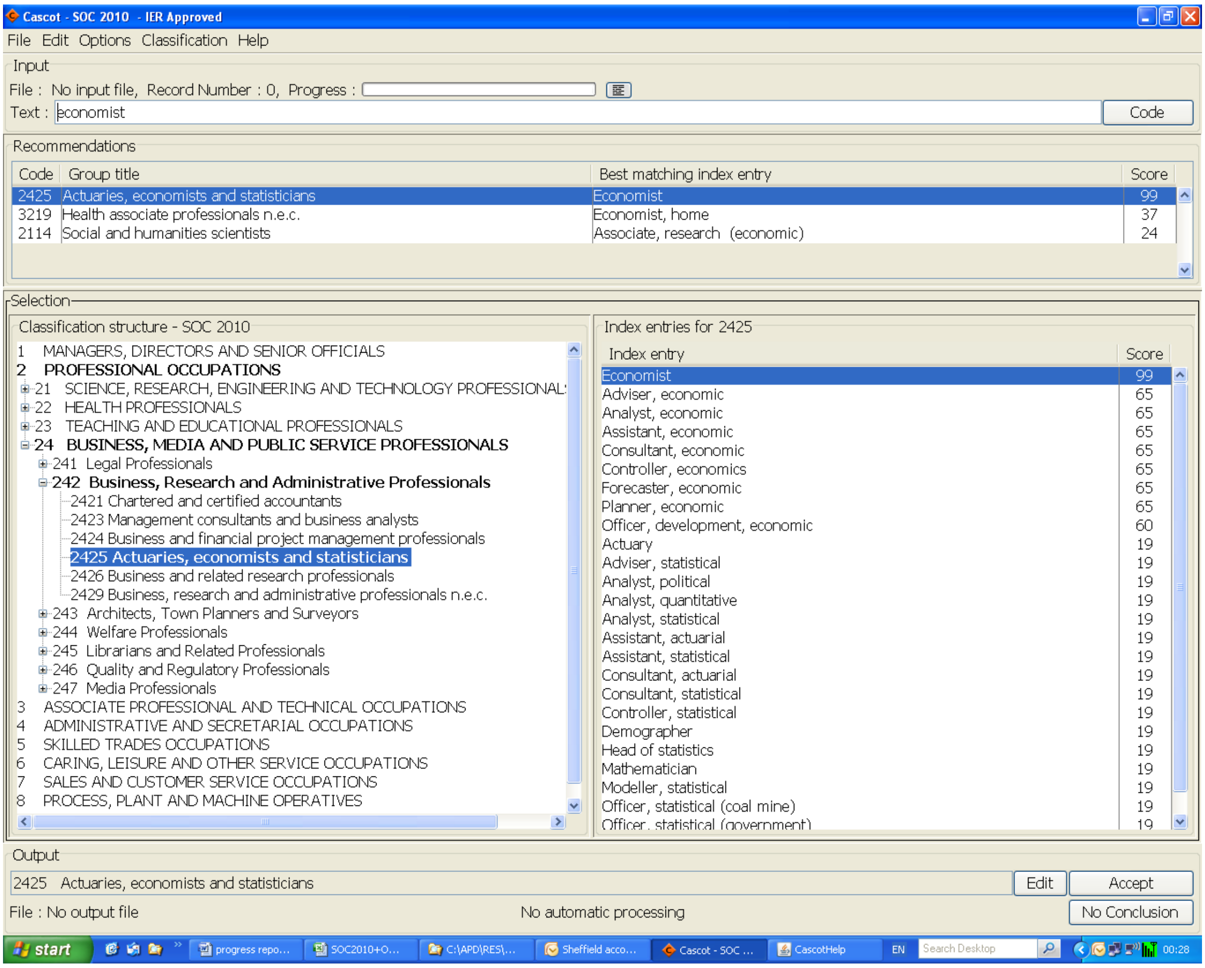
Sometimes Cascot will not list any recommendations at all. This may be because of a rule defining the input text to be unclassifiable (eg ‘Mother’ for SOC2000), or because the input text bears little relationship to any other term in the index or the classification.

Source: *CASCOT help documentation*

As an example, using the UK SOC2010 classification, Figure 2 displays the CASCOT output in response to the term ‘Economist’ as the text input in the top left hand corner. It returns a score of 99 for ‘SOC 2425: Actuaries, economists and statisticians’, and this is its top recommendation since it is the ‘best match’. The other index entries for SOC 2425 are listed in the lower right hand panel together with their scores for ‘economist’.

However, there are other possibilities too and these are listed with lower scores under the entry for SOC 2425 – these are SOC 3219 and SOC 2114 corresponding to index entries 'home economist' and 'research associate (economics)' respectively. These additional choices enable an experienced operator to select accordingly if the top recommendation does not look appropriate.

Figure 2: CASCOT Classification of 'Economist' to SOC2010

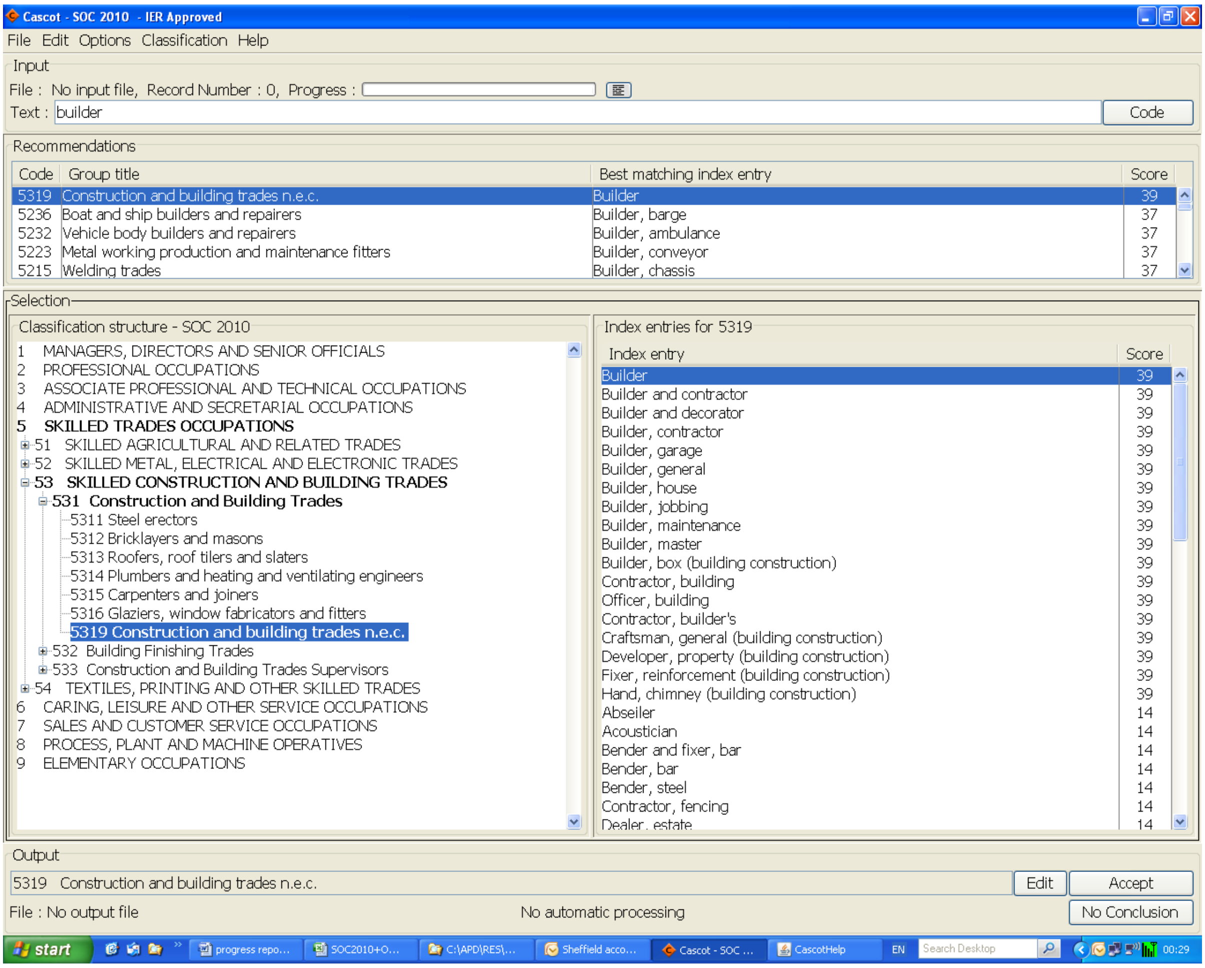


Source: *CASCOT, SOC2010 classification.*

Unfortunately, not all input text can be classified so unambiguously. Figure 3 reports the CASCOT response to the entry 'builder', This is clearly an ambiguous term – the term 'builder' is indexed under a number of categories, and this is recognised in the score that CASCOT allocates. While 'SOC 5319: Construction and building trades n.e.c.' is chosen as the top category, it is only given a score of 39 reflecting its ambiguity i.e. the fact that builders can build many different things as shown in the recommendation list. Ambiguous text is limited to a score of 39 in CASCOT – essentially this is regarded as a 'fail'.

This 'best match' classification system can be automated – i.e. multiple text entries can be supplied to CASCOT and processed in batch – and this is the approach that we have implemented.

Figure 3: CASCOT Classification of 'Builder' to SOC2010



Source: *CASCOT, SOC2010 classification.*

3.3 Matching

Given suitable dictionaries for UK SOC2010 and for O*NET-SOC2009, matching can, in principle, be either from UK SOC into O*NET SOC or vice versa. The advantage of matching from UK SOC into O*NET SOC is that it will ensure that all UK SOC occupations get matched (at least with some score) to O*NET occupations, since the UK SOC would be used as the input text. Coding from O*NET SOC into UK SOC may result in some UK SOC codes not being matched at all, and our ultimate purpose is to be able to match all of the 369 unit groups (4-digit occupations) in UK SOC in order to be able to construct occupational skills profiles for all 4 digit occupations in SOC2010.

The major disadvantage of matching from UK SOC to O*NET SOC is that the rules that exist to help improve the matching apply for coding *into* UK SOC rather than into any other occupational classification, and these rules will not necessarily apply well when coding into O*NET SOC[17]. However, even in the absence of a set of rules for O*NET coding, matching in this 'reverse' way can be used to provide some information relevant for validation and robustness checks on the coding from O*NET SOC into UK SOC.

CASCOT already has a SOC2010 classification dictionary available. This has 27,739 index entries (job titles) associated with the 369 4-digit unit groups in SOC2010, an average of 75 job titles per occupational unit group. However, in order to perform 'reverse' matching from UK SOC *into* O*NET SOC, we have had to construct our own O*NET classification dictionary. O*NET provides a 'Lay Title File' (which was revised in August 2010 to be compatible with Version 15.0 of O*NET-SOC2009). This has 56,634 lay titles (job titles) associated with the 1,102 O*NET-SOC2009 occupational titles, an average of 51 per occupational group. The CASCOT editor was then used to construct the O*NET classification dictionary from the job titles and their associated codes which together describe the O*NET structure. It is then possible to code into O*NET using this as the classification dictionary.

We have undertaken 4 variants of matching:

Variant 1: Matched 1,102 O*NET-SOC2009 occupational titles into SOC2010 using the SOC2010 classification dictionary.

Variant 2: Matched 369 SOC2010 unit group titles into O*NET SOC using the O*NET classification dictionary.

Variant 3: Matched 56,634 O*NET-SOC2009 job titles into SOC2010 using the SOC2010 classification dictionary.

[17] Of course, a set of rules for O*NET classification could be developed, but this would be a lengthy task – the rule set for SOC2010 contains over 2,000 rules as noted above.

Variant 4: Matched 27,739 SOC2010 job titles into O*NET SOC using the O*NET classification dictionary.

One potential advantage of Variants 1 and 2 is that, given the relatively small number of text entries to be classified, it is possible to perform this matching 'by hand' – that is, if appropriate, to intervene in the matching process to select other than the best match SOC2010 unit group or O*NET occupational title that CASCOT selects. An experienced CASCOT user/operator can process this number of entries relatively quickly (hours rather than days). However, these variants produced quite a large number of anomalies where CASCOT did not appear to find a good match between O*NET SOC and UK SOC2010, and also between UK SOC2010 and O*NET SOC.

One possible cause of the problems in identifying a good match between the two classification systems using occupational/unit group titles is that CASCOT is specifically designed to classify *jobs*, but the names given to the 369 SOC2010 unit group titles (and also the 1,102 O*NET-SOC2009 occupational titles) are not jobs *per se*. Rather, they are labels provided by the those responsible for designing the occupational classification system to be generic descriptions of the group of jobs that the particular occupation code encompasses. It is therefore not really appropriate to attempt to match these to index entries since they are not actually jobs, but rather are simply words and phrases used to describe a collection of job titles. For this reason, we have subsequently disregarded these first two forms of matching in what follows. However, it will be possible to use Variants 1 and 2 as part of the validation and robustness assessment at a later stage of the project.

Our main focus will therefore be on Variants 3 and 4 above. In particular, Variant 3 benefits from the rules incorporated in the SOC2010 classification dictionary to classify the O*NET job titles since the rule set has been developed explicitly for matching into SOC2010. Note that it is not practical to intervene in this matching exercise given its scale – almost 60,000 O*NET job titles are classified – and hence some post-matching processing of the resulting information is required. It is also important to examine whether all 369 SOC2010 unit groups are matched with at least some O*NET-SOC2009 job titles since our aim is to provide a complete set of SOC2010 occupational skills profiles.

3.4 Outcomes

Some summary statistics for the outcomes for each of the four variants of matching described above are presented in Table 4.

Table 4: Matching Variants – Coverage and Summary Score Statistics

Input text format:	**OCCUPATIONAL TITLES**		**JOB TITLES**	
Matching variant:	**Variant 1**	**Variant 2**	**Variant 3**	**Variant 4**
Matching direction: source → target:	O*NET SOC →UK SOC	UK SOC → O*NET SOC	O*NET SOC →UK SOC	UK SOC → O*NET SOC
Number of input entries:	1,102	369	56,634	27,739
Score percentiles: *10%*	36	34	34	38
25%	44	40	40	42
(median) 50%	58	47	55	48
75%	74	56	72	59
90%	92	73	91	80
Mean score:	59.8	50.3	57.4	53.0
Number with score 100:	0	0	418	70
Number with score 0 (unmatched):	4	1	432	134
Allocation of source classification:	1,098 of 1,102	368 of 369	56,202 of 56,634	27,605 of 27,739
Coverage of target classification:	289 of 369	250 of 1,102	368 of 369	956 of 1,102
Missing of target classification:	80 of 369	852 of 1,102	1 of 369	146 of 1,102

Matching is, in general, better when matching into UK SOC than when matching into O*NET SOC as expected given the rules that are designed to assist in matching text entries into SOC2010. Matching into SOC using O*NET occupational titles (Variant 1) or job titles (Variant 3) produces higher mean and median scores than when matching into O*NET using SOC occupational titles (Variant 2) or job titles (Variant 4). In all variants, most of the input text entries – whether occupational titles or jobs – are matched with at least some positive score. However, around 25% of the scores are at less than 40 – which means that the match is ambiguous according to the CASCOT scoring conventions. There are a few unclassified input entries (score 0 – unmatched) in all variants, but these are never a significant proportion. Some of these are a result of differences in spelling or language, while others are because there is no corresponding occupation or job. Some examples include:

- Variant 1: O*NET occupation title 'Anesthesiologists' is one of only 4 input entries classified as 'no conclusion' (score 0) in SOC2010, but using the UK spelling ('Anaetheologists'), it obtains a score of 27 for '2211: Medical practitioners'.
- Variant 3: O*NET job title 'Tire molder' is classified as 'no conclusion' (score 0) in SOC2010, but using the UK spelling ('Tyre moulder'), it obtains a score of 96 for '8115: Rubber process operatives'.
- Variant 4: words such as labourer and cabbie are unrecognised, as are some UK-specific qualifications and abbreviations used as job titles (eg SEN, GP, MP), and so are classified as no conclusion. Similarly, specific professional and senior grades in the civil service are not recognised within O*NET and hence receive very low matched scores.

It is apparent that extending the set of rules (or rather, developing a set of rules specifically for the O*NET classification) would eliminate many if not most of these remaining difficulties.

As can be seen from Table 4 above, in terms of the coverage of the target classification, Variant 1 which matches O*NET occupational titles into SOC results in only 289 of the 369 SOC2010 unit groups being matched with (at least one) O*NET occupation, a coverage ratio of 78%. This means that 80 SOC unit groups would be unmatched and would require some intervention to link to O*NET-SOC2009 if we used this variant of matching to generate the occupational skills profiles for the SOC unit groups. And while Variant 2 which matches SOC unit groups titles into O*NET only fails to find a match for 1 of the 369 unit groups when these are used as input, the matches with O*NET for the other 368 unit groups are with less than one quarter of the O*NET occupational groups – just (250/1102=) 23% of the O*NET occupational data would be utilised if we use this as the matching variant.

While Variants 3 and 4 have numerically more unmatched input entries (mainly due to the very country-specific nature of some jobs (eg snowmaker) or differences in language as noted above, using the considerably more numerous job titles as input text ensures much greater coverage of the target classification. In particular, in Variant 3, all but one of the 369 SOC2010 unit groups is matched to at least one O*NET job title with a positive score. The one unmatched unit group is '3115: Police community support officers' which is never selected as the best match to any of the 56,634 O*NET job titles. It would be relatively straightforward to assign O*NET occupational codes to this one remaining SOC unit group however. For the 368 matched unit groups, there is a mean (median) of 153 (83) O*NET jobs per SOC unit group, but the distribution is very uneven, and ranges from a minimum 1 job in a SOC unit group (there are two instances of this: O*NET job title 'architectural technologist' which is the only match to SOC2010 2435: chartered architectural technologist (score 94); and 'window cleaner', which is again the only match to 9231: Window Cleaners (score 94)), through to a maximum of 1,858 jobs in a unit group which is for SOC 8125: Metal working machine operatives (average score 49.5).

In what follows, we thus take the matching outcomes from Variant 3 as the basis of the mapping or correspondence between the US and UK occupational classifications. CASCOT provides us with a matrix of 56,634 rows (the US job titles) and 27,739 columns (the UK job titles). Each row has one and only one entry corresponding to the top score that CASCOT allocates to the quality of the match between the US job title and one of the UK job titles. The next Chapter of the report describes how this job-job matching information has been amalgamated to produce a weighting scheme between US and UK occupational titles.

4 Assigning Job Skills and Abilities to UK Occupations

Chapter Summary

The second stage of the project devises suitable schemes for amalgamating the CASCOT output on the US-job-title to UK-job-title matches to provide an O*NET-detailed-occupation to UK-SOC-occupational-unit-group correspondence. This is necessary since each occupation in O*NET SOC typically comprises several thousand jobs as does each occupation in UK SOC.

There are a number of different ways in which this can be accomplished and there is no optimal method. We illustrate a variety of the possible schemes using a simulated example, and then illustrate the resulting patterns in weights between O*NET and UK SOC occupations.

Having selected a weighting scheme between O*NET and UK SOC occupations, it is then possible to assign O*NET occupational information to the UK SOC unit group occupations.

4.1 Weighting Schemes

Having matched the 56,634 O*NET job titles to the 27,739 UK SOC2010 job titles using CASCOT, the next stage of the project is to use the resulting matching matrix to produce a weighting scheme that can be used to map the skills and other job and worker domains which are recorded at the O*NET detailed occupational level (for the data level occupations) into the SOC unit group occupations. That is, we need to amalgamate the 56,634 rows (US job titles) by 27,739 columns (UK job titles) in the job-job matching matrix into a matrix of 965 rows (O*NET-SOC2009 data-level occupations) and 369 columns (UK SOC2010 unit group occupations). There are a number of issues here that need to be considered:

- Should the job-job matching be ignored when scores are less than 40 (which is essentially regarded as a ‘fail’ in CASCOT – usually due to ambiguity)?
- Should weighting be based simply on the scores derived from CASCOT? Alternatively, should relative employment (in the US SOC category or as estimated in the O*NET-SOC category where this is not available) as an indicator of the importance of the occupational group also be taken into account?

- Should the job-job matching take account of non-matched jobs in averaging or aggregating to occupation-occupation weights (i.e. should these blanks be ignored or treat as zero scores)?
- Should weighting take into account the number of job-job matches in the occupation-occupation match rather than (or as well as) just the scores for the match?
- Should weighting take into account the aggregate scores across all job-job matches in the occupation-occupation match rather than average scores?

By way of illustration of these considerations, the next section implements a variety of aggregation methods for a simulated example, and graphically presents the resulting weighting schemes.

4.2 An Illustrative Example

This subsection presents an illustrative example of the matching process and alternative aggregation methodologies. For reference, the notation used is as follows:

O• = O*NET occupations (O1, O2, ...)

S• = UK SOC occupations (S1, S2, ...)

oj• = O*NET jobs (oj1, oj2, ...)

sj• = UK SOC jobs (sj1, sj2,...)

Figure 4 presents a simulated example which is for matching 7 O*NET jobs (oj1-oj7), which are in 3 O*NET occupational groups (O1-O3), to 5 SOC jobs (sj1-sj5) in 2 SOC2010 occupational groups (S1 and S2). This corresponds directly to Variant 3 in the CASCOT matching described in the previous Chapter (i.e. job title matching from O*NET SOC to UK SOC). The matching matrix of scores (SCORES) is shown in the matrix at the top of the figure. For each O*NET job oj, there will be *one and only one* entry in a row corresponding to the maximum score as generated by CASCOT with one and only one SOC job sj. (In practice, some of these maximum scores will be less than 40, but this is ignored in what follows). Some jobs may not match at all (as shown in Table 4, 432 of the 56,634 jobs in Variant 3 did not match to any SOC job in CASCOT).

Figure 4: Weighting Schemes Based on Scores

SCORES		S1	S1	S2	S2	S2
		sj1	sj2	sj3	sj4	sj5
O1	oj1	100				
O1	oj2	40				
O1	oj3				60	
O2	oj4					
O2	oj5		50			
O3	oj6			100		
O3	oj7	50				

	skill	employment
O1	7	50
O2	5	10
O3	1	40

SCHEME A **OCCUPATION AVERAGE SCORES (IGNORING BLANKS)**

SCORES	S1	S2
O1	70.0	60.0
O2	50.0	0.0
O3	50.0	100.0

SCORES ONLY

WEIGHT	S1	S2
O1	0.412	0.375
O2	0.294	0.000
O3	0.294	0.625
sum	1	1
skill	4.647	3.250

SCORES & EMPLOYMENT

WEIGHT	S1	S2
O1	0.583	0.429
O2	0.083	0.000
O3	0.333	0.571
sum	1	1
skill	4.833	3.571

SCHEME B **OCCUPATION AVERAGE SCORES (BLANKS AS ZEROS)**

SCORES	S1	S2
O1	23.3	6.7
O2	12.5	0.0
O3	12.5	16.7

SCORES ONLY

WEIGHT	S1	S2
O1	0.483	0.286
O2	0.259	0.000
O3	0.259	0.714
sum	1	1
skill	4.931	2.714

SCORES & EMPLOYMENT

WEIGHT	S1	S2
O1	0.651	0.333
O2	0.070	0.000
O3	0.279	0.667
sum	1	1
skill	5.186	3.000

SCHEME C **OCCUPATION TOTAL SCORES**

SCORES	S1	S2
O1	140.0	60.0
O2	50.0	0.0
O3	50.0	100.0

SCORES ONLY

WEIGHT	S1	S2
O1	0.583	0.375
O2	0.208	0.000
O3	0.208	0.625
sum	1	1
skill	5.333	3.250

SCORES & EMPLOYMENT

WEIGHT	S1	S2
O1	0.737	0.429
O2	0.053	0.000
O3	0.211	0.571
sum	1	1
skill	5.632	3.571

Reading down the oj's: oj1 matches with sj1 (score 100); oj2 also matches with sj1 (score 40); oj3 matches with sj4 (score 60); oj4 does not match with any of the sj1-sj5; oj5 matches with sj2 (score 50); oj6 matches with sj3 (score 100); oj7 matches with sj1 (score 50). Reading across the columns, all sj's except sj5 are matched with at least one oj.

Finally, each of O1-O3 has some matched jobs to at least one of S1 or S2 with the exception that no jobs in O2 match with any jobs in S2.

Corresponding to this SCORES matrix, Figure 5 counts the number of job-job matches which are shown as either 1 or blank in the COUNT matrix at the top of the Figure. Using counts rather scores reflects the number of matches in any occupation-occupation cell regardless of the 'quality' of the match as indicated by the CASCOT scoring algorithm.

Figure 5: Weighting Schemes Based on Counts

	COUNT	S1	S1	S2	S2	S2
		sj1	sj2	sj3	sj4	sj5
O1	oj1	1				
O1	oj2	1				
O1	oj3				1	
O2	oj4					
O2	oj5		1			
O3	oj6			1		
O3	oj7	1				

	skill	employment
O1	7	50
O2	5	10
O3	1	40

SCHEME A **OCCUPATION AVERAGE COUNTS**

COUNT	S1	S2
O1	1	1
O2	1	0
O3	1	1

COUNTS ONLY

WEIGHT	S1	S2
O1	0.333	0.500
O2	0.333	0.000
O3	0.333	0.500
sum	1	1
skill	4.333	4.000

COUNTS & EMPLOYMEN

WEIGHT	S1	S2
O1	0.500	0.556
O2	0.100	0.000
O3	0.400	0.444
sum	1	1
skill	4.400	4.333

SCHEME B **OCCUPATION AVERAGE COUNTS**

COUNT	S1	S2
O1	0.33	0.11
O2	0.25	0.00
O3	0.25	0.17

COUNTS ONLY

WEIGHT	S1	S2
O1	0.400	0.400
O2	0.300	0.000
O3	0.300	0.600
sum	1	1
skill	4.600	3.400

COUNTS & EMPLOYMEN

WEIGHT	S1	S2
O1	0.571	0.455
O2	0.086	0.000
O3	0.343	0.545
sum	1	1
skill	4.771	3.727

SCHEME C **OCCUPATION TOTAL COUNTS**

COUNT	S1	S2
O1	2	1
O2	1	0
O3	1	1

COUNTS ONLY

WEIGHT	S1	S2
O1	0.500	0.500
O2	0.250	0.000
O3	0.250	0.500
sum	1	1
skill	5.000	4.000

COUNTS & EMPLOYMEN

WEIGHT	S1	S2
O1	0.667	0.556
O2	0.067	0.000
O3	0.267	0.444
sum	1	1
skill	5.267	4.333

Below these matching matrices, we also provide some simulated data for each occupation O on employment and a measure of job skill. We consider a variety of weighting schemes between the Os and the Ss derived from the CASCOT matching scores/counts, possibly in combination with relative employment shares, in order to produce a weighted skill score for each occupation S. That is, we examine different methods of summarising the 7 × 5 matrix of job-job scores as would be produced by CASCOT into a 3 × 2 matrix of occupation-occupation weights. The weighting schemes differ in terms of how the job-job matching scores are combined.

- In Scheme A, averages across all identified (i.e. non-zero) matches in the occupation-occupation cell are taken (i.e. ignoring the blanks (non-matches)).
- In Scheme B, the non-matches are explicitly taken into account by averaging over all job-job cells (i.e. regarding blanks (non-matches) as scores of zero). In terms of relative weights, this is also equivalent to averaging over the row sums of oj within each S (since each oj has a maximum of one entry in any row).
- In Scheme C, the aggregate total score in each occupation-occupation cell is computed.

The results of implementing these three schemes are shown in the first column of occupation-occupation grids in Figure 4. For the count-based scoring in Figure 5, combining the counts in each scheme uses the same principles as for the scores. Thus for Scheme A, all occupation-occupation averages are either 1 or 0. Including zeroes as in Scheme B is equivalent to taking into account the number of oj that make a match within any S. Finally Scheme C counts the total number of matches in each occupation-occupation cell.

Having combined the scores and counts, the relative weights can be calculated. These are just the normalised (i.e. sum to unity) scores/counts for each occupation S (i.e. for each column). This enables any S to be expressed as a weighted average of all of the O's; the weight will be zero if there are no matches between an O*NET occupation and a SOC occupation as shown in the O2-S2 cells in the second column of occupation-occupation grids in Figures 4 and 5.

These weights do not take into account the relative importance of the Os in terms of their share of employment. Thus the final grid in each scheme combines the relative scores and counts with relative employment in each O, again normalising so that the weights sum to unity for each S. Thus O1 which has 5 times more employment than O2, now gets a greater weight when the score based weights are combined (multiplicatively) with relative employment. This is also true of the count based weights too. The results from combining the score/counts with employment are shown in the third column of occupation-occupation grids in Figures 4 and 5.

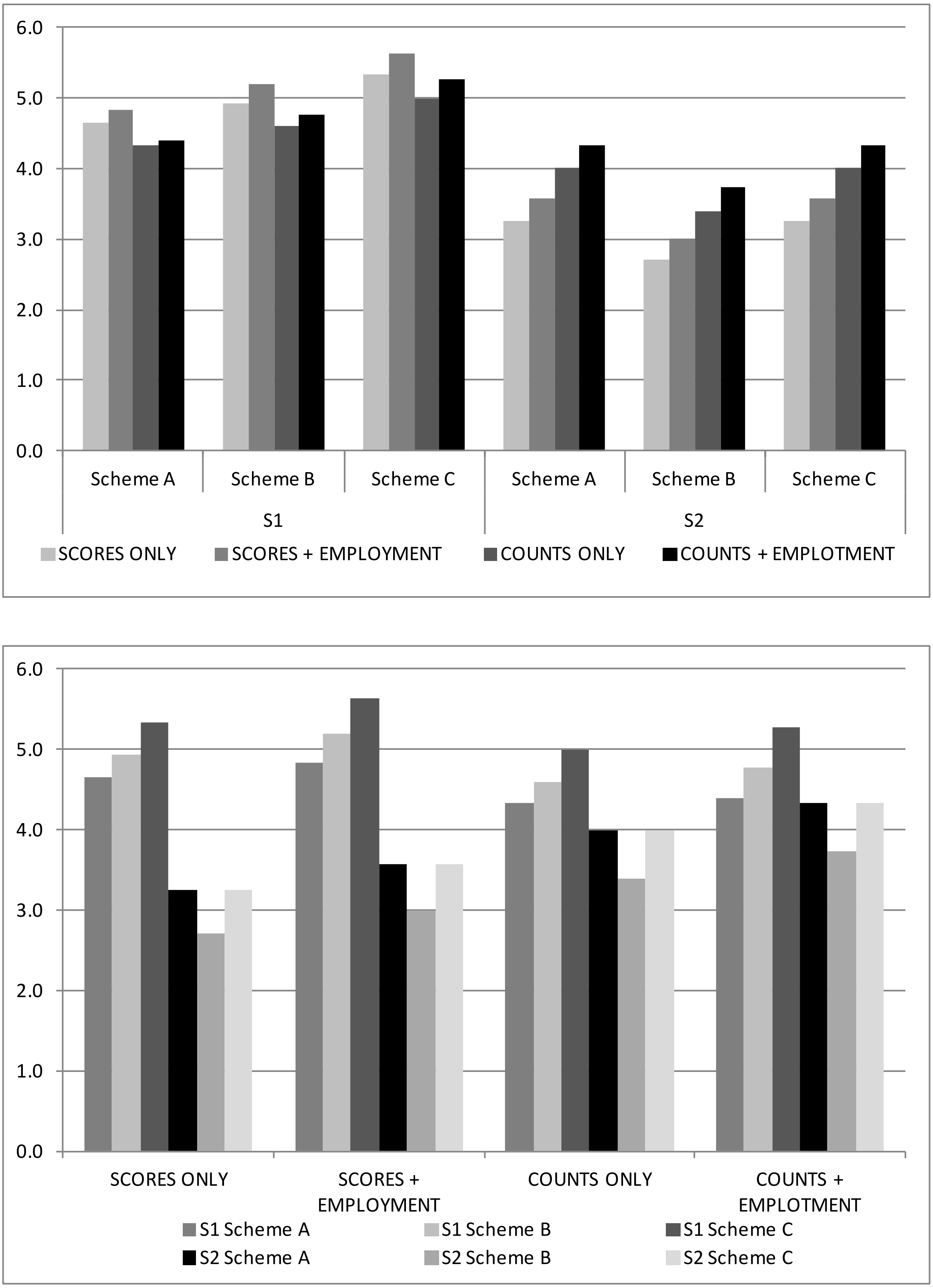
For each of the 12 weighting methods (3 (Schemes A, B, C) × 2 (without or with employment) × 2 (SCORES or COUNTS)), we illustrate the impact of the different methods using a measure of ‘skill’ at the O*NET level. The ‘skill’ row in Figures 4 and 5 beneath the weights is the weighted average skill for each S based on the weights in the grids immediately above. These weighted average skill scores are illustrated in Figure 6.

The top graph presents for the 4 methods of weighting (scores only, scores+employment, counts only, and counts+employment) the estimated skill score for S1 and S2 using Scheme A, B and C. Average Skill is consistently greater in S1 than in S2 which suggests that (at least in these limited data), the weighting does not changing the rankings, although the difference in skill scores vary quite a lot between Schemes, and between methods. Within method, the difference is greatest for Scheme B (averaging including blanks as zeroes) and smallest for Scheme A (averaging ignoring blanks). Indeed, for the final method (counts+employment) and Scheme A (ignoring the blanks) the average skill is actually the same in S1 and S2.

The bottom graph presents the same data organised rather differently. Now, for each S and Scheme, the average skill for the four different methods (scores and counts, without or with relative employment) are shown.

The alternative weighting schemes capture the extent to which matching is prevalent between pairs of occupations, the degree to which there is good ‘coverage’ of the SOC occupation, and the quality of the matching. Ultimately, no one method has a clear superiority over any other and we therefore perform sensitivity analyses to investigate the extent to which the choice of weighting makes significant differences to the results we obtain for the occupational skills profiles.

Figure 6: Comparison of Weighting Schemes



4.3 Comparison of Weights

In this section, we apply the 12 weighting methods described in the example above to the output from CASCOT and describe and compare the resulting sets of weights. For the schemes which take into account the relative importance of each O*NET occupation in terms of its employment, we use employment estimates for 2008 (the latest available) from the Bureau of Labour Statistics[18]. This is available for the 821 detailed occupations in the US SOC categories. Where O*NET-SOC provides greater detail than US SOC, we have no further information on employment in the subcategories (the so-called detailed O*NET-SOC occupations). For these, we simply distribute the employment of the parent SOC evenly across the detailed O*NET-SOC sub-categories. The results of the amalgamation of the 56,634 × 27,739 matrix of job-job scores from CASCOT into a 1,102 × 368[19] matrix of occupation-occupation weights are summarised below.

First, the mean (median) number O*NET occupations linked to each of the UK SOC occupation unit groups is 49.2 (34.5). The reveals the importance of the job-job matching that we have undertaken – rather than simply matching at the occupation level as in Variant 1 and Variant 2 as described in Chapter 3 (and as used by some other researchers when using information from the O*NET system). US-UK occupational groups do not match one-to-one but rather more 'fuzzily', and the weighting schemes that we have devised explicitly take this more complex matching between US and UK occupational groups into account.

Second, there are three occupational groups which match uniquely i.e. the UK SOC is matched with one and only one O*NET occupation. These are:

1. SOC 2435: Chartered Architectural Technologist which matches uniquely with O*NET 17-3011.01: Architectural Drafters. This matching arises from there being just one job ('Architectural Technologist') of the 35 jobs in 17-2011.01 which matches to 2435, although there are 15 other SOC codes which also link to 17-3011.01 through the other 34 jobs.
2. SOC 3216: Dispensing opticians which matches uniquely with O*NET 29-2081.00: Opticians, Dispensing. This matching arises from six jobs (Optical Dispenser, Dispensing Optician, Licensed Dispensing Optician, Dispensing Optician Apprentice, Dispensing and Measuring Optician, Licensed Optical Dispenser) of the 23 jobs in 29-2081.00 which match to 3216, although there are 7 other SOC codes which also link to 29-2081.00 through the other 17 jobs.
3. SOC 9231: Window Cleaners which matches uniquely with O*NET37-2011.00: Janitors and Cleaners, Except Maids and Housekeeping Cleaners. This matching

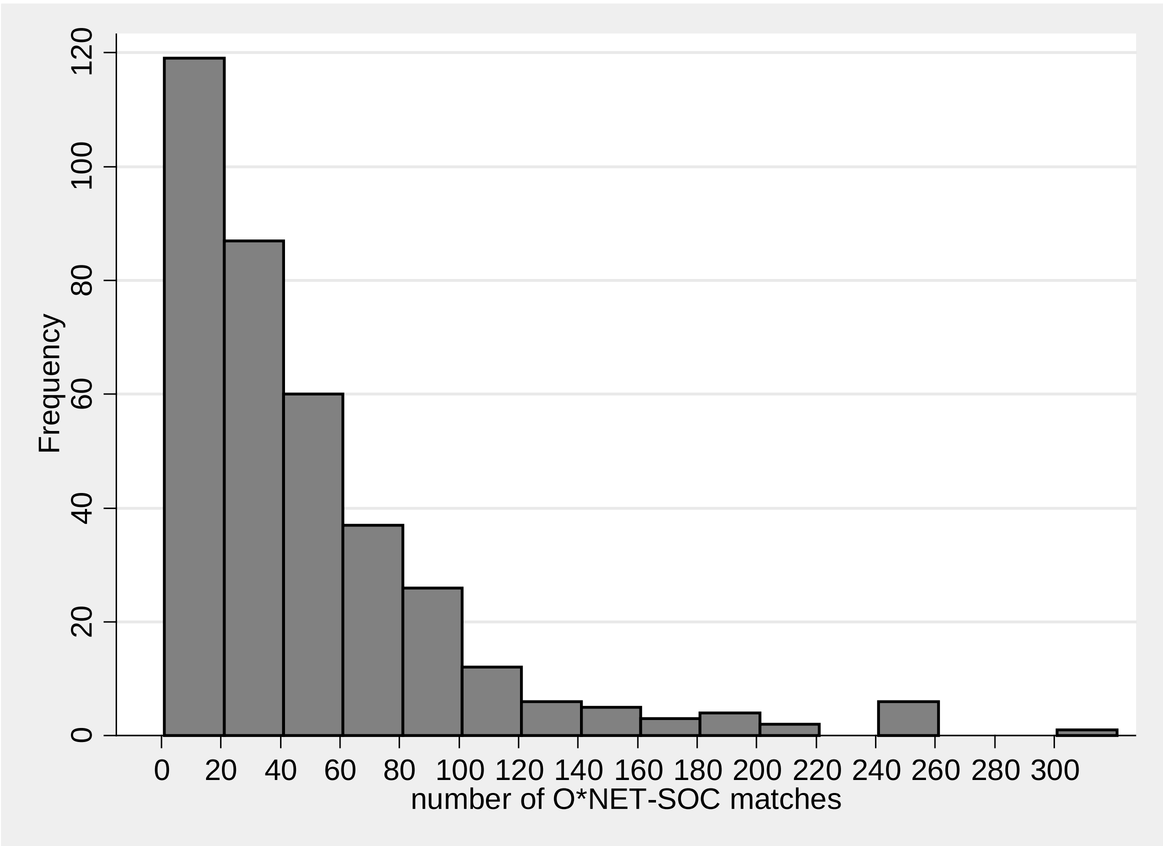
[18] Source: BLS
[19] At this stage, we have not provided a match to the one UK SOC occupation unit group that is not matched to any job in O*NET - '3115: Police community support officers'.

arises from there being just one job ('Window Cleaner') of the 129 jobs in 37-2011.00 which matches to 9231 although there are 41 other SOC codes which link to 37-2011.00 through the other 128 jobs.

At the other extreme, there is one UK SOC which matches with 307 different O*NET occupations. This is SOC 3119: Science, engineering and production technicians n.e.c and arises from 853 different O*NET jobs having a best match with this (very broad) SOC unit group.

The overall distribution of O*NET-SOC occupation matches is depicted in Figure 7 for the 368 SOC occupation unit groups: 119 SOC occupations (approximately 33 per cent of the total) have between 1 and 20 matches, while only 39 (10 per cent) have more than 100 matches. This distribution arguably reflects a reasonably close correspondence between the two occupational classifications – (386,338/405,536 =) 95 per cent the cells in the 1,102 × 368 occupational matching matrix are empty. Clearly the correspondence could be made closer still if we ignored low scores from CASCOT (those less than 40 for example), and improved the SOC2010 rule set to take account of differences in spelling etc between the US and the UK as noted in Chapter 3.

Figure 7: Distribution of O*NET-SOC Occupation Matches



Third, the correlations between the 12 different weighting methods described above as applied to the O*NET-SOC matching are presented in Table 5.

Table 5: Correlations between weights

Panel A: Weights based on scores

A1: All weights (i.e. including zeroes) N=405,536

		Scores only			Scores and employment		
Scheme		A	B	C	A	B	C
Scores only	A	1.0000					
	B	0.6598	1.0000				
	C	0.7140	0.8360	1.0000			
Scores & emp	A	0.6074	0.3045	0.4311	1.0000		
	B	0.5516	0.6737	0.7016	0.6600	1.0000	
	C	0.5371	0.4628	0.6855	0.7958	0.8435	1.0000

A2: Non-zero weights only N=18,094

		Scores only			Scores and employment		
Scheme		A	B	C	A	B	C
Scores only	A	1.0000					
	B	0.5928	1.0000				
	C	0.6497	0.8098	1.0000			
Scores & emp	A	0.5337	0.2066	0.3438	1.0000		
	B	0.4822	0.6346	0.6643	0.6207	1.0000	
	C	0.4676	0.3990	0.6475	0.7734	0.8278	1.0000

Panel B: Weights based on counts

B1: All weights (i.e. including zeroes) N=405,536

		Counts only			Counts and employment		
Scheme		A	B	C	A	B	C
Counts only	A	1.0000					
	B	0.6548	1.0000				
	C	0.7148	0.8261	1.0000			
Counts & emp	A	0.6037	0.3040	0.4352	1.0000		
	B	0.5424	0.6627	0.6884	0.6663	1.0000	
	C	0.5309	0.4504	0.6779	0.8030	0.8387	1.0000

B2: Non-zero weights only N=18,094

		Counts only			Counts and employment		
Scheme		A	B	C	A	B	C
Counts only	A	1.0000					
	B	0.5784	1.0000				
	C	0.6416	0.7950	1.0000			
Counts & emp	A	0.5268	0.1995	0.3422	1.0000		
	B	0.4671	0.6195	0.6468	0.6263	1.0000	
	C	0.4562	0.3804	0.6364	0.7805	0.8218	1.0000

Note: The 12 different weighting methods are described in Section 4.2.

Panel A presents the correlations between the weights based on scores, while Panel B presents the correlations between the weights based on counts. Correlations between the weights produced both without and with employment are shown. In Panel A1 and B1, all possible (1,102×368=) 405,536 weights are considered, including the 95% that are zero because there are no matches between the occupational groups. In Panel A2 and B2, these zero weights are ignored and the correlations presented for only the positive weights.

These correlations between the possible weighting methods considered are high in general, although it is difficult to discern any clear patterns. For example, score-based weighting schemes do not produce weights that are necessarily more highly correlated with each other than count-based schemes. However, it is the case that those methods which also incorporate employment shares to reflect the ‘importance’ of the O*NET occupation in employment are slightly more strongly correlated with each other than those which do not. Clearly, just as in the simulation considered above, in practice these weights capture slightly different nuances in the measurement of the ‘quality’ of the matching between US and UK occupational groups, and thus give rise to some differences in the weights as a consequence. How important these differences are can only really be assessed when the results of applying the weights to the job skills and abilities metrics in O*NET are examined.

5 Summarising Job Skills and Abilities for UK Occupations

Chapter Summary

The third stage of the project involves selecting a set of dimension/descriptors from the O*NET system. These are then combined using the weighting schemes devised in the previous Chapter to provide relevant and informative skills profiles for UK SOC occupational groups.

In order to inform our selection of descriptors from the O*NET system, we briefly review previous studies which have devised taxonomies based on the DOT and earlier versions of the O*NET database.

Focussing on just the skills and abilities sets of descriptors from O*NET, we present the resulting occupational skills profiles for two illustrative selections – one based on distinguishing the skills relevant for 'Data-People-Things', and one based on STEM (Science, Technology, Engineering and Mathematics) skills and abilities.

5.1 Selection of Descriptors from O*NET

Having matched the O*NET database to UK SOC2010, the objective of Stage 3 of the project is to produce some useful measures of skills used in different occupations for the UK. Given that there are 239 dimensions or descriptors of individual skills, job tasks and characteristics recorded in the O*NET system, clearly some selection and amalgamation/aggregation is required.

Table 6 documents some of the ways in which this has been accomplished previously using the DOT, O*NET, plus other surveys with similar characteristics to the O*NET surveys at the individual or occupational level. 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 Autor et al, 2003) is also popular.

Table 6: Summarising Skills, Tasks and Work Activities: Examples from the Literature

Reference	Taxonomy	Data	Measures/Methods	Notes/Findings
Autor, Levy and Murnane (ALM) (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 (mimeo 2009)	Cognitive tasks Interpersonal tasks Physical job tasks (aka data- people-things as used in DOT)	Princeton Data Improvement Initiative (PDII) O*NET v. (not specified) 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.
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.

Reference	Taxonomy	Data	Measures/Methods	Notes/Findings
Goos, Manning and Salomons (AER 2009 and CEP DP 1026 Nov 2010)	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.

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.

For our purposes here, the selection of descriptors with an appropriate aggregation method, together with a weighting scheme as devised in the previous Chapter enables us to characterise all UK unit group occupations in terms of a set of skills/job tasks i.e. to create a set of occupational skills profiles.

5.2 Creating a Taxonomy of Job Skills for the UK

In the following two sub-sections, we present two illustrative examples. First, using a taxonomy closely related to 'Data-People-Things', we create a set of profiles based on an allocation of all of the 35 Skills descriptors across these three broad categories. Second, we create an occupational indicator of STEM skills using a selection of indicators taken from the 35 Skills descriptors and the 52 Abilities descriptors.

5.2.1 Data-People-Things

In this sub-section, we use the O*NET Skills Domain only which comprises 35 items from Domain 2A (Basic Skills) and Domain 2B (Cross-functional Skills) (see Table 2) to create three broad indicators of job skills corresponding to Data, People and Things. Recall that Version 15 of the O*NET Database is the first version for which these job skills are recorded exclusively by job analysts, rather than being the average score recorded by job incumbents. We use only the skills *Importance* items in the analysis presented below – the skills *Levels* items could also be used since both are recorded for the Skills domain , but these tend to be highly correlated with the importance items as Handel (2010) and others have noted previously.

Table 7 lists the 35 Skills descriptors and their allocation into three broad sets which provide indicators of 'Data' skills, 'People' skills and 'Things' skills[20]. We then take the first principal component of each set of skill descriptors (21 items for Data, 7 items for People and 7 items for Things) in order to generate three summary measures at the O*NET occupational group level. The first principal component contains 55%, 71% and 77% of the variance of the set of items for Data, People and Things respectively. Thus, for each set of items, the first principal component would appear to be a good summary measure of the constituent set of skill descriptors.

[20] Note that, technically, it is not necessary to utilise all of the 35 descriptors as here – a selection could be used.

Table 7: Classification of O*NET Skills Domain Descriptors into Data-People-Things

Row	Code	Descriptor	Allocation
	2.A	**Basic Skills**	
	2.A.1	**Content**	
1	2.A.1.a	Reading Comprehension	Data
2	2.A.1.b	Active Listening	People
3	2.A.1.c	Writing	Data
4	2.A.1.d	Speaking	People
5	2.A.1.e	Mathematics	Data
6	2.A.1.f	Science	Data
	2.A.2	**Process**	
7	2.A.2.a	Critical Thinking	Data
8	2.A.2.b	Active Learning	Data
9	2.A.2.c	Learning Strategies	Data
10	2.A.2.d	Monitoring	Data
	2.B	**Cross-Functional Skills**	
	2.B.1	**Social Skills**	
11	2.B.1.a	Social Perceptiveness	People
12	2.B.1.b	Coordination	Data
13	2.B.1.c	Persuasion	People
14	2.B.1.d	Negotiation	Data
15	2.B.1.e	Instructing	People
16	2.B.1.f	Service Orientation	People
	2.B.2	**Complex Problem Solving Skills**	
17	2.B.2.i	Complex Problem Solving	Data
	2.B.3	**Technical Skills**	
18	2.B.3.a	Operations Analysis	Data
19	2.B.3.b	Technology Design	Data
20	2.B.3.c	Equipment Selection	Things
21	2.B.3.d	Installation	Things
22	2.B.3.e	Programming	Data
23	2.B.3.g	Operation Monitoring	Things
24	2.B.3.h	Operation and Control	Things
25	2.B.3.j	Equipment Maintenance	Things
26	2.B.3.k	Troubleshooting	Data
27	2.B.3.l	Repairing	Things
28	2.B.3.m	Quality Control Analysis	Things
	2.B.4	**Systems Skills**	
29	2.B.4.e	Judgment and Decision Making	Data
30	2.B.4.g	Systems Analysis	Data
31	2.B.4.h	Systems Evaluation	Data
	2.B.5	**Resource Management Skills**	
32	2.B.5.a	Time Management	Data
33	2.B.5.b	Management of Financial Resources	Data
34	2.B.5.c	Management of Material Resources	Data
35	2.B.5.d	Management of Personnel Resources	People

We then use the weights based on the CASCOT matching scores between O*NET-SOC2009 and SOC2010 occupational groups (using job-job matching as described in the previous Chapter) and occupational employment (i.e. we use weighting Scheme A, using scores and employment). This gives us measures of Data skills, People skills and Things skills for each of the 368 SOC2010 unit groups. Finally, we standardise the measures to have zero mean and unit variance for ease of comparison (there is more variance in the 'Data' measure in part because it is constructed from more dimensions). The 10 top and bottom ranked 4-digit occupations separately by Data, People and Skills are listed in Table 8.

The resulting measures for Data, People and Things are depicted in Figure 8. Higher level occupations (SOC2010 Major Groups 1, 2 and 3) have more intensive use of 'Data' and 'People' skills and less intensive use of 'Things' skills than lower level occupations (SOC2010 Major Groups 5-9). Similarly, SOC2010 Major Group 5: Skilled trades occupations are most intensive in their use of 'Things' but utilise lower 'Data' and 'People' skills. Finally, SOC2010 Major Groups 8 and 9 (Process, plant and machine operatives and Elementary occupations, respectively) exhibit the lowest relative use of 'Data' and 'People' skills.

These patterns in the relative usage of Data-People-Things skills are clearer still when we aggregate to the SOC2010 Sub-Major and Major group levels as shown in Figures 9 and 10 which average the Data, People and Things skills measures across the constituent SOC2010 unit groups.[21] The patterns in skills utilisation would appear to be consistent with our priors based on the Sub-Major and Major group descriptors (see Table B1).

The contribution that this project has made is to be able to provide this taxonomy at the 4-digit unit group level. Moreover, changes in the patterns of relative skills utilisation between (4 digit) occupations, and within Major (1 digit) and Sub-major (2 digit) groups can be traced over time by (i) recalculating the weights to allow for changing employment patterns and/or (ii) using previous versions of the O*NET database to allow for changes in the importance of different skills within O*NET occupational groups.

[21] Ideally, we would aggregate by taking a weighted average according to employment in each SOC2010 unit group. However, estimates of employment by SOC2010 unit group occupations is not yet available and so we have taken simple arithmetic means at this stage.

Table 8: Top and Bottom Ranked Occupations for Data-People-Things

Rank	DATA SOC	DATA Description	PEOPLE SOC	PEOPLE Description	THINGS SOC	THINGS Description
Top 10 Occupations						
Top 1	4124	Finance officers	1241	Health care practice managers	8135	Tyre, exhaust and windscreen fitters
2	1241	Health care practice managers	1181	Health services & public health managers	5237	Rail and rolling stock builders and repairers
3	2134	IT project and programme managers	1116	Elected officers and representatives	3116	Planning, process and production technicians
4	3234	Housing officers	3234	Housing officers	8141	Scaffolders, stagers and riggers
5	1116	Elected officers and representatives	4124	Finance officers	1252	Garage managers and proprietors
6	1181	Health services & public health managers	1133	Purchasing managers and directors	8143	Rail construction & maintenance operatives
7	1133	Purchasing managers and directors	2212	Psychologists	3113	Engineering technicians
8	2222	Occupational therapists	6146	Senior care workers	5119	Agricultural and fishing trades n
9	2141	Conservation professionals	2232	Midwives	3112	Electrical and electronics technicians
10	2212	Psychologists	1115	Chief executives and senior officials	5225	Air-conditioning and refrigeration engineers
Bottom 10 Occupations						
Bottom 1	9231	Window cleaners	9239	Elementary cleaning occupations n	2212	Psychologists
2	9239	Elementary cleaning occupations n	9231	Window cleaners	2114	Social and humanities scientists
3	9271	Hospital porters	9271	Hospital porters	2222	Occupational therapists
4	9236	Vehicle valeters and cleaners	9236	Vehicle valeters and cleaners	4114	Officers of non-governmental organisations
5	6231	Housekeepers and related occupations	5234	Vehicle paint technicians	2425	Actuaries, economists and statisticians
6	8223	Agricultural machinery drivers	8223	Agricultural machinery drivers	2311	Higher education teaching professionals
7	5114	Groundsmen and greenkeepers	6231	Housekeepers and related occupations	4124	Finance officers
8	9275	Leisure and theme park attendants	5114	Groundsmen and greenkeepers	4121	Credit controllers
9	9233	Cleaners and domestics	3116	Planning, process and production technicians	6126	Educational support assistants
10	9251	Shelf fillers	2433	Quantity surveyors	2312	Further education teaching professionals

Figure 8: Data-People-Things Taxonomy – 4-digit SOC2010

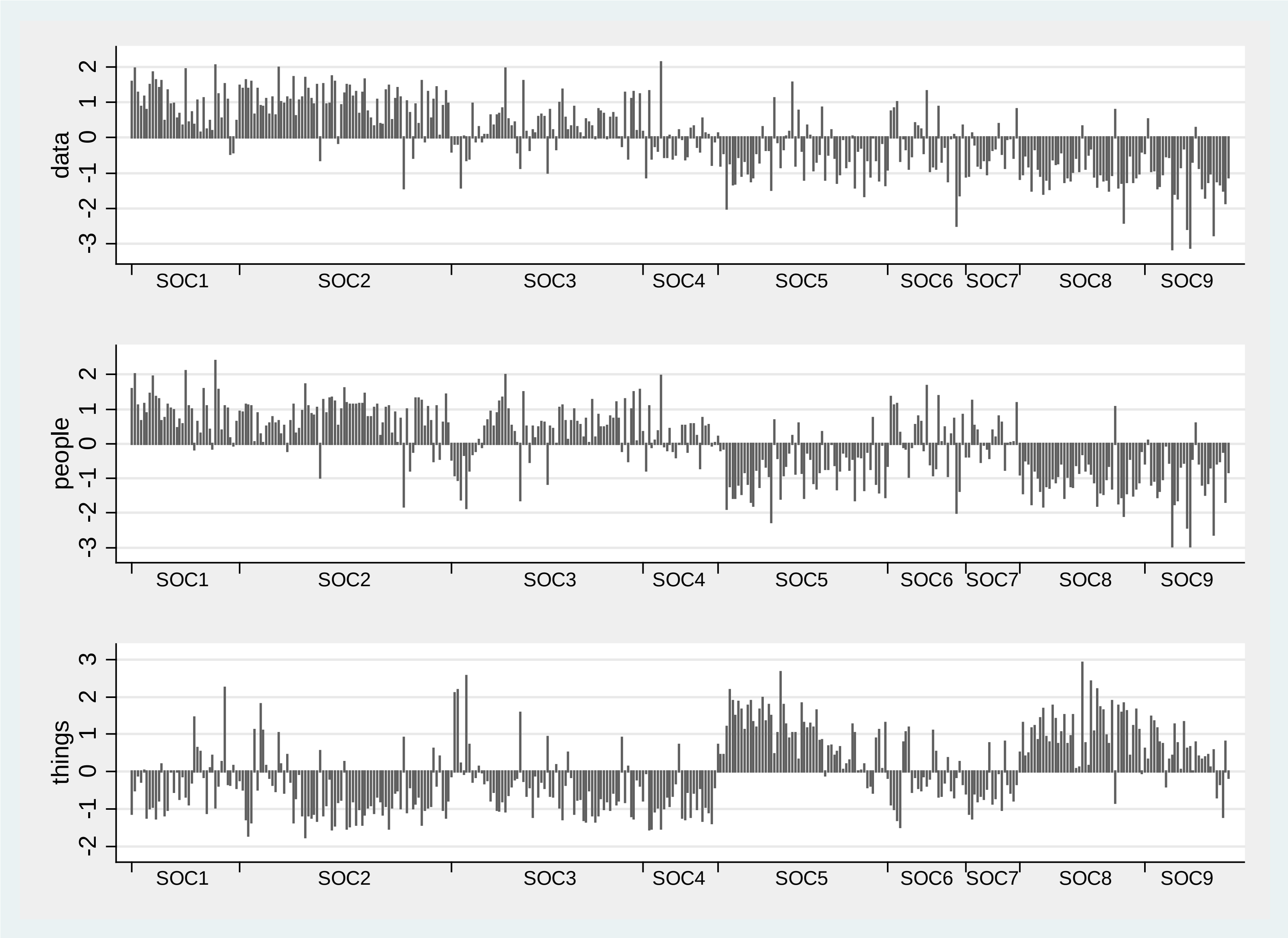


Figure 9: Averaging Data-People-Things Taxonomy to SOC2010 Sub-Major Groups

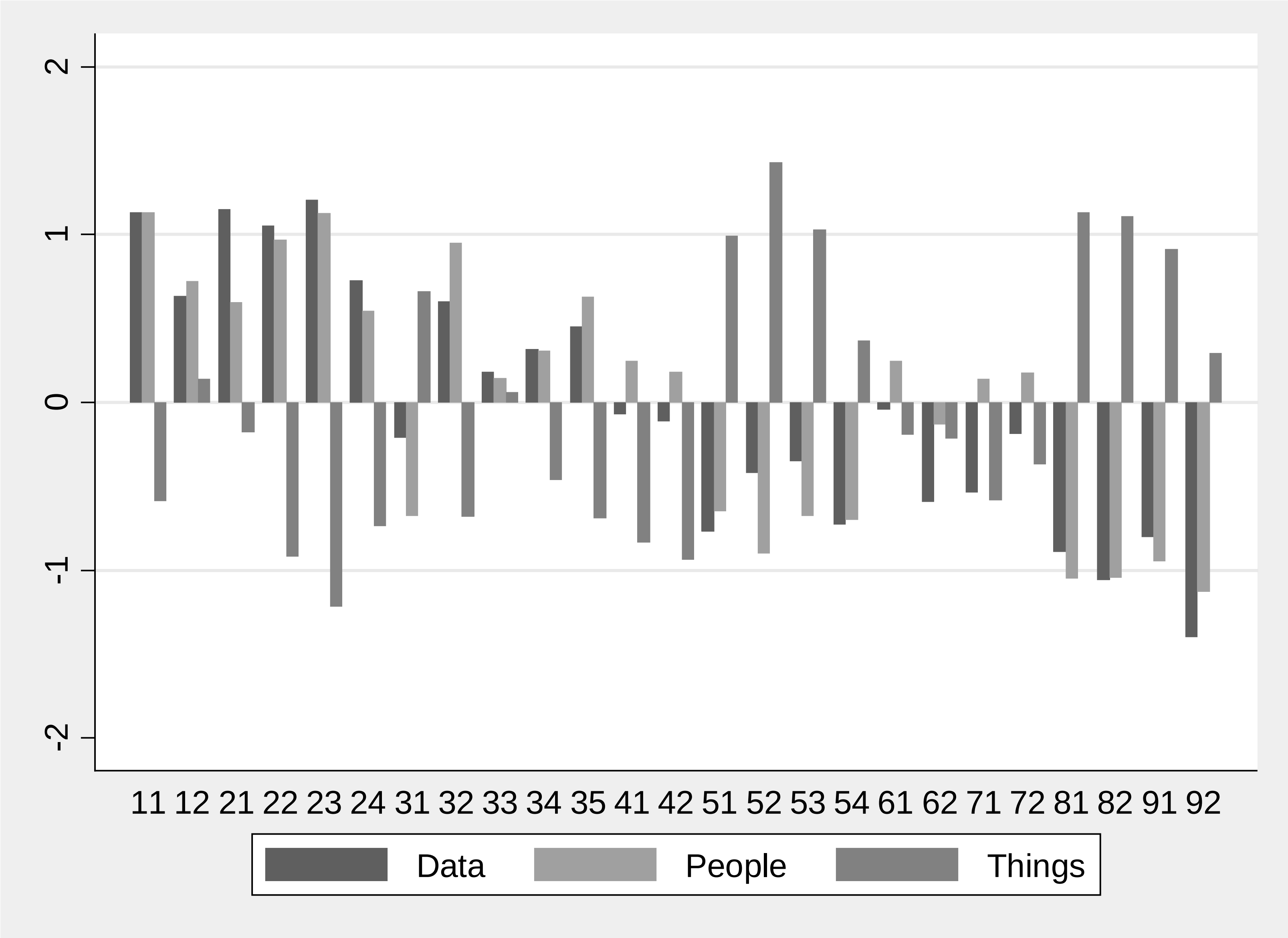
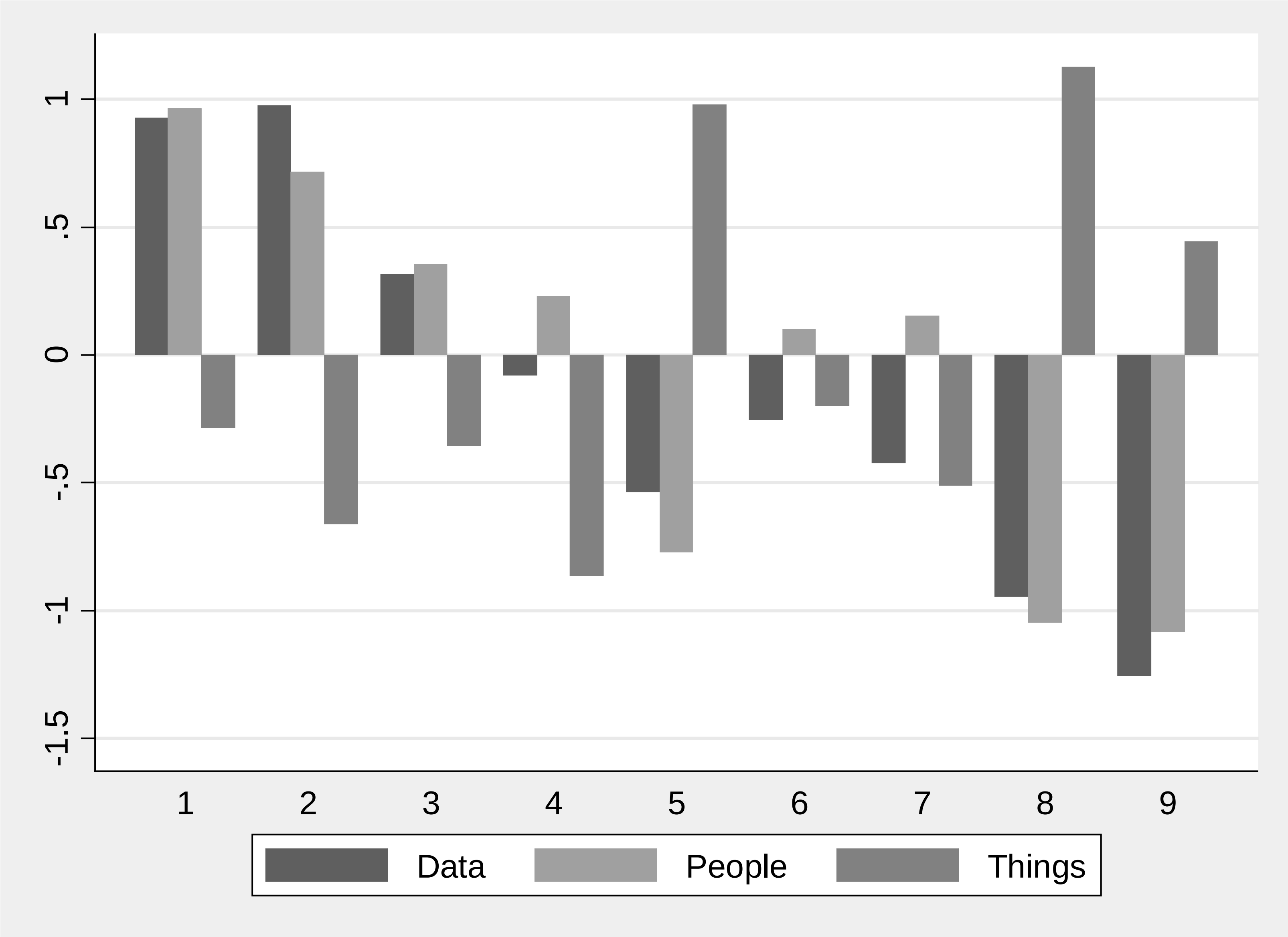


Figure 10: Averaging Data-People-Things Taxonomy to SOC2010 Major Groups



5.2.2 STEM Skills

Our second illustrative exercise takes relevant indicators from the Skills and Abilities domain to provide a STEM[22] skills occupational profile. We first identify the relevant descriptors from the Skills Domain as listed in Table 7 and the Abilities Domain listed in Table 9.

Table 9: O*NET Abilities Domain Descriptors

Row	Code	Descriptor
1	1.A.2.a.1	Arm-Hand Steadiness
2	1.A.4.b.2	Auditory Attention
3	1.A.1.b.7	Category Flexibility
4	1.A.2.b.1	Control Precision
5	1.A.1.b.4	Deductive Reasoning
6	1.A.4.a.6	Depth Perception
7	1.A.3.c.2	Dynamic Flexibility
8	1.A.3.a.3	Dynamic Strength
9	1.A.3.a.2	Explosive Strength
10	1.A.3.c.1	Extent Flexibility
11	1.A.4.a.2	Far Vision
12	1.A.2.a.3	Finger Dexterity
13	1.A.1.e.2	Flexibility of Closure
14	1.A.1.b.1	Fluency of Ideas
15	1.A.4.a.7	Glare Sensitivity
16	1.A.3.c.3	Gross Body Coordination
17	1.A.3.c.4	Gross Body Equilibrium
18	1.A.4.b.1	Hearing Sensitivity
19	1.A.1.b.5	Inductive Reasoning
20	1.A.1.b.6	Information Ordering
21	1.A.2.a.2	Manual Dexterity
22	1.A.1.c.1	Mathematical Reasoning
23	1.A.1.d.1	Memorization
24	1.A.2.b.2	Multilimb Coordination
25	1.A.4.a.1	Near Vision
26	1.A.4.a.4	Night Vision
27	1.A.1.c.2	Number Facility
28	1.A.1.a.1	Oral Comprehension
29	1.A.1.a.3	Oral Expression
30	1.A.1.b.2	Originality
31	1.A.1.e.3	Perceptual Speed
32	1.A.4.a.5	Peripheral Vision
33	1.A.1.b.3	Problem Sensitivity
34	1.A.2.b.4	Rate Control
35	1.A.2.c.1	Reaction Time
36	1.A.2.b.3	Response Orientation
37	1.A.1.g.1	Selective Attention
38	1.A.4.b.3	Sound Localization
39	1.A.1.f.1	Spatial Orientation
40	1.A.4.b.5	Speech Clarity
41	1.A.4.b.4	Speech Recognition
42	1.A.1.e.1	Speed of Closure
43	1.A.2.c.3	Speed of Limb Movement
44	1.A.3.b.1	Stamina
45	1.A.3.a.1	Static Strength

[22] STEM: science, technology, engineering and mathematics.

Row	Code	Descriptor
46	1.A.1.g.2	Time Sharing
47	1.A.3.a.4	Trunk Strength
48	1.A.4.a.3	Visual Color Discrimination
49	1.A.1.f.2	Visualization
50	1.A.2.c.2	Wrist-Finger Speed
51	1.A.1.a.2	Written Comprehension
52	1.A.1.a.4	Written Expression

The eight descriptors selected are as follows:

Skills Domain:		
Row	*Code*	*Descriptor*
5	2.A.1.e	Mathematics
6	2.A.1.f	Science
19	2.B.3.b	Technology Design
22	2.B.3.e	Programming
Abilities Domain:		
Row	*Code*	*Descriptor*
5	1.A.1.b.4	Deductive Reasoning
20	1.A.1.b.6	Information Ordering
22	1.A.1.c.1	Mathematical Reasoning
27	1.A.1.c.2	Number Facility

We focus again on the importance measures only given the high correlation between importance and level measures as previously noted (typically greater than 0.9 for the level and importance of the eight measures considered in this sub-section). We can aggregate these eight measures into a single index of STEM skills in two obvious ways. First, we consider simple averaging across the eight descriptors – this giving each the same weight in the STEM index (all eight measures are recorded on the same importance scale ranging from 1 (not important) to 5 (extremely important)). A second possibility is to take the first principal component of the eight measures. For the eight indicators, the first principal component summarises 60% of the variance, and the correlation between this component and the simple average index is 0.99. Hence there is little difference in the associated STEM skills index whichever aggregation method is selected.

We then weight the STEM index to SOC2010 as before using the weights based on average CASCOT matching scores between O*NET-SOC2009 and SOC2010 occupational groups based on job-job matching as described in the previous Chapter, and relative employment shares (i.e. weighting Scheme A, using scores and employment). Finally we standardise the resulting STEM measure and examine the relative distribution of STEM skills by 2, 3 and 4 digit SOC2010 occupation. Figure 11 presents the resulting profiles for unit groups, sub-major groups and major groups using the simple average of the eight STEM skill indicators.

Figure 11: Averaging STEM skills to SOC2010 Unit, Sub-Major and Major Groups

Figure 11A: Unit groups

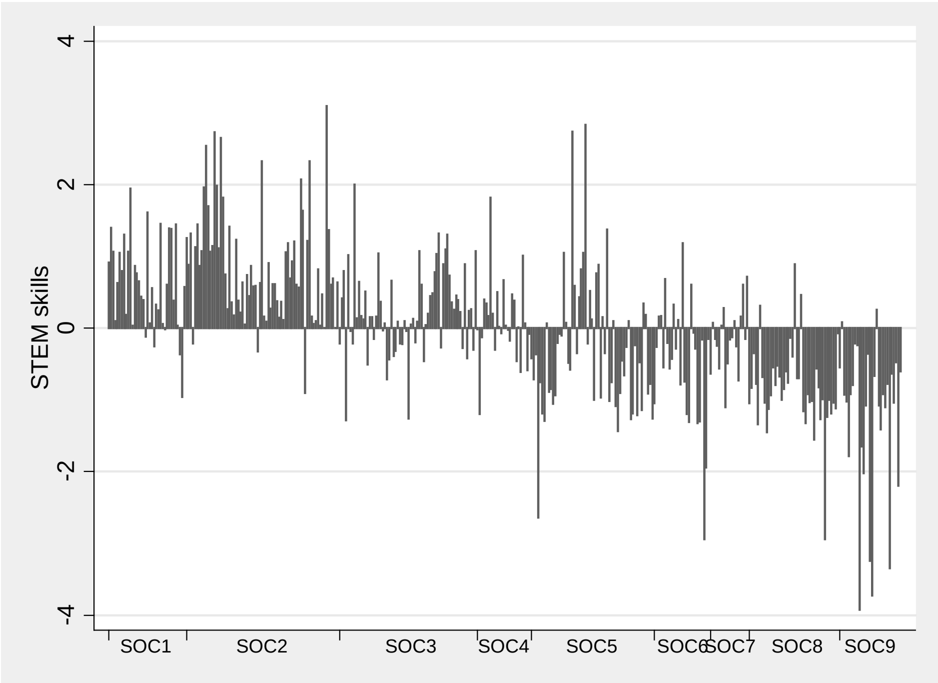


Figure 11B: Sub-Major groups

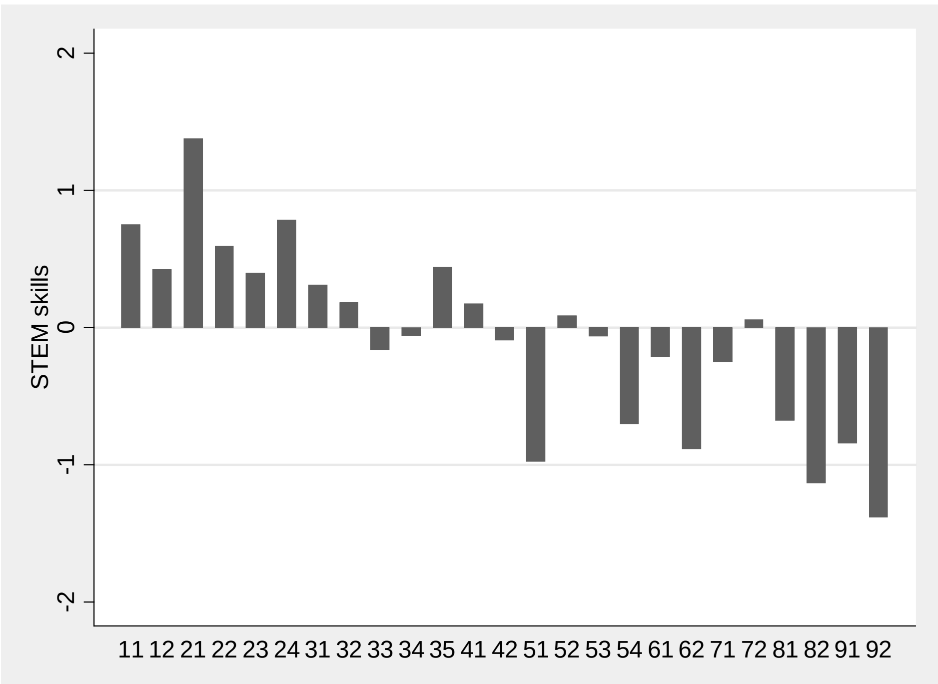
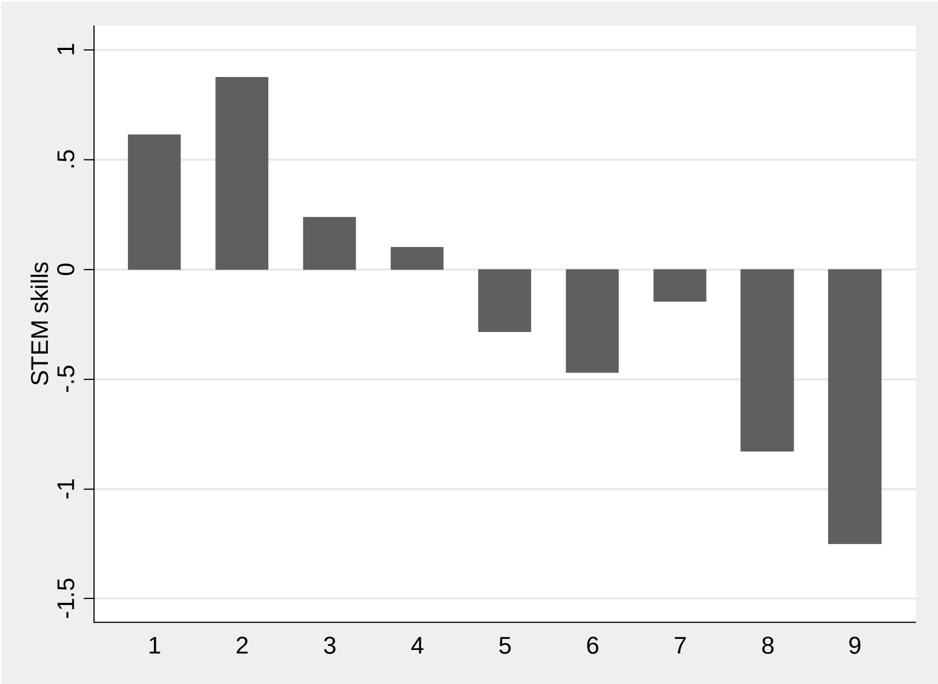


Figure 11C: Major groups



As would be anticipated, STEM skills are more prevalent in jobs located in professional occupations, especially Sub-major group 21: Science, Research, Engineering and Technology Professionals. In contrast, workers in Major groups 8 and 9 – operatives and elementary occupations – utilise very few STEM skills in their jobs.

5.3 Skills profiles

The two preceding examples serve to illustrate the way in which a selection of descriptors from the different domains in the O*NET database can be combined to provide occupational skills profiles for the UK. Any number of different profiles can be constructed using the set of domains/descriptors in the O*NET according to the particular demands and interests of the researcher. Single indices (i.e. based on only one descriptor), or combinations of descriptors can be used to define the skill set of interest.

There remains the question of how best to summarise and present the data on all 239 dimensions recorded in O*NET to potential users. One possibility is a 368×239 matrix with each column being one of the 239 dimensions/descriptors from O*NET weighted to the 368 SOC2010 unit groups.[23]

[23] Of course, there could be 12 of these matrices corresponding to each of the 12 different weighting schemes discussed above. But, as we show in the next Chapter, this is probably unnecessary.

6 Assessment and Validation

Chapter Summary

In this final stage of the project, we assess the validity of the methodology that we have developed and described in the earlier chapters.

We first compare some derived skills profiles for education, training and learning with independent estimates of these three measures as provided by the 2006 Skills Survey.

Second, we investigate the robustness of the resulting profiles to the choices made about weighting schemes and method of aggregation.

6.1 Assessment

Before making further or more widespread use of these new occupational skills profiles, it is important to attempt to assess the validity of the matching and merging methodology developed in this project. There are a number of possible ways in which this can be investigated. First, and most directly, we can compare the generated O*NET-based skills profiles for a particular measure of skills with the corresponding occupational profile derived from UK-based survey evidence. There are a number possibilities, but perhaps one of the most direct is to compare the derived SOC2010 occupational skills profiles for the various O*NET required levels of education, experience and training dimensions with the required qualification, training time and learning time indices derived from the 2006 UK Skills Survey (Felstead *et al*, 2007). Second, we can examine the sensitivity and robustness of the skills profiles to the various assumptions that are made with respect to weights, aggregation (e.g. averaging or principal component analysis (PCA)), etc. Third, we can examine past (and potentially future) trends in skills utilisation to see the extent to which the predicted trends as produced by the methodology accord with the actual and anticipated trends predicted by other forms of analysis. Note that it is not possible to examine past trends in skills utilisation under the SOC2010 classification as employed in this report since historic employment data is not available for SOC2010 occupations. Instead, we would need to undertake all of the CASCOT matching analysis again using a UK-SOC2000-based dictionary.[24]

The next section presents an examination of the occupational profiles for the required level of education, experience and training variables (in sub-section 6.2.1), and makes

[24] We could additionally use earlier O*NET databases together with previous US distributions of employment to produce skills and job task profiles for the past and match these to UK-SOC2000. Alternatively, we can simply assume that, at the fine level of SOC disaggregation we have utilised, change is mostly captured by the changing occupational distribution of employment, rather than between or within occupation changes in the nature of jobs.

direct comparisons with the corresponding UK Skills Survey measures. We then summarise the sensitivity of both the derived data-people-things taxonomy of skills and the STEM skills profile to the various choices of weights and aggregation method in sub-section 6.2.2.

6.2 Validation

6.2.1 Required Level of Education, Experience and Training Profiles

There are four descriptors within the O*NET content model which record required education, experience and training. All four descriptors are derived from responses to questions answered mainly by job incumbents. Within the Worker Requirements domain (see Table 2) , item 2.D.1 is the required level of education, defined on a 12 point scale as the level of education required to perform a job. Within the Experience Requirements domain (see Table 2), there are three descriptors of relevance: 3.A.1 records the amount of related work experience required to get hired for the job (on an 11 point scale); 3.A.2 measures the amount of on-site or in-plant training (such as organized classroom instruction) required to perform the job (on a 9 point scale); and 3.A.3 assesses the amount of on the job training required to perform the job (again, on a 9 point scale). The scales for each of these four descriptors are defined in Table 10.

Table 10: Definitions for O*NET Required Education, Experience and Training Scales

CODE and DESCRIPTOR Scale and definition	
2.D.1: REQUIRED LEVEL OF EDUCATION	
1	Less than a High School Diploma
2	High School Diploma (or GED or High School Equivalence Certificate)
3	Post-Secondary Certificate - awarded for training completed after high school (for example, in Personnel Services, Engineering-related Technologies, Vocational Home Economics, Construction Trades, Mechanics and Repairers, Precision Production Trades)
4	Some College Courses
5	Associate's Degree (or other 2-year degree)
6	Bachelor's Degree
7	Post-Baccalaureate Certificate - awarded for completion of an organized program of study; designed for people who have completed a Baccalaureate degree, but do not meet the requirements of academic degrees carrying the title of Master
8	Master's Degree
9	Post-Master's Certificate - awarded for completion of an organized program of study; designed for people who have completed a Master's degree, but do not meet the requirements of academic degrees at the doctoral level
10	First Professional Degree - awarded for completion of a program that: requires at least 2 years of college work before entrance into the program, includes a total of at least 6 academic years of work to complete, and provides all remaining academic requirements to begin practice in a profession
11	Doctoral Degree
12	Post-Doctoral Training
3.A.1: RELATED WORK EXPERIENCE	
1	None
2	Up to and including 1 month
3	Over 1 month, up to and including 3 months

CODE and DESCRIPTOR Scale and definition	
4	Over 3 months, up to and including 6 months
5	Over 6 months, up to and including 1 year
6	Over 1 year, up to and including 2 years
7	Over 2 years, up to and including 4 years
8	Over 4 years, up to and including 6 years
9	Over 6 years, up to and including 8 years
10	Over 8 years, up to and including 10 years
11	Over 10 years
3.A.2: ON-SITE OR IN-PLANT TRAINING	
1	None
2	Up to and including 1 month
3	Over 1 month, up to and including 3 months
4	Over 3 months, up to and including 6 months
5	Over 6 months, up to and including 1 year
6	Over 1 year, up to and including 2 years
7	Over 2 years, up to and including 4 years
8	Over 4 years, up to and including 10 years
9	Over 10 years
3.A.3: ON-THE-JOB TRAINING	
1	None or short demonstration
2	Anything beyond short demonstration, up to and including 1 month
3	Over 1 month, up to and including 3 months
4	Over 3 months, up to and including 6 months
5	Over 6 months, up to and including 1 year
6	Over 1 year, up to and including 2 years
7	Over 2 years, up to and including 4 years
8	Over 4 years, up to and including 10 years
9	Over 10 years

The responses to each item are recorded using percentages or proportions falling into each category. Given the strong hierarchical nature of the definitions in each case, we aggregate for each O*NET occupation using a simple weighted average of the category numeric levels, with weights given by the proportions in each category. These are then reweighted to SOC2010 using the weights based on the average CASCOT scores between O*NET-SOC2009 and SOC2010 occupational groups. Once again, we only report the results based on job-job matching and relative employment shares (i.e. weighting Scheme A, using scores and employment), although choice of weighting scheme made no substantive differences to the findings presented below. Finally, the scores were aggregated to the SOC2010 Major Group level.

The *2006 Skills Survey* is described in detail in Felstead et al (2007). It is a large representative sample survey of working individuals living in the UK aged 20-65 and was undertaken in 2006-07. Its aim was to gather information on the skills used in work through survey questions directed at the workers themselves. A total of 7,787 individuals in the UK were surveyed. Weights are provided to ensure that any analysis is representative of the UK employed labour force, and these weights are utilised as appropriate in all of the analysis which follows. Amongst the many measures of the skills that individuals use in their jobs, three indices of broad skills can be derived: the qualification level required on entry into jobs; the training time required to do the type of work carried out; and the leaning time needed to do the job well. Measures of all of these dimensions of skills are obtained directly from the survey respondents' answers to a range of questions in the *Skills Survey*.

The first measure of broad skills is the qualifications required to get the job (as perceived by the individual currently doing that job). Note that this differs from the qualifications that an individual may possess but not necessarily require in order to get the job. It also may not correspond to the qualifications used to screen applicants for the job if it was vacant. The respondents were asked: "If they were applying today, what qualifications, if any, would someone need to get the type of job you have now?" A range of qualifications were shown and these were subsequently converted to the five major NQF equivalents. The required qualifications index is a numeric measure based on scoring the NQF levels from level 0 (for no qualifications required) to level 4 (for NQF level 4+, equivalent to first degree or higher).

The second measure is based on a series of questions relating to the training time required for the particular job of work performed by the survey respondent. The amount of training time required is presumed to reflect the knowledge and skills demanded by the job. Specifically, respondents were asked: "Since completing full-time education, have you ever had, or are you currently undertaking, training for the type of work that you currently do?" If they answered yes, they were then asked: "How long, in total, did/will that training last?" Given the distribution of responses, a training time index was constructed corresponding to: 0 – no training for job; 1 – up to 1 month; 2 – 1 month up to 3 months; 3 – 3 months up to 6 months; 4 – 6 months up to 1 year; 5 – 1 year up to 2 years; and 6 – over 2 years training.

The third broad skill measure which captures the time required to learn to do the job well was constructed in a similar fashion. It is presumed that the amount of time it takes to learn to do the job well is an indicator of the level of skills required in the job, although it is possible that less able individuals might take a longer time to learn how to do a job well. Respondents were asked: "How long did it take for you, after you first started doing this type of job, to learn to do it well?" and if they suggested that they were still learning, the supplementary question asked:"How long do you think it will take?" As with the training time index, the learning time index is based on converting the respondents responses to a numeric scale similar to that used for the training time index above.

These three Skills Survey measures of broad skills were then aggregated to the SOC2000 Major Group level.

Figure 12: Comparing Required Education and Training Levels at SOC Major Group Level

Figure 12A: Required qualifications, training time and learning time – UK Skill Survey

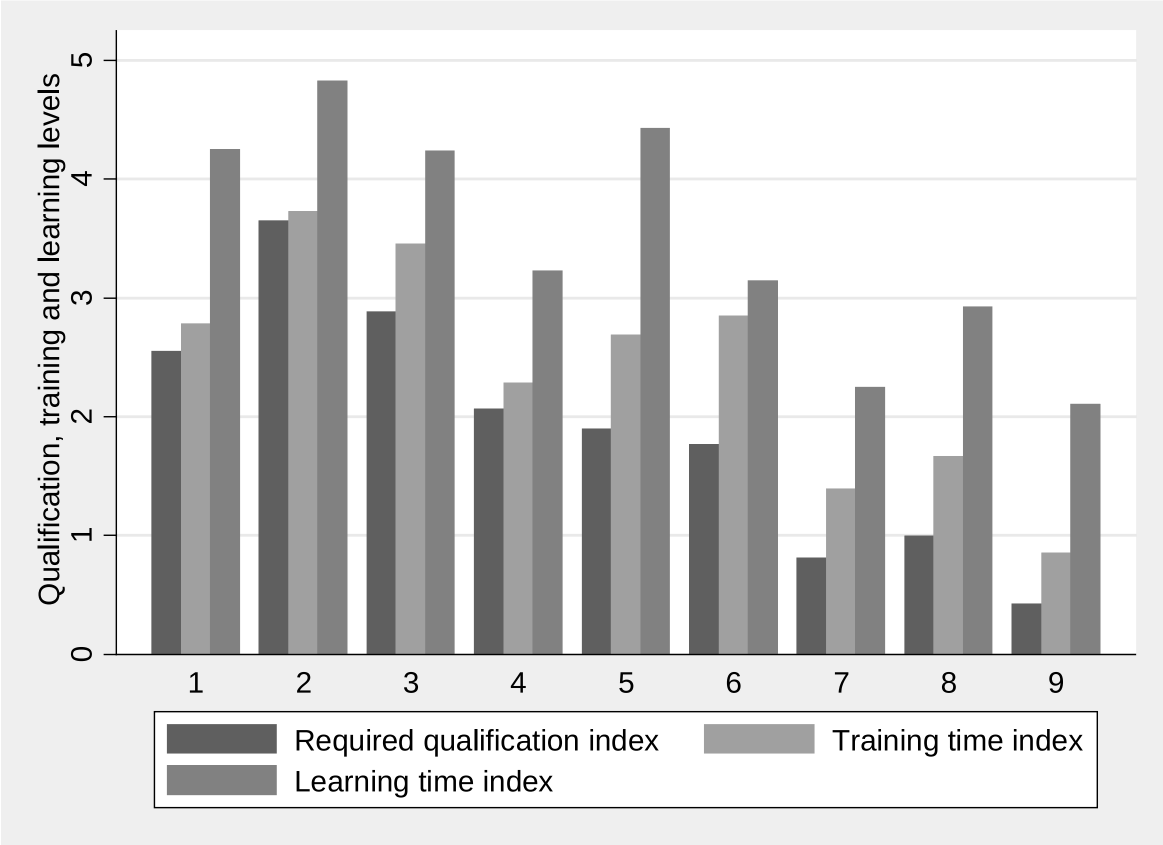


Figure 12B: Required education, experience and training – O*NET-based profiles

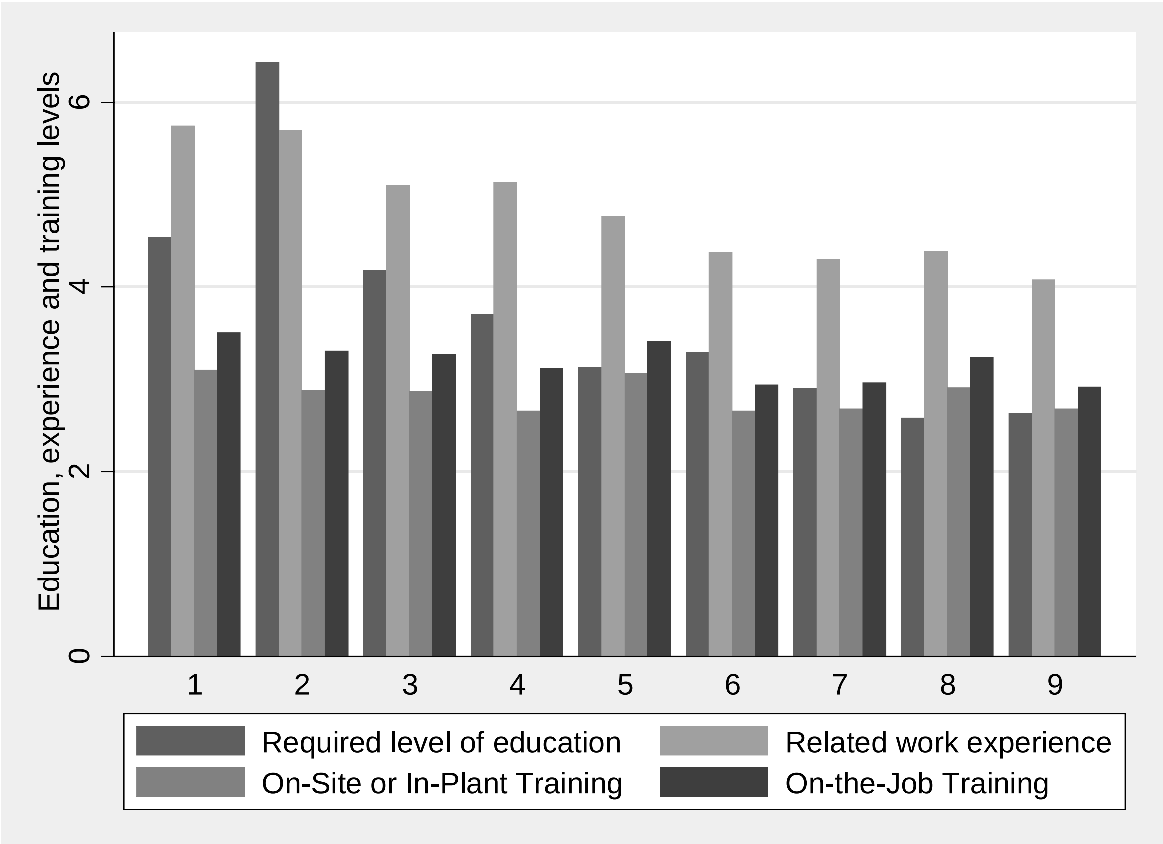
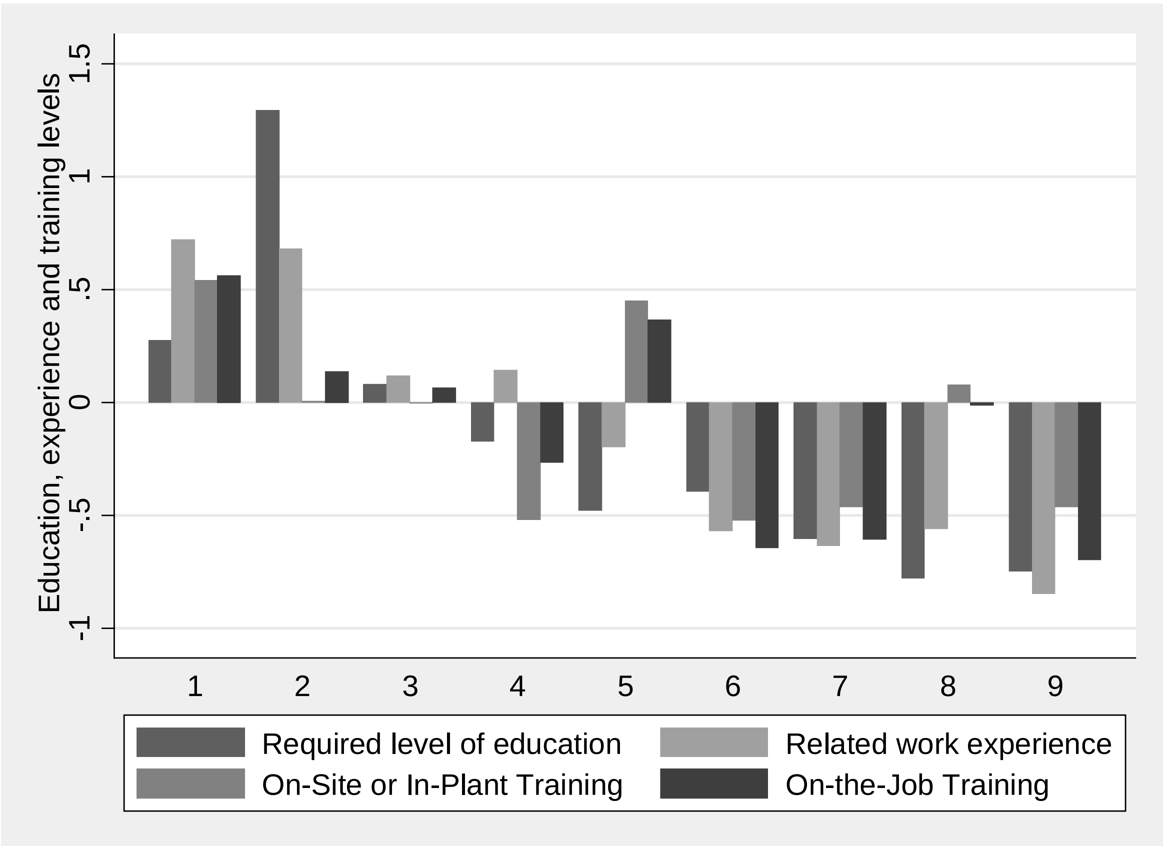


Figure 13: Standardised Education and Training Levels at SOC Major Group Level

Figure 13A: Required qualifications, training time and learning time – UK Skill Survey

Figure 13B: Required education, experience and training – O*NET-based profiles



Figures 12 and 13 present the results of these two exercises for the levels and for these levels standardised to facilitate a simpler visual comparison. In the first graph in each case, the required qualification, training time and learning time indices as computed from the Skills Survey data are presented. The patterns here are familiar. In the second graph, the profiles derived from the O*NET descriptors are presented. The patterns depicted are clearly very similar, and clearly the indices obtained are highly correlated with each other. For example, the correlation between the Skills Survey required qualification index and the O*NET-based required level of education index is 0.915, while that between the Skills Survey learning time index and the O*NET-based on-the job training measure is 0.836. Moreover, these are the strongest correlation for each of these variables. For example, while the learning time index is also positively correlated with the required qualification and training time indices, these are weaker (at 0.630 and 0.579 respectively) than its correlation with the on-the-job training index. Similar patterns are evident between the other measures considered.[25]

These results suggest that the matching methodology described and developed in this report is **valid and appropriate** (at least as far as these particular measures are concerned). The patterns produced replicate quite closely the results obtain from directly measuring these skills across the population as in the Skills Survey using very similar (although not identical) questions.

6.2.2 Sensitivity and Robustness

In this sub-section, we examine the sensitivity and robustness of our findings to the choices over weighting and aggregation that we can make as described in Chapter 4. Results are presented for both the data-people-things taxonomy of skills, as well as for the index of STEM skills.

Table 11 reports the correlations between the data, people and things skills at the SOC2010 unit group level for the measures derived using weights based on scores (Panel A) and weights based on counts (Panel B). Similarly, Table 12 reports the correlations between the STEM skills scores aggregated using simple averaging and derived using the first principal component of the eight contributing items. Once again, the correlations are presented separately for weights based on scores (Panel A) and weights based on counts (Panel B).

[25] Note that our O*NET-based profiles are based on SOC2010 whereas the 2006 Skills Survey is based on SOC2000. However, at the 1 digit Major Group level, there are no real differences between the two SOC classifications.

Table 11: Correlations between Data-People-Things measures based on different weights

Panel A: Weights based on scores

A1: Data skills scores

		Scores only			Scores and employment		
Scheme		A	B	C	A	B	C
Scores only	A	1.0000					
	B	0.9281	1.0000				
	C	0.9523	0.9670	1.0000			
Scores & emp	A	0.8979	0.8006	0.8671	1.0000		
	B	0.8925	0.9016	0.9258	0.9146	1.0000	
	C	0.8877	0.8494	0.9207	0.9568	0.9582	1.0000

A2: People skills scores

		Scores only			Scores and employment		
Scheme		A	B	C	A	B	C
Scores only	A	1.0000					
	B	0.9424	1.0000				
	C	0.9609	0.9688	1.0000			
Scores & emp	A	0.8858	0.8076	0.8689	1.0000		
	B	0.8816	0.8930	0.9214	0.9152	1.0000	
	C	0.8800	0.8505	0.9212	0.9529	0.9610	1.0000

A3: Things skills scores

		Scores only			Scores and employment		
Scheme		A	B	C	A	B	C
Scores only	A	1.0000					
	B	0.9411	1.0000				
	C	0.9607	0.9766	1.0000			
Scores & emp	A	0.9094	0.8488	0.8767	1.0000		
	B	0.8876	0.9222	0.9265	0.9118	1.0000	
	C	0.9064	0.8974	0.9349	0.9429	0.9655	1.0000

Panel B: Weights based on counts

B1: Data skills counts

		Counts only			Counts and employment		
Scheme		A	B	C	A	B	C
Counts only	A	1.0000					
	B	0.9257	1.0000				
	C	0.9501	0.9662	1.0000			
Counts only	A	0.8956	0.7957	0.8644	1.0000		
	B	0.8842	0.8928	0.9198	0.9156	1.0000	
	C	0.8818	0.8437	0.9175	0.9563	0.9583	1.0000

B2: People skills counts

		Counts only			Counts and employment		
Scheme		A	B	C	A	B	C
Counts only	A	1.0000					
	B	0.9405	1.0000				
	C	0.9597	0.9675	1.0000			
Counts only	A	0.8842	0.8029	0.8661	1.0000		
	B	0.8733	0.8863	0.9161	0.9140	1.0000	
	C	0.8739	0.8456	0.9185	0.9510	0.9609	1.0000

B3: Things skills counts

		Counts only			Counts and employment		
Scheme		A	B	C	A	B	C
Counts only	A	1.0000					
	B	0.9405	1.0000				
	C	0.9594	0.9759	1.0000			
Counts only	A	0.9044	0.8461	0.8758	1.0000		
	B	0.8823	0.9170	0.9230	0.9118	1.0000	
	C	0.8975	0.8914	0.9306	0.9428	0.9640	1.0000

Note: The 12 different weighting methods are described in Section 4.

Table 12: Correlations between STEM measures based on different weights

Panel A: Weights based on scores

A1: Mean STEM skills scores

Scheme		Scores only A	Scores only B	Scores only C	Scores and employment A	Scores and employment B	Scores and employment C
Scores only	A	1.0000					
	B	0.6305	1.0000				
	C	0.8470	0.6952	1.0000			
Scores & emp	A	0.7473	0.4134	0.6888	1.0000		
	B	0.5887	0.7447	0.6719	0.6920	1.0000	
	C	0.6918	0.4557	0.8194	0.8797	0.7359	1.0000

A2: PCA STEM skills scores

Scheme		Scores only A	Scores only B	Scores only C	Scores and employment A	Scores and employment B	Scores and employment C
Scores only	A	1.0000					
	B	0.9055	1.0000				
	C	0.9364	0.9613	1.0000			
Scores & emp	A	0.8493	0.7214	0.8007	1.0000		
	B	0.8399	0.8617	0.8887	0.8704	1.0000	
	C	0.8385	0.7977	0.8802	0.9318	0.9479	1.0000

Panel B: Weights based on counts

B1: Mean STEM skills counts

Scheme		Counts only A	Counts only B	Counts only C	Counts and employment A	Counts and employment B	Counts and employment C
Counts only	A	1.0000					
	B	0.6347	1.0000				
	C	0.8482	0.6814	1.0000			
Counts only	A	0.7323	0.4090	0.6931	1.0000		
	B	0.5880	0.7299	0.6688	0.7019	1.0000	
	C	0.6740	0.4358	0.8147	0.8841	0.7359	1.0000

B2: PCA STEM skills counts

Scheme		Counts only A	Counts only B	Counts only C	Counts and employment A	Counts and employment B	Counts and employment C
Counts only	A	1.0000					
	B	0.9009	1.0000				
	C	0.9326	0.9609	1.0000			
Counts only	A	0.8427	0.7124	0.7959	1.0000		
	B	0.8281	0.8464	0.8803	0.8727	1.0000	
	C	0.8277	0.7860	0.8720	0.9328	0.9494	1.0000

Note: The 12 different weighting methods are described in Section 4.

The correlations between the derived occupational skills measures are all very high whichever method of weighting is used. The correlations are somewhat higher within broad method (e.g. scheme A, B and C within scores and employment) than between methods (e.g. scores only vs. scores and employment). Somewhat surprisingly, taking into account relative employment makes little difference to the correlations – while they are marginally higher for the methods 'with employment', the difference is small in general.

Note that the correlations between the different skills measures we have analysed are all rather higher than the correlations between the weights themselves as reported in Table 6.

Finally, when comparing the STEM mean skills scores with the principal component analysis (PCA) skills scores, it would appear that simple averaging produces greater variation by scheme than PCA. Certainly there are more differences between the derived skills measures (as reflected in the weaker correlations between the skills measures) when means are used as compared to the resulting profiles when PCA is utilised. Whether scores or counts are used in the weighting scheme does not appear to affect this conclusion.

From this inspection of the occupational skills measures, we conclude that the methodology appears to be **quite insensitive to the choice of weighting scheme, and relatively robust to the method of aggregation**.

6.3 Further work

There are a number of possible additional refinements that could be made to the matching and aggregation processes described above and which might further enhance the resulting measures of occupational skills. First, within the matching process, CASCOT could be 'taught' (i,e. further rules added) to recognise US spellings etc. This would improve the quality of the matching process to better reflect the equivalences between jobs in the US and the UK, and thus between O*NET and UK SOC occupations. Second, poor matches (scores less than 39) in CASCOT could be rejected (i.e. treated as being unmatched). This is likely to produce a sparser occupational-to-occupation matrix of weights, so that each O*NET occupation will match with typically fewer UK SOC occupations. Finally, we could restrict the whole process to job titles associated with data-level occupations only – currently, all job titles from the full set of 1,102 O*NET occupational titles are utilised, even though there is only O*NET data gathered on 965 of these (see Annex C for details).[26]

[26] In practice, while the 137 non-data level occupations receive some weight when matching to UK SOCs, these weights are small once employment in the non-data level occupations is taken in to account as in our preferred weighting scheme.

As noted in Chapter 2, the O*NET system is being updated on a 5 year rolling basis, so that new information on approximately 20% of O*NET-SOC occupations is being gathered and made available every year. At a minimum, the matching and reweighting exercise described in this report should be repeated at 5 year intervals therefore. However, given the largely automated processes involved, it would not be difficult to update the profiles every year to reflect the changing distribution of employment across occupations both in the US and in the UK, as well as the new information on one fifth of the O*NET occupations.

7 Conclusions

Until we have a much better understanding of the skills that are in relative shortage (or excess demand) and surplus (excess supply), then it will not be possible to really make significant progress in closing the gaps and addressing the mismatch between the skills that individuals possess and those that they require in employment in the UK labour market today. Given the paucity of information and the inherent weaknesses in conventional measures of skills, this project has the potential to substantially advance our understanding and knowledge of the nature of skills demand and utilisation in the UK.

In this report, we have demonstrated the feasibility of adopting the detailed US O*NET system to describe the skills and other characteristics of individuals working in jobs in the UK today. We develop a systematic and transparent 'mapping' between the US and UK occupational classifications, and assign the job tasks, skills and other content of the US Occupational Information Network (O*NET) system to the matched UK occupations. This report serves to demonstrate the feasibility of the methodology, and provides some illustrative occupational skills profiles using the method that we have developed.

In order to assess the validity of the resulting profiles, we have compared derived measures of required qualifications and training time with similar measures taken from the 2006 Skill Survey. The correspondence between similar measures derived from the two different sources are very high – at least at the SOC Major Group level – giving us confidence in the validity and robustness of the methodology. Other patterns in derived profiles – for example for STEM skills – also conform to our priors, and also suggest that the method that we have developed has validity.

Using this methodology, we are able to provide a summary and assessment of myriad of job skills and attributes for the UK, and at a high level of occupational disaggregation. There are a number of possible uses for occupational profiles developed in this manner:

- Assessment of trends in skills demand as recorded by their changing utilisation in employment (rather than recording only the change in occupational composition or qualifications of the workforce)
- Provide estimates of future skills demand (rather than simply providing estimates of future patterns in employment by sector and occupation) by linking to *Working Futures* projections
- Supplying useful information to Information, Advice and Guidance (IAG) practitioners and careers advisors – and also to individuals – on the types of skills that are necessary for, and useful in employment today, and likely to be of importance and value in the future in terms of labour market outcomes

- Providing a detailed description of the spatial distribution of skills demand/utilisation
- Providing a more nuanced assessment of the 'job polarisation' debate which suggests that, increasingly, the structure of employment is being dominated by 'high skill' and 'low skill' jobs
- Estimating the value of skills in employment. For example, Black and Spitz-Oener (2010) have suggested that the changing nature of skills being used in the different jobs dominated by men and women has contributed significantly to reducing the gender pay gap in Germany
- Extending the information available to the Migration Advisory Committee (MAC) on the measurement of skills, and on the specific skills that are in shortage

ANNEX A: THE O*NET SYSTEM (Wilson, 2009)

A.1 Overview

The O*NET system is the primary source of occupational competency information in the US. It is available to all users online[27]. At its core is the O*NET database which contains detailed data on a large range of occupation-specific indicators, including tasks undertaken, pay and technical requirements. The database is updated on a continuous basis, drawing upon customised surveys and other material. The data collection and validation process for O*NET is complex – some details are given in Wilson (2009), Annex A.

O*NET has been described as a 'common language and dynamic system for describing the world of work for both the public and private sectors'. It is a comprehensive system for collecting, organising and disseminating information on occupational and worker requirements, based around the notion of competency, with emphasis on skills transferability.

A.2 Content

The content of the O*NET model is summarised in Figure A1. It looks at things from both an individual worker and an employer (job) perspective. It covers six main domains: Worker characteristics; Worker requirements; Experience requirements; Occupational requirements; Workforce characteristics; and Occupation specific information. These are described in turn:

Worker Characteristics cover enduring characteristics that may influence both work performance and the capacity to acquire knowledge and skills, including: abilities; occupational interests (encompassing personality traits); work values; and work styles.

Worker Requirements are descriptors referring to work-related attributes acquired and/or developed through experience and education. These include: basic skills; generic skills; knowledge; and prior education.

Experience Requirements relate to work related experience that may be needed, including: specific prior experience and training; basic and cross-functional skills and entry requirements; licensing and other certificates, registrations or credentials needed.

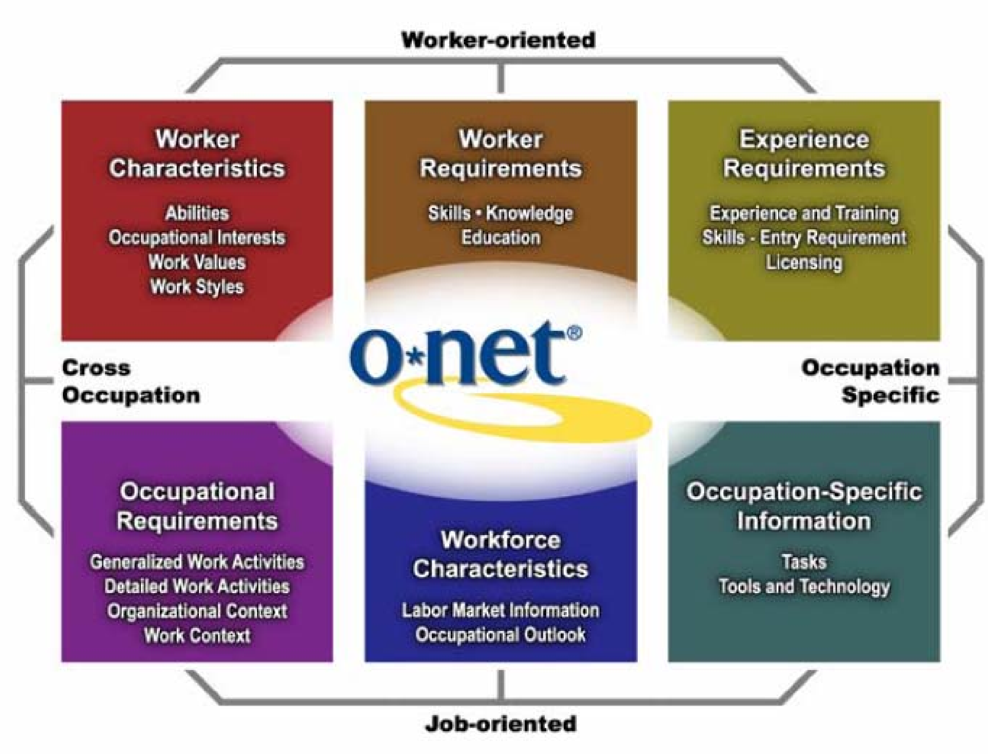
Occupation-Specific Information includes other Content Model elements needed in specific occupations, such as: particular tasks; and use of special tools, technology or machines that workers may need to function in the workplace.

[27] O*NET Resource Center (www.O*NETcenter.org).

Workforce Characteristics cover variables that define and describe the general characteristics of the occupations concerned, including: *Labour Market Information* (current labour force characteristics of those employed in the occupations); and *Occupational Outlook* (prospects for these occupations).

Occupational Requirements is a detailed set of elements that describe what various occupations require, covering: generalized and specific work activities (types of behaviours occurring in many different jobs); and organizational and general work context (the latter covering physical and social factors that influence the nature of work in the job).

Figure A1: The O*NET Content Model



Source: *Tippins and Hilton (2010), p.8.*

Occupation is defined using an extended version of the US Standard Occupational Classification. The O*NET SOC has been developed by a multi-agency initiative. It is structured for comparability, with four hierarchical levels. The O*NET-SOC currently distinguishes over 800 occupational categories. It is constantly growing, to include important new and emerging (N & E) occupations.

The level of detail enables much more subtle and sophisticated analysis of changing skill demands than is possible when only aggregate data are available. For example studies such as those by EMSI (2009) show how this kind of information allows a very detailed analysis of the possible impacts of climate change on employment patterns (both positive and negative), the opportunities for generating green jobs and what this means in terms of skill requirements and training priorities.

A.3 Updating

Information in the O*NET system is updated and published on a regular basis, with a new database released at least annually (for 2001-2006 there were two releases per year). Information on at least 100 occupations is refreshed on each release. A maximum of 5 years is allowed before information on any **key** occupation is refreshed. The average age of the information set for all occupations is 2.59 years, so the information is generally pretty up to date. Key occupations are defined as:

- Identified as ‘in-Demand’ by US Department of Labor (DOL);
- ‘Top 50 occupations’ as identified by DOL;
- High growth rates and/or large employment numbers;
- Linked to technology, maths and science, computers, engineering and innovation;
- Linked to ‘Job Zones’; or
- ‘Green’ occupations.

A.4 Dissemination

The database is freely available via the O*NET OnLine web-based application. This enables users to explore a huge range of occupational information, such as employment and pay levels, including job prospects, and skill requirements.

Users can see individual occupational summaries, but also related occupations, with similar requirements in terms of skills, knowledge, and tasks (which BLS refer to as ‘crosswalks’). The system includes a number of separate databases covering:

- Occupational classification and ‘crosswalks’;
- Occupational coding assistant;
- Training and e-Learning;
- Technical assistance (including testing & assessment guides);
- Questionnaires;
- Research & technical reports;
- Links to related sites, including national/state information.

A.5 Tools

Within the main website there is a variety of tools design to help users to make maximum use of the information. This includes the O*NET *Toolkit for Business*, which provides technical information detailing O*NET's many uses for employers and HR specialists. This includes detailed job descriptions, as well as information to help with succession planning, training needs analysis, career development, and general workforce development[28]. There is also a range of other tools aimed at workers and students searching for work or looking to change career direction (so called *Career Exploration Tools* (covering interests and abilities)[29].

O*NET draws upon the national industrial and occupational employment projections. The 10-year horizon projections produced by BLS are widely used in career guidance, as well as in education and training programme planning, and by all those interested in long-range employment trends. They continue a 60-year tradition of providing labour market information to individuals making choices about education and training, as well as to those entering the job market or changing careers.

In addition to the national level projections the DOL and BLS also support State and local level employment projections. The Department's Employment and Training Administration (ETA) provides funding for states to develop medium to long-term (10-year horizon) as well as short-term (2-year horizon) projections. Much of this information is made available via State Web sites.

The Projections Managing Partners Consortium, which includes representatives from BLS and ETA, as well as the individual States, helps to organise and coordinate these activities. They help to provide structure and guidance, as well as some software. This helps to ensure that the results are comparable.

Other websites also offer common resources and tools to facilitate strategic planning and benchmarking at regional level and to assist economic development and recovery programmes. These include:

- the 'Workforce Information and Economic Analysis' website, which recognises the role of good quality LMI and related and economic analysis for making sound local and regional economic development decisions, including strategic planning, benchmarking economic competitiveness, and measuring outcomes;[30]
- the WIN-WIN Network Community of Practice, which was established to advance the application and integration of data, analysis, and research to decision making in

[28] For details see: http://www.O*NETcenter.org/toolkit.html
[29] The Career Exploration Tool set includes: Ability Profiler; Interest Profiler; Computerised Interest Profiler; Work Importance Locator; and O*NET Work Importance Profiler. See http://www.O*NETcenter.org/dev_tools.html for further details.
[30] The website (http://www.workforce3one.org/page/wiea) offers resources and tools aimed at assisting employers, economic developers, and educational institutions, as well as jobseekers to understand and make best use of such LMI.

national and regional workforce and economic development and economic recovery efforts;[31]

- The DOL CareerOneStop offers career exploration resources and related LMI to job seekers, students, businesses, and workforce professionals, including matching skills and supporting career transitions);[32,33]
- The ETA's Industry Competency Model Initiative is aimed at promoting an understanding of the skill sets and competencies that are essential to educate and train a globally competitive workforce, and includes a number of elements related to competency models, their many uses, and how to exploit them from both a employer and an individual perspective.[34]

A.6 Users and Uses

O*NET is used by a wide range of different individuals and organisations, including:

- Students;
- Young people and other labour market entrants;
- Job seekers;
- Employers in general;
- Business analysts;
- Workforce and economic development specialists;
- Organisational consultants;
- HR professionals;
- Training specialists;
- Careers counsellors;

[31] The WIN-WIN network is intended to identify best practices and to share this information about methods and techniques with both data producers and consumers. It is also intended to promote capacity building and the development of new tools and technology that support analysis and research. For details see: http://winwin.workforce3one.org/

[32] The main site is at www.CareerOneStop.org. Highlights include: *America's Career InfO*NET* which is designed to help individuals explore career opportunities and make informed employment and education choices (see www.CareerInfO*NET.org); *Certification Finder* which is an online directory of occupational certification standards, provide by authoritative bodies (see http://www.careerinfO*NET.org/certifications_new/default.aspx);*Skills Profiler* which allows users to identify skills and activities used in a job, including a feature that identifies similarities and differences between the selected occupation and any other occupations (see http://www.careerinfO*NET.org/skills/default.aspx?nodeid=20); and the *Worker ReEmployment Portal* which is designed to assist and support workers made redundant (see www.careeronestop.org/ReEmployment).

[33] OneStop partners include: career counsellors; interviewers; rehabilitation counsellors; Veterans' representatives; training providers; and business consultants.

[34] The initiative has produced various resources and tools including key documents which explain the concept and illustrate practical uses as well as providing tools (see http://wdr.doleta.gov/directives/corr_doc.cfm?docn-2743). The *Competency Model Clearinghouse* (CMC) website also offers developers and users of competency models a variety of resources, tools, and links including easy to follow examples and illustrations (for details see: http://www.careeronestop.org/competencymodel/). The CMC site also offers interactive tools, including the *Build a Competency Model*) which enables users to customise national industry competency models to reflect specific workforce needs in a particular region or sub-industry) and *Career Ladder Tools* (designed to display the sequence of jobs or occupations with specific careers in a particular industry, including documentation of the requirements for each job and the critical development experiences needed to move up the career ladder).

- Government officials and policy makers;
- The military;
- Education and training providers;
- Teachers and lecturers;
- Researchers.

Amongst employers, O*NET is used for:

- Job matching, recruitment and training activities (including writing job descriptions, identifying competencies skills gaps and training needs);
- Developing training programmes and curriculum;
- Other human resources planning and related activities;
- Business forecasting and analysis.

It is widely used in large organisations and corporations, in both private and public sectors, including many famous names such as Boeing, Manpower and Microsoft. But its availability via the net also make it accessible to small and medium size enterprises and individuals.

Individuals use O*NET for career exploration and development, job search and employment transitions. O*NET enables people to learn what jobs might fit their personal interests, skills and experience as well as highlighting the different skills required for different jobs and which occupations and industries are in demand based on the latest workforce information. The system identifies success factors associated with different occupations, including the types of qualifications and competences need to enter and advance in that particular job.

A.7 New & Emerging and 'Green' Jobs

O*NET and the OES are being revamped to focus on so called 'Green Occupations'[35]. The BLS has focused on three main aspects:

- the extent to which 'green' issues are shaping patterns of economic activity and technology, increasing the demand for some existing industries and occupations;
- the way that they are altering the nature of the tasks and competences needed for existing jobs; and
- the ways in which they may generate new work and worker requirements.

[35] The BLS defines the 'Green Economy' as covering "Economic activity related to reducing the use of fossil fuels, decreasing pollution and greenhouse gas emissions, increasing the efficiency of energy usage, recycling materials and developing and adopting renewable sources of energy."

The former are referred to as ‘Green Increased Demand Occupations’. Generally these face few (if any) significant changes in tasks or job requirements. The second group are termed ‘Green Enhanced Skills Occupations’. For these, even though the essential nature of the job remain the same, tasks, skills, knowledge, and external elements, such as credentials, may have changed so that there are some significant changes in worker requirements. The final group are called ‘Green New & Emerging Occupations’ by BLS. They reflect jobs where the impact of green economy activities and technologies is creating new types of activity and work, possibly resulting in the generation of new occupational titles.

The O*NET ‘New & Emerging’ project identifies ‘new’ occupations. These are defined as:

- significantly different from existing occupations and not adequately reflected in the current SOC;
- have significant employment; and
- are expected to see positive projected growth.

The work to identify such occupations is undertaken in conjunction with education and certification programmes, and involving related professional associations.

A.8 Monitoring of Use and Evaluation

The BLS monitors use of O*NET and its other systems. It reports extensive downloading of data and reports and widespread and intense use of its websites. Response rates to the O*NET surveys are improving over time (2001-2008), all of which points to a service which is regarded as of great value by it users. An official, independent and comprehensive evaluation of O*NET (at Federal level) has recently been undertaken (Tippins and Hilton, 2010).

Many of the users of O*NET are very positive. For example the Brookings Foundation says:

> “We find that O*NET is indispensable to the development of the nation’s workforce. By providing a common taxonomy and highly detailed information on the characteristics of 800 occupations, O*NET:
>
> serves as the foundation for critical workforce delivery systems,
>
> enables interaction and cooperation across the workforce development community, and,
>
> most importantly, allows jobseekers, employers, educators, and workforce professionals to make more informed choices.” [36]

[36] See: http://www7.nationalacademies.org/cfe/Andrew%20Reamer%20Comments%20on%20O*NET.pdf and for some other typically positive endorsements: http://www7.nationalacademies.org/cfe/Mercer%20Associates%20Comments%20on%20O*NET.pdf.

The current budget for O*NET is over $6 million per annum (Tippins and Hilton, 2010, p.14).

Annex B: UK SOC2010 Occupational Classification

Full details on the UK SOC2010 can be obtained from the ONS at: http://www.ons.gov.uk/about-statistics/classifications/current/soc2010/index.html

Box B1: UK-SOC2010 Classification

Jobs are classified into groups according to the concept of 'skill level' and 'skill specialisation'. As in SOC2000 and its predecessor SOC90, skill level is defined with respect to the duration of training and/or work experience recognised in the field of employment concerned as being normally required in order to perform the activities related to a job in a competent and efficient manner.

Skill specialisation is defined as the field of knowledge required for competent, thorough and efficient conduct of the tasks. In some areas of the classification it refers also to the type of work performed (for example materials worked with, tools used).

Skill levels are approximated by the length of time deemed necessary for a person to become fully competent in the performance of the tasks associated with a job. This, in turn, is a function of the time taken to gain necessary formal qualifications or the required amount of work-based training. Apart from formal training and qualifications, some tasks require varying types of experience, possibly in other tasks, for competence to be acquired. Within the broad structure of the classification major groups and sub-major groups reference can be made to these four skill levels:

- The first skill level equates with the competence associated with a general education, usually acquired by the time a person completes his/her compulsory education and signalled via a satisfactory set of school-leaving examination grades. Competent performance of jobs classified at this level will also involve knowledge of appropriate health and safety regulations and may require short periods of work-related training. Examples of occupations defined at this skill level within the SOC2010 include postal workers, hotel porters, cleaners and catering assistants

- The second skill level covers a large group of occupations, all of which require the knowledge provided via a good general education as for occupations at the first skill level, but which typically have a longer period of work-related training or work experience. Occupations classified at this level include machine operation, driving, caring occupations, retailing, and clerical and secretarial occupations.

- The third skill level applies to occupations that normally require a body of knowledge associated with a period of post-compulsory education but not normally to degree level. A number of technical occupations fall into this category, as do a variety of trades occupations and proprietors of small businesses. In the latter case, educational qualifications at sub-degree level or a lengthy period of vocational training may not be a necessary prerequisite for competent performance of tasks, but a significant period of work experience is typical.

- The fourth skill level relates to what are termed 'professional' occupations and high level managerial positions in corporate enterprises or national/local government. Occupations at this level normally require a degree or equivalent period of relevant work experience.

Source: http://www.ons.gov.uk/about-statistics/classifications/current/soc2010/soc2010-volume-1-structure-and-descriptions-of-unit-groups/index.html

The structure of UK-SOC2010 comprises nine major groups (1 digit) , 25 sub-major groups (2 digit), 90 minor groups (3 digit) and 369 unit groups (4 digit). To illustrate, the first 5 unit group occupations in SOC2010 are as follows:

Major Group	Submajor Group	Minor Group	Unit Group	Group Title
1				**MANAGERS, DIRECTORS AND SENIOR OFFICIALS**
	11			**CORPORATE MANAGERS AND DIRECTORS**
		111		**Chief Executives and Senior Officials**
			1115	Chief executives and senior officials
			1116	Elected officers and representatives
		112		**Production Managers and Directors**
			1121	Production managers and directors in manufacturing
			1122	Production managers and directors in construction
			1123	Production managers and directors in mining and energy

The major and sub-major groups, and associated skill level specialisations are presented in Table B1. As can be seen, the skill specialisation criterion is used to distinguish (sub major) groups of occupations within each skill level.

Table B1: SOC2010 Major Groups, Sub-major Groups and Skill Specialisation levels

	Major group		Sub-Major Groups	Skill level
1	Managers, directors and senior officials	11	Corporate managers and directors	4
		12	Other managers and proprietors	3
2	Professional occupations	21	Science, research, engineering and technology professionals	4
		22	Health professionals	4
		23	Teaching and educational professionals	4
		24	Business, media and public service professionals	4
3	Associate professional and technical occupations	31	Science, engineering and technology associate professionals	3
		32	Health and social care associate professionals	3
		33	Protective service occupations	3
		34	Culture, media and sports occupations	3
		35	Business and public service associate professionals	3
4	Administrative and secretarial occupations	41	Administrative occupations	2
		42	Secretarial and related occupations	2
5	Skilled trades occupations	51	Skilled agricultural and related trades	3
		52	Skilled metal, electrical and electronic trades	3
		53	Skilled construction and building trades	3
		54	Textiles, printing and other skilled trades	3
6	Caring, leisure and other service occupations	61	Caring personal service occupations	2
		62	Leisure, travel and related personal service occupations	2
7	Sales and customer service occupations	71	Sales occupations	2
		72	Customer service occupations	2
8	Process, plant and machine operatives	81	Process, plant and machine operatives	2
		82	Transport and mobile machine drivers and operatives	2
9	Elementary occupations	91	Elementary trades and related occupations	1
		92	Elementary administration and service occupations	1

Source: *SOC 2010: Volume 1: Structure and Description of Unit Groups*

Annex C: US SOC and O*NET-SOC Occupational Classifications

The O*NET-SOC is a slightly modified (i.e. extended) version of the US Standard Occupational Classification US-SOC. Full details on the O*NET occupational classification system can be found at: http://www.onetcenter.org/. Both classifications and their inter-relationship are described in Box C1 below.

Box C1: US-SOC2009 and O*NET-SOC2009

The structure of the US-SOC2009 system includes four levels of aggregation: 23 major groups, 96 minor groups, 449 broad occupations and 821 detailed occupations. All SOC occupations are assigned a six-digit code. The 23 major groups of the SOC are as follows:

11-0000 Management Occupations
13-0000 Business and Financial Operations Occupations
15-0000 Computer and Mathematical Occupations
17-0000 Architecture and Engineering Occupations
19-0000 Life, Physical, and Social Science Occupations
21-0000 Community and Social Services Occupations
23-0000 Legal Occupations
25-0000 Education, Training, and Library Occupations
27-0000 Arts, Design, Entertainment, Sports, and Media Occupations
29-0000 Healthcare Practitioners and Technical Occupations
31-0000 Healthcare Support Occupations
33-0000 Protective Service Occupations
35-0000 Food Preparation and Serving Related Occupations
37-0000 Building and Grounds Cleaning and Maintenance Occupations
39-0000 Personal Care and Service Occupations
41-0000 Sales and Related Occupations
43-0000 Office and Administrative Support Occupations
45-0000 Farming, Fishing, and Forestry Occupations
47-0000 Construction and Extraction Occupations
49-0000 Installation, Maintenance, and Repair Occupations
51-0000 Production Occupations
53-0000 Transportation and Material Moving Occupations
55-0000 Military Specific Occupations

SOC minor groups, broad occupations, and detailed occupations are assigned codes related to the corresponding major groups. For example:

19-0000 Life, Physical, and Social Science Occupations (SOC major group)
    19-4000 Life, Physical and Social Science Technicians (SOC minor group)
        19-4050 Nuclear Technicians (SOC broad occupation)
            19-4051 Nuclear Technicians (SOC detailed occupation)

In the O*NET-SOC2009 taxonomy, an occupation that is directly adopted from the SOC system is assigned the six-digit SOC code, along with a **.00** extension. If directly adopted from the SOC, the SOC title and definition are also used. Hereafter, these are referred to as SOC-level occupations.

If the O*NET-SOC occupation is more detailed than the original SOC detailed occupation, it is assigned the six-digit SOC code from which it originated, along with a two-digit extension starting with **.01**, then **.02**, **.03** and so on, depending on the number of detailed O*NET-SOC occupations linked to the particular SOC detailed occupation. For example, Nuclear Technicians is a SOC detailed occupation to which two detailed O*NET-SOC occupations are linked. See the occupational codes and titles for this example below.

19-4051.00 Nuclear Technicians (SOC-level)
- 19-4051.01 Nuclear Equipment Operation Technicians (detailed O*NET-SOC occ.)
- 19-4051.02 Nuclear Monitoring Technicians (detailed O*NET-SOC occupation)

Both 19-4051.01 Nuclear Equipment Operation Technicians and 19-4051.02 Nuclear Monitoring Technicians are data-level occupations in the O*NET taxonomy. Data-level occupations are those for which the O*NET program collects data from job incumbents, occupational experts, and occupational analysts on a wide variety of variables and scales, such as occupational characteristics and worker requirements drawn from the O*NET Content Model.

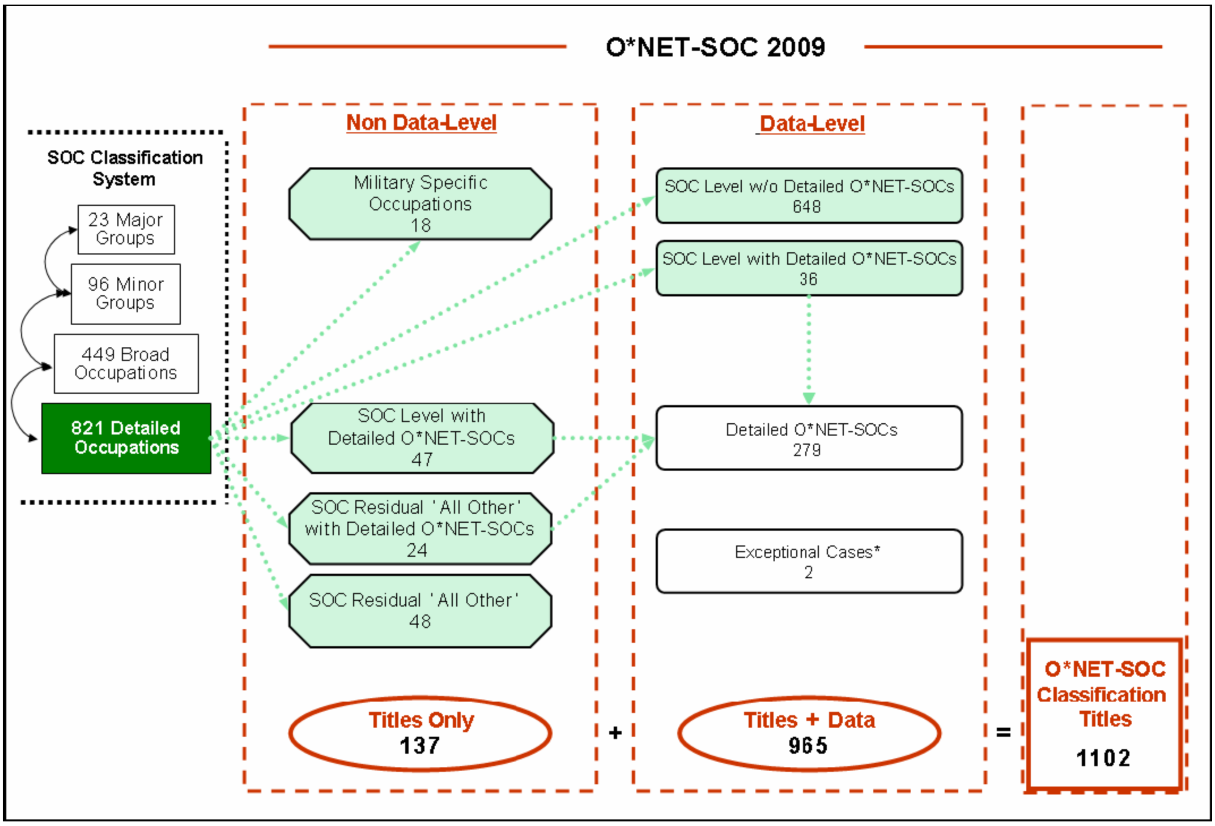
In the example above, the two detailed O*NET-SOC occupations, 19-4051.01 Nuclear Equipment Operation Technicians and 19-4051.02 Nuclear Monitoring Technicians, are data-level occupations, whereas the SOC detailed occupation, 19-4051.00 Nuclear Technicians, is not an O*NET data-level occupation.

Source: *National Center for O*NET Development (2009, 2010).*

The precise relationship between the US SOC2009 and O*NET-SOC2009 is depicted in Figure C1. As shown in Figure C1, the O*NET-SOC2009 taxonomy contains 1,102 occupational titles, 965 of which represent data-level occupations. Of the 965 data-level occupations (which includes 159 New & Emerging occupations identified within 17 in-demand industry clusters), 648 are SOC-level occupations adapted directly from the SOC, 36 are SOC-level occupations adapted directly from the SOC and also contain more detailed O*NET-SOC occupations, 279 are detailed O*NET-SOC occupations, and 2 occupations are exceptional cases. The remaining 137 are non-data level occupations (occupational titles only): 18 military occupational titles, 48 SOC residual ‘All Other’ occupational titles, 24 SOC residual ‘All Other’ occupational titles to which more detailed

O*NET-SOCs are linked, and 47 SOC detailed occupations to which more detailed O*NET-SOC occupations are linked[37].

Figure C1. Summary of the O*NET-SOC2009 Taxonomy



** The 2 exceptional cases include detailed O*NET-SOC occupations subsumed under broad-level*

***Source:** National Center for O*NET Development (2009, p.15).*

[37] For US-SOC and O*NET-SOC, we are using the current 2009 classifications. However, it should be noted that the O*NET system is about to update its SOC classification to a new O*NET-SOC2010 version. This new taxonomy will be used with the next release (Version 15.1) of the O*NET database. The O*NET-SOC2010 taxonomy is designed to be compatible with changes made to the US SOC2010 and to align the two classification systems. This modification to the O*NET SOC will not cause any immediate problems for our project but will have implications for potential future revisions. The O*NET-SOC2010 taxonomy will have 1,110 occupational titles, 974 of which will have data within the O*NET system. Much of the information for O*NET-SOC2009 will carry over, but it may be appropriate to iterate our matching of job titles to the O*NET-SOC2010 once the new O*NET occupational classification is operational and the job title files have been revised accordingly. The modification and implementation of the O*NET-SOC2010 taxonomy is described in National Center for O*NET Development (2010).

Bibliography

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

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

Autor D H and Handel M J (2009), "Putting tasks to the test: Human capital, job tasks and Wages", NBER Working Paper No. 15116, June 2009.

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

EMSI (2009). *Green Pathways: A data-driven approach to defining, quantifying, and harnessing the green economy*. Green Jobs, Part 3, Data Spotlight, March, 2009. http://www.economicmodeling.com/resources/wp-content/uploads/emsi-green-jobs-3-green-pathways.pdf

Everitt B S (1993, 3rd edition). *Cluster Analysis*, Arnold: London.

Felstead A, Gallie D, Green F and Zhou Y (2007), *Skills at Work 1986-2006*, ESRC Centre on Skills, Knowledge and Organisational Performance, SKOPE, Universities of Oxford and Cardiff.

Goos M, Manning A, and Salomons A (2009), "Job polarization in Europe", *American Economic Review Papers and Proceedings*, 99 (2), pp.58-63.

Goos M, Manning A, and Salomons A (2010), "Explaining Job Polarization in Europe: The Roles of Technology, Globalization and Institutions", CEP DP 1026, November 2010.

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

Handel M J (2008), "Measuring job content: Skills, technology, and management practices", Institute for Research on Poverty Discussion Paper No. 1357-08. Madison: University of Wisconsin.

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

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

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

Kim J O and Mueller C W (1978), Introduction to Factor Analysis: What it is and How to do it, Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-013, Sage: Thousand Oaks, CA.

Koucký J, Kovařovic J and Lepič M (2010), “Forecasting skills supply and demand in Europe to 2020: Occupational skills profiles – methodology and application”, mimeo February 2010, Education Policy Centre (EPC), Charles University in Prague.

National Center for O*NET Development (2009), “New and Emerging Occupations of the 21st Century: Updating the O*NET-SOC Taxonomy - Summary and Implementation”, Prepared for U.S. Department of Labor Employment and Training Administration, Washington, DC, March, 2009

National Center for O*NET Development (2010) “Updating the O*NET-SOC Taxonomy: Incorporating the 2010 SOC Structure - Summary and Implementation”, Prepared for U.S. Department of Labor Employment and Training Administration, Washington, DC, December 2010.

Sattinger M (1979), “Differential Rents and the Distribution of Earnings”, *Oxford Economic Papers*, 31(1), pp.60-71.

Sattinger M (1993), “Assignment models of the distribution of earnings”, *Journal of Economic Literature*, XXXI, 831–80.

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

Spitz-Oener A (2006), “Technical Change, Job Tasks, and Rising Educational Demands: Looking Outside the Wage Structure,” *Journal of Labor Economics*, 24(2), pp.235-270.

Taylor P J, Li W, Shi K & Borman W C (2008) “The transportability of job information across countries”. *Personnel Psychology*, 61(1), pp.69-111.

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

Wilson R A (2009), *Lessons from America: What can we learn from US approaches to anticipation of labour market change?,* Think-piece as a contribution to the UK Commission Evidence Base on Skills and Employment in the UK.

Wilson R A (2010), *Forecasting skill supply and demand in Europe: Inception Report*. Cedefop (European Centre for the Development of Vocational Training), Thessaloniki.

Evidence Reports present detailed findings of the research produced by the UK Commission for Employment and Skills. The reports contribute to the accumulation of knowledge and intelligence on skills and employment issues through the review of existing evidence or through primary research. All of the outputs of the UK Commission can be accessed on our website at www.ukces.org.uk

UKCES
Renaissance House
Adwick Park
Wath upon Dearne
South Yorkshire
S63 5NB
T +44 (0)1709 774 800
F +44 (0)1709 774 801

UKCES
28-30 Grosvenor Gardens
London
SW1W 0TT
T +44 (0)20 7881 8900
F +44 (0)20 7881 8999

This document is available at
www.ukces.org.uk under 'Publications'

ISBN 978-1-908418-04-3