A review of occupation-based social classifications for social research

Vernon Gayle
Roxanne Connelly
Paul Lambert
ABSTRACT

This working paper is a review of issues associated with measuring occupations and using occupation-based socio-economic classifications in social science research. The review is orientated towards researchers who undertake secondary analyses of large-scale micro-level social science datasets. The paper begins with an outline of how to handle raw occupational information. This is followed by an introduction to the two main approaches to measuring occupations, and a third lesser known but intellectually innovative approach. The three approaches are social class schemes, social stratification scales and the microclass approach. International comparisons are briefly described and a discussion of intersectionality with other key variables such as age and gender is provided.

We are careful to emphasise that we are not advocating the uncritical adoption of any one particular socio-economic measure over and above other alternatives. Rather we are advocating that researchers should choose from the portfolio of existing socio-economic measures in an informed and empirically defensible way and we strongly advocate undertaking sensitivity analyses. We conclude that researchers should always use existing socio-economic measures that have agreed upon and well documented standards. We strongly advise researchers not to develop their own measures without strong justification, nor to use existing measures in an un-prescribed or ad hoc manner.

KEYWORDS

Measuring Occupations; Social Stratification; Social Classification; Social Class; Microclass.

EDITORIAL NOTE

Professor Vernon Gayle is Professor of Sociology and Social Statistics at the University of Edinburgh.

Dr Roxanne Connelly is a Research Fellow at the University of Edinburgh.

Professor Paul Lambert is Professor of Sociology at the University of Stirling.

Corresponding author: Professor Vernon Gayle, Vernon.Gayle@ed.ac.uk.
ACKNOWLEDGEMENTS

We gratefully acknowledge the comments of Professor Geoff Payne, Professor John Field, Dr Alasdair Rutherford and Dr Christopher Playford.

© Vernon Gayle, Roxanne Connelly and Paul Lambert all rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This Working Paper Series is edited by Teresa McGowan.
# TABLE OF CONTENTS

1. INTRODUCTION .................................................................................... 1

2. THE RATIONALE FOR OCCUPATION-BASED SOCIAL CLASSIFICATIONS ................................................................................ 2

3. STANDARDISED OCCUPATIONAL CLASSIFICATIONS .......................... 5

4. OCCUPATION-BASED SOCIAL CLASSIFICATIONS ................................. 7

5. SOCIAL CLASS SCHEMES ...................................................................... 8

6. SOCIAL STRATIFICATION SCALES ...................................................... 13

7. ‘MICROCLASS’ APPROACHES .............................................................. 16

8. SELECTING AN APPROPRIATE SOCIO-ECONOMIC MEASURE .............. 18

9. SOCIO-ECONOMIC CLASSIFICATIONS AND RELATIONSHIPS WITH DEMOGRAPHIC STRUCTURES ......................................................... 18

10. CONCLUSION ........................................................................................ 21

REFERENCES ................................................................................. 23
1. INTRODUCTION
Occupations are a key element of contemporary social life and occupation-based indicators are central to social research. This working paper is a review of issues associated with measuring occupations and using occupation-based socio-economic classifications in social science research. Theoretical and empirical research on social inequalities or ‘social stratification’ has been one of the hallmarks of UK sociology since it burgeoned as an academic discipline after the Second World War (Pevalin and Rose 2002). Central to this field is the recognition that the occupational structure is an important foundation for the main dimension of social stratification (Blau and Duncan 1967). Within sociology there is a longstanding recognition that in industrialised societies occupations are the most powerful single indicator of levels of material reward, social standing and life chances (Parkin 1971).

Despite the sociological consensus that occupations are central to understanding social stratification, there is no single obvious and agreed upon way of measuring occupations. The discipline has produced a wide range of measures which are often linked with mainstream sociological theories and concepts, most notably related to ‘social class’, and debates have been exacerbated because of the complex nature of the occupational structure in contemporary labour markets. The theorisation, measurement and operationalisation of occupation-based measures have received great attention in the specialised field of social stratification research. The use and interpretation of occupational classifications has received less attention in other areas of sociology and in neighbouring social science disciplines, despite the centrality of occupations for many empirical analyses. It is our conjecture that appropriately measuring occupations and using occupational information in analyses requires thought and in-depth knowledge of these measures. The aim of this review is to document issues associated with utilising occupation-based social classifications in social research, and to provide some clear prescriptions for data analysts who are not experts in this field.

We begin with an outline of how to handle raw occupational information. This is followed by an introduction to the two major approaches to measuring occupations, and a third lesser known but intellectually innovative approach. The three approaches are social class
schemes, social stratification scales and the microclass approach. In addition, occupation-based measures for international comparisons are briefly described. We conclude with a discussion of the intersectionality of other key variables such as age and gender, and discuss a range of issues related to the inclusion of these measures in social science analyses.

2. THE RATIONALE FOR OCCUPATION-BASED SOCIAL CLASSIFICATIONS
The most common justification for using occupation-based socio-economic measures is that they make reliable, parsimonious indicators of the social positions of individuals (Parkin 1971; Rose and Pevalin 2003). To most sociologists, occupation-based socio-economic measures do not simply act as a proxy where income data themselves are unavailable, they are measures designed to help us better understand fundamental forms of social relations and inequalities to which income is merely epiphenomenal (Rose and Pevalin 2003). Empirical inquiries using repeated contacts data have convincingly shown that there is a high degree of income churning from year to year which makes income data unlikely to consistently represent individuals’ positions in industrial economies (Jarvis and Jenkins 1997, Jenkins and Van Kerm 2009). By contrast occupation-based socio-economic measures are more stable and therefore better describe lifetime earnings profiles (Rose and Pevalin 2003, Goldthorpe and McKnight 2006).

Hauser and Warren (1997) contend that the social sciences have been suffering from a preoccupation with measures of income and poverty. This focus possibly stems from the assumed utility of monetary measures for impact or ‘real world’ relevance. The focus on income might also reflect the relative disciplinary esteem of the field of economics within the social sciences. Bourguignon (2006) and Goldthorpe (2012) both assert that it is possible that this economic focus may have diverted some social scientists from major and consequential dimensions of social inequality which are not captured by focusing purely on income.

It is necessary to note at this point that some contemporary sociologists dispute the continued relevance of occupation-based social classifications. Against the backdrop of a vast
quantity of empirical results charting continued class-based inequalities (e.g. Erikson, Goldthorpe et al. 1979, Goldthorpe, Llewellyn et al. 1980, Goldthorpe, Llewellyn et al. 1987, Erikson and Goldthorpe 1992, Wright 1997), a parallel stream of literature has claimed that ‘class as a concept is ceasing to do any useful work’ (Pahl 1989) or is indeed ‘dead’ (e.g. Holton and Turner 1989, Clark and Seymour 1991, Kingston 1994, Joyce 1995, Lee and Turner 1996, Pakulski and Waters 1996). These theories generally argue that the lives and experiences of individuals in modern society are too fluid and transient, and too influenced by the processes of globalisation, to fit neatly within class categories. Pakulski and Waters’ (1996) account of the ‘death of class’ centres on two main ideas. First, that class based divisions peaked in industrial society and have been declining ever since. Second, that although there are inequalities in modern society these are not aligned with traditional social classes.

There are however a number of weaknesses in the ‘end of class’ thesis. Goldthorpe and Marshall (1992) note that the concept of class which is being attacked is a concept which is never clearly defined, is most aligned to the Marxist tradition, and which differs greatly from the more recent sociological concepts of social class that have been developed and deployed in empirical studies. The more nuanced aspects of the concept of social class are largely overlooked by those who argue that class is dead, and social class is often represented in a caricatured and simplistic manner (Goldthorpe and Marshall 1992). Platt (2011, pp.15) highlights that central concerns in contemporary class analysis include the notion of changes in the influence of class, the declining importance of class, and the intersectionality of other variables such as gender and ethnicity. We would add that a central theme in contemporary class analysis is the study of the extent to which the influence of social class has decreased over time in relation to major economic and social changes (e.g. Shavit and Blossfeld 1993, Hochstadt 1999, Breen 2004, Zijdeman 2009, Erikson and Goldthorpe 2010).

Whilst class analysts have researched these issues in some depth, Goldthorpe and Marshall (1992) assert that sociologists who claim that class is dead have provided little convincing evidence to support their arguments. Furthermore, Goldthorpe and Marshall (1992) have also noted that there has been no attempt to provide longitudinal evidence of change in the nature or influence of class to provide adequate support for the ‘death of class’ argument. It is
also notable that many theoretically oriented sociologists, from very different standpoints, have continued to describe the importance and relevance of class in contemporary society (Giddens 1981, Skeggs 1997, Sayer 2005).

Savage et al. (2013) recently proposed a new model of social class based on sociological theory advanced by the French sociologist Bourdieu. From this theoretical standpoint occupations are not the main indicators of social stratification positions. This work has led to an extensive amount of discussion over the last year. Savage et al. (2013) have developed a new model of class based on the concepts of economic capital (e.g. income and wealth), cultural capital (e.g. engagement with cultural goods and activities), and social capital (e.g. social contacts and networks) (see Bourdieu 1984). These concepts are measured using indicators such as household income, savings, property value, the number of social contacts held and the occupations of these social contacts, engagement with ‘highbrow culture’ (e.g. visiting museums or listening to classical music) and engagement with ‘emerging cultural capital’. Emerging cultural capital describes activities once considered ‘lowlbrow culture’ but that may now be more ubiquitous. Examples might include using social networking websites, going to a gym, or on-line gaming.

Rather than theorising occupations as the main basis of the opportunity structure, Bourdieu (1984) argues that the three capitals can be used to better explain the processes of social reproduction. Based on this theory Savage et al. (2013) contend that by measuring individuals’ levels of these capitals, a far more informative social class scheme can be developed than the traditional occupation-based measures that are widely used. Payne (2013) has noted that the seven ‘new’ classes proposed by Savage et al. (2013) are very similar to the established UK National Statistics Socio-Economic Classification categories, but with the added distinction of an ‘elite’ category. Therefore this new scheme does not represent a revolutionary re-working of the established social class categories. The measure proposed by Savage et al. (2013) has been extensively critiqued (Lambert and Griffiths 2013, Mills 2013, Payne 2013, Rose and Harrison 2013, Bradley 2014, Mills 2014, Rollock 2014). Despite the critiques offered by the ‘end of class’ theorists, and the direction proposed by Savage et al. (2013), there is no strong empirical
evidence that dissuades us of the extremely high value of using existing occupation-based measures in the secondary analysis of large-scale social surveys.

3. STANDARDISED OCCUPATIONAL CLASSIFICATIONS
Modern industrial economies are characterised by a wide range of occupations each of which contain numerous different jobs. Collecting job related information (e.g. job titles) is routinely the first step in collecting occupational information within social surveys. The next step is usually marshalling this information into a recognised occupational scheme.

The raw occupational information in large-scale micro-level social surveys is usually coded into a standardised occupational unit group scheme. In the UK, it is common for social survey data to be coded into the Office for National Statistics Standardised Occupational Classification, which are known as SOC codes (see Office for National Statistics 2010). Some surveys are coded into the International Labour Organisation’s International Standard Classification of Occupations, which are known as ISCO codes (see Ganzeboom 2010). This information is often augmented with additional employment data such as employment status (e.g. self-employed or supervisory). Occupational unit group codes are produced by matching original textual occupational descriptions (e.g. from survey question responses) with a standardised list of occupations. We consider that it is extremely important that data collectors maintain occupational data in the form of an established protocol (e.g. using SOC codes). Coding raw occupational data (e.g. textual descriptions) directly into socio-economic measures is highly unsatisfactory because it will result in the loss of valuable detailed occupational information. As Lambert (2002) demonstrates without the use of an established protocol for coding raw occupational information (e.g. using SOC codes) it is later impossible to test for comparability between both current and future occupation-based measures.

Translating raw survey data into unit group codes can be a time consuming exercise, but the burden is greatly reduced through the use of computer assisted and computer automated
coding procedures (Elias, Halstead et al. 1993). The Computer Aided Structured Coding Tool\(^1\) (CASCOT) is an online resource for the quick and reliable coding of occupational descriptions, which was developed by the Institute of Employment Research at the University of Warwick (Jones 2004). The CASCOT program compares the text in the description of an occupation with the text in standardised descriptors for occupational classifications. The software then presents a list of recommended matches. CASCOT also provides a score for the matches indicating the degree of certainty that the given occupational code is correct. The Office for National Statistics also publish an open access online coding tool\(^2\) that operates in a similar manner, and these and other coding software are available as off-line packages suitable for bulk-processing large volumes of data.

Schemes of unit group codes are updated periodically and the current nationally specific UK scheme is SOC2010\(^3\). Another example of a standardised occupational code is the International Standard Classification of Occupations ISCO\(^4\) (International Labour Organization 2010). ISCO is also widely used in both cross-national and nationally specific survey datasets (Bergman and Joye 2005). ISCO represents an important effort to develop internationally comparable occupational codes, which facilitate cross-national comparisons in social surveys (Elias 1997). In many countries, cross-walks’ are available, that enable values of the national standardised occupational unit group scheme (e.g. SOC2010) to be translated into ISCO\(^5\). These cross-walks are usually written by researchers and/or by national statistics agencies.

\(^1\) CASCOT can be accessed here: http://www2.warwick.ac.uk/fac/soc/ier/software/cascot/.


\(^3\) Although SOC2010 is the most up to date UK scheme, some surveys may use coding guidelines that are based on previous schemes such as SOC2000, SOC90 or CO80. Further details of SOC2010 are available here: http://www.ons.gov.uk/ons/guide-method/classifications/currentstandardclassifications/soc2010/index.html.


\(^5\) Exemplary resources are available on the webpages of Professor Harry Ganzeboom: http://www.harryganzeboom.nl/ismf/ismf.htm.
4. OCCUPATION-BASED SOCIAL CLASSIFICATIONS

Organising occupational information into standardised occupational codes is only the first stage in the process. Secondary data analysts will generally not need to engage with this stage of the process when they are using existing survey resources that are well curated. The British Household Panel Survey, is a prime example of a large-scale social survey with well curated occupational data, that is readily available for secondary data analysis (Taylor, Brice et al. 2010). The second stage in the process of developing occupation-based social classifications is to convert these data from standardised lists of occupational unit groups (e.g. SOC codes) into the required socio-economic measure.

The means to convert unit group codes and employment status data into occupation-based measures is typically supplied in a listing of occupational unit groups alongside the corresponding levels of an occupation-based measure. This may take the form of a table, textual description, statistical software command file, or a matrix of data for matching (see Lambert and Bihagen 2012 for a more extended description). In order to carry out these operations the researcher will usually require some skills in the use of statistical software for data management (see Treiman 2009, Mitchell 2010).

Lambert and Bihagen (2014) report that there has been a great deal of inconsistency in how researchers have organised occupational information into occupation-based social classifications. For instance they claim that upwards of a thousand different measures based upon occupations have been used in contemporary social science. It is unsurprising that this surfeit of measures may initially seem daunting for researchers, especially those that are not social stratification specialists. It is notable that many social classifications emerge from very different sociological theories, which influence both their conceptual and empirical foundations.

It is well observed that in many secondary analyses of large and complex social surveys the analysts tend to select a single occupation-based measure. The choice of measure might be made as a result of theoretical fiat, or due to more practical operational issues, or even a mixture of both. Our methodological advice is that researchers should utilise existing measurement options whenever possible, and should avoid producing their own measures, or using existing
measures in an *ad hoc* manner, unless they have very strong and clearly justified reasons. Social survey methodologists have heralded this same dire warning on several previous occasions (Blumer 1956, Stacey 1969, Burgess 1986, Bulmer, Gibbs *et al.* 2010, Lambert and Bihagen 2014).

Our advice is based on the following premises. First, the field of stratification research is highly specialised and a great deal of theoretical thought and empirical testing has been directed towards the development of occupation-based measures. Therefore it is probable that a measure suitable for most analyses already exists. Second, the adoption of an existing measure is almost always more time efficient. Third, and most importantly in our view, existing measures have agreed and documented standards, and therefore facilitate replication. As Lambert and Bihagen (2012) assert this locates firmly within the culture of cumulative social scientific endeavour.

5. SOCIAL CLASS SCHEMES
Social Class schemes are very widely used in sociological research. Social class schemes can generally be regarded as socio-economic measures that divide the population into unequally rewarded categories (Crompton 2008). Social class schemes are not necessarily hierarchical, although an ordinal structure is often evident (Glass 1954, Carlsson 1958). There are a plethora of social class schemes, and these measures are often informed by different theoretical standpoints (see Crompton 2008). Wright (2005) distinguished between groups of social class measures which could be classified as being either Marxist, Weberian or Durkhiemian in their approach. As we have described above, more recently the measure proposed by Savage *et al.* (2013) has its genesis in the theoretical work of Bourdieu. However, Marxist and Bourdiesusian socio-economic measures are not usually readily derived from occupational information alone, and do not ordinarily feature in social survey datasets.

Readers might be interested to know that Marxist class schemes generally consider social relations of economic production as the basis upon which class groups can be defined, rather than the ‘technical’ divisions of labour (i.e. occupations) (see Wright and Perrone 1977, Wright, Costello *et al.* 1982, Wright and Martin 1987, Wright 1989, Wright and Cho 1992, Western and Wright 1994, Wright 1997, Wright 2005).
Many of the earliest published social class schemes focussed upon differences in the skill levels of occupations, and defined social categories in those terms. Skill categories were sometimes calculated based on typical qualification requirements, but their identification was also often associated with evaluations of the relative prestige or social standing of the occupation. A prominent example is the UK’s long standing ‘Registrar General’s Social Classification’ (e.g. Szreter 1984). There is evidence that skill based measures are empirically very powerful, and they remain a popular choice in social research (see Elias and McKnight 2001, Tahlin 2007).

The work of John Goldthorpe has arguably generated the most influential social class scheme in sociology and allied disciplines (Evans 1992). The Erikson-Goldthorpe-Portocarero (EGP) scheme (see Erikson, Goldthorpe et al. 1979) has become a widely used measure in social research (Evans 1992). The theoretical principles of the EGP approach led to the development of subsequent cognate schemes, including CASMIN (Erikson and Goldthorpe 1992), the UK’s National Statistics Socio-Economic Classification (Rose and Pevalin 2003) and the European Socio-Economic Classification (Rose and Harrison 2007). In this tradition, employment relations in the labour market are held to be of key importance to the allocation of individuals into social class categories (Erikson and Goldthorpe 1992). Individuals within a social class are considered to share similar ‘market situation’ (e.g. levels of income, economic security, chances for economic advancement) and ‘work situation’ (e.g. authority and control) (Goldthorpe 1980). Accordingly, those individuals within a social class are thought to hold similar life chances and often lifestyles.

In its least aggregated form the EGP scheme identifies eleven classes, although a seven class version is widely used in empirical analyses (see Table 1). Erikson and Goldthorpe (1992) recommend that researchers move between the seven, five and three class versions based on the need to balance explanatory comprehensiveness with explanatory parsimony. They state that the scheme could be extended to include more classes if there was good reason to do so (Erikson and Goldthorpe 1992). The use of varied forms of the EGP scheme is consistent with the claim that the measure is an instrument du travail rather than a definitive representation of social class groupings in the UK (Erikson and Goldthorpe 1992).
In 1994 the UK Economic and Social Research Council (ESRC) commissioned a review of social classifications. As a result of this review, the EGP approach was adopted as the basis of a new government measure of social class (Rose 1995, Rose and Pevalin 2003). Consequently, the National Statistics Socio-Economic Classification (NS-SEC) was developed, and since 2001 this occupation-based measure, described in Table 2, has been used in official statistics and government research in the UK (Office for National Statistics 2010). In congruence with the EGP approach, the NS-SEC approach comprises of aggregate groupings of individuals who are considered to share similar life chances and lifestyles. Similarly, various reduced versions of the scheme are recommended (see Table 2).
<table>
<thead>
<tr>
<th>Full Version</th>
<th>Collapsed Versions</th>
<th>Five-class version</th>
<th>Three-class version</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Higher-grade professionals, administrators and officials; managers in large industrial establishments; large proprietors</td>
<td>I+II</td>
<td>I-III</td>
</tr>
<tr>
<td>II</td>
<td>Lower-grade professionals, administrators and officials; higher-grade technicians; managers in small industrial establishments; supervisors of non-manual employees</td>
<td>III</td>
<td>Service class: professionals, administrators and managers; higher-grade technicians; supervisors of non-manual workers</td>
</tr>
<tr>
<td>IIIa</td>
<td>Routine non-manual employees, higher grade (administration and commerce)</td>
<td>IV</td>
<td>Routine non-manual workers: routine non-manual employees in administration and commerce; sales personnel; other rank-and-file service workers</td>
</tr>
<tr>
<td>IIIb</td>
<td>Routine non-manual employees, lower grade (sales and services)</td>
<td>IVb</td>
<td>Petty bourgeoisie: small properties and artisans, etc., with and without employees</td>
</tr>
<tr>
<td>Iva</td>
<td>Small proprietors, artisans, etc., with employees</td>
<td>IVa+b</td>
<td>Petty bourgeoisie: small properties and artisans, etc., with and without employees</td>
</tr>
<tr>
<td>IVb</td>
<td>Small proprietors, artisans, etc., without employees</td>
<td>IVc</td>
<td>Farmers: farmers and small holders and other self-employed workers in primary production</td>
</tr>
<tr>
<td>IVc</td>
<td>Farmers and smallholders; other self-employed workers in primary production</td>
<td>IVc+V</td>
<td>Farm workers</td>
</tr>
<tr>
<td>V</td>
<td>Lower-grade technicians; supervisors of manual workers</td>
<td>V+V</td>
<td>Skilled workers: lower-grade technicians; supervisors of manual workers, skilled manual workers</td>
</tr>
<tr>
<td>VI</td>
<td>Skilled manual workers</td>
<td>V+VI</td>
<td>Skilled workers</td>
</tr>
<tr>
<td>VII</td>
<td>Semi-skilled and unskilled manual workers (not in agriculture, etc.)</td>
<td>VIIa</td>
<td>Non-skilled workers: semi and unskilled manual workers (not in agriculture, etc.)</td>
</tr>
<tr>
<td>VIIa</td>
<td>Agricultural workers and other workers in primary production</td>
<td>VIIb</td>
<td>Agricultural labourers: agricultural and other workers in primary production</td>
</tr>
<tr>
<td>VII</td>
<td>Agricultural workers and other workers in primary production</td>
<td>VII</td>
<td>Agricultural labourers: agricultural and other workers in primary production</td>
</tr>
</tbody>
</table>

Table 1: Thorpe class Scheme (Erikson & Goldthorpe, 1992, pp. 38-39).
<table>
<thead>
<tr>
<th>Eight-class version</th>
<th>Five-class version</th>
<th>Three-class version</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Higher managerial, administrative and professional occupations</td>
<td>1 Higher managerial, administrative and professional occupations</td>
<td>1 Higher managerial, administrative and professional occupations</td>
</tr>
<tr>
<td>1.1 Large employers and higher managerial and administrative occupations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2 Higher professional occupations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Lower managerial, administrative and professional occupations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Intermediate occupations</td>
<td>2 Intermediate occupations</td>
<td>2 Intermediate occupations</td>
</tr>
<tr>
<td>4 Small employers and own account workers</td>
<td>3 Small employers and own account workers</td>
<td></td>
</tr>
<tr>
<td>5 Lower supervisory and technical occupations</td>
<td>4 Lower supervisory and technical occupations</td>
<td>3 Routine and manual occupations</td>
</tr>
<tr>
<td>6 Semi-routine occupations</td>
<td>5 Semi-routine and routine occupations</td>
<td></td>
</tr>
<tr>
<td>7 Routine occupations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Never worked and long-term unemployed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: The National Statistics Socio-economic Classification (NS-SEC)

In a recent and related exercise, a ‘European Socio-Economic Classification’ (ESeC) has been developed (Harrison and Rose 2006, Rose and Harrison 2010). This is a social class scheme, based upon the EGP model, which is designed to facilitate cross-nationally comparative research. ESeC7 comprises a nine-class categorical measure, with recommended reduced versions of five or three classes, which can be readily operationalised from data coded into the three-digit version of the ISCO occupational unit group scheme. The ‘ESeC’ scheme is specifically designed for international research, although other UK oriented versions of the EGP scheme have also been exploited in cross-nationally comparative studies (e.g. Erikson and Goldthorpe 1992, Ganzeboom 1996, Breen 2004, Blossfeld and Hofmeister 2005).

7 Full details of the ESeC scheme are available here: https://www.iser.essex.ac.uk/archives/esec/user-guide.
The EGP social class scheme and its derivatives are widely used in British sociology, and several studies have provided evidence of acceptable construct\(^8\) and criterion\(^9\) validity for this measure (e.g. Evans 1992, Evans and Mills 1998, Evans and Mills 2000). Nevertheless, these social class schemes have also been evaluated critically. Questions have been asked about the degree of within-class homogeneity in social class categories, and concerns have been raised that individuals placed within the same social class can hold very different positions within social hierarchies (Blackburn and Prandy 1997, Bergman and Joye 2005). Penn (1981), Hout and Hauser (1992) and Blackburn and Prandy (1997) have also argued that the EGP scheme’s categories downplay the key element of hierarchy in social stratification. Lastly, measures from the EGP scheme have a relatively a high number of categories, so it is sometimes suggested that they are not well suited to incorporation in multivariate statistical analyses (e.g. regression models) because having a large number of categories inhibits the estimation of interaction effects. A consequence is that this limitation might encourage researchers, *de facto*, into the less desirable practice of simplifying the measure into a more coarse-grained format.

6. SOCIAL STRATIFICATION SCALES

Having introduced categorical social class schemes we now turn our attention to occupation-based scales. The principal difference between categorical social class schemes and stratification scales is that rather than placing individuals into qualitatively distinct categories, social stratification scales place individuals at some point on a continuous or gradational one-dimensional hierarchy (Bergman and Joye 2005). This single dimension has often been labelled ‘status’, but more generally reflects ‘relative social advantage’ (Jonsson, Grusky *et al.* 2009). Scaling approaches also, technically, accommodate measurements that assign scores to occupations in more than one dimension (e.g. Levine and Spadaro 1988), but in practice scales are only regularly used in a single dimensional framework in applied research.

---

\(^8\) Construct validity is based on the assessment of whether a measure reflects the underlying construct of interest (Cronbach and Meehl 1955).

\(^9\) Criterion validity is based on the assessment of whether a measure behaves in the expected fashion, given the theory underlying the measure (Carmines and Zeller 1979)
A notable example of a social stratification scale is the Cambridge Social Interaction and Stratification Scale (CAMSIS) (Stewart, Prandy et al. 1980, Prandy 1990). This is an empirically constructed scale that is based on the theoretical idea that there is a stratification order derived from a hierarchical structure of advantage (and disadvantage) arising from the unequal distribution of social, cultural and economic resources. In the CAMSIS approach, individuals are embedded in social networks of relationships within which they engage in social, cultural, political and economic interactions. These social interactions are circumscribed by the social distance of these social actors. The idea of the centrality of ‘social space’ is not unique to the CAMSIS approach, and has a long history in the sociological literature. Chan (2010) describes a more recent project that constructed occupation-based scales by analysing social interaction patterns using a very similar approach to the CAMSIS perspective.

The CAMSIS approach is based on examining patterns of social interaction between occupations that are uncovered by examining the frequency of links between people in different occupations. These links are typically defined either by friendship or by marriage/cohabitation. CAMSIS scales are formed using statistical analysis of ‘dimensions’ within the social interaction structure (Prandy 1999). In this approach, country and time specific scales are usually calculated, and different CAMSIS scales can also be generated for men and women, and could also be generated for other important socio-demographic differences if desired (e.g. ethnic groups or geographical regions). Lambert et al. (2008) conclude that this quality of ‘specificity’ has attractive properties for a wide range of analyses where more detailed resolution might be illuminating.

Two other important stratification scales are the Standard International Occupational Prestige Scale (SIOPS) (see Ganzeboom and Treiman 1996, Ganzeboom and Treiman 2003) and the International Socio-Economic Index (ISEI) (Ganzeboom, De Graaf et al. 1992). SIOPS is devised by taking survey information on prestige ratings given by respondents to samples of jobs, and calculating averages

---

10 For example Sorokin (1927) states that ‘man’s social position is the totality of his relationships towards all groups of a population, within each of them, towards its members’.  
11 Detailed guidance for the translation of occupational codes and employment status information into CAMSIS measures can be found on the project’s website: http://www.camsis.stir.ac.uk/.
within and across societies (Treiman 1977). Treiman’s original analysis compared ratings from over sixty societies, and drew the important conclusion that variation from society to society, and across time, in the prestige allocated to occupations was minimal. This axiom is often referred to as the ‘Treiman constant’ within stratification research12 (Hout and DiPrete 2006). SIOPS provides a hierarchical ranking from the least to the most esteemed occupations according to average ratings, and scores are shown to correlate strongly with the socio-economic circumstances of individuals who hold these occupations. ISEI by contrast calculates scores for occupations based upon their average profiles in terms of the income and educational qualifications held by their incumbents (with some adjustment for age profiles). Further examples of social stratification scales include scales based only upon the average income of occupations (e.g. Sobek 1995), upon career prospects in terms of average wage growth (e.g. Bihagen and Ohls 2004), or upon job quality or desirability (e.g. Jencks, Perman et al. 1988, Mills 2007).

As continuous measures, all of the occupation-based social stratification scales lead to numeric values being attached to occupations, but the relative importance of a specific value is only meaningful in comparison with other occupations on the same scale. For example, the CAMSIS scales are usually standardised to a mean of 50 and standard deviation of 15 in each version, but the SIOPS and ISEI measures are scaled in terms of their original measurement and they typically have a mean of around 40, and a standard deviation of about 14, in a nationally representative sample. CAMSIS scales tend to be specific to particular societies whereas ISEI and SIOPS are designed to be ‘universal’ (i.e. the same scores are applicable to the same occupations across different societies). However, socio-economic index and prestige scales can also be calculated ‘specifically’ within a society, and a ‘universal’ version of the CAMSIS scale is also available (De Luca, Meraviglia et al. 2010).

12 Coxon and Jones (1978, 1979, 1979) have critiqued Treiman’s approach at length, based on analyses which focus on the cognitive issues involved when asking individuals to rank occupations. Coxon and Jones considered the types of distinctions people might draw between occupations. They experimented with tasks such as asking respondents to sort occupations into groups and asking the respondents to describe the criteria by which occupations could be ordered. They argue that these evaluative tasks should have preceded Treiman’s protocol to rank the occupations by prestige, in order to ensure the respondents had a clear basis by which to rank the occupations (Coxon and Jones 1978). Coxon and Jones (1978, 1979, 1979) also present evidence that the ranking of occupations can vary between individuals and groups which may be overlooked when producing average rankings or scales. For example, individual’s exhibit a pattern of ‘occupational egoism’ whereby they give more favourable ratings to their own occupations and occupations similar to their own (Coxon and Jones 1978).
A major attraction of all scale-based socio-economic measures is their relative parsimony in statistical analyses. Scales typically need only a single parameter (i.e. for a linear effect) or two parameters (e.g. for a curvilinear effect) to summarise their influence in a statistical modelling approach (for example a standard regression model). In many circumstances this parsimony offers a major improvement over the inclusion of a categorical social class scheme. In statistical analyses where the socio-economic measure may interact with other explanatory variables included in the analysis the parsimony that emerges from using a scale is further emphasised.

Another important attraction of using stratification scales as occupation-based measures is that their numerical functional form lends them to arithmetic standardisation strategies that may aid comparative evaluations of social change. For instance, in some analyses in stratification research it is common practice to apply mean-standardisation to scale scores within the country or year of a particular dataset. Subsequent results such as association statistics are expressed in terms of their relative influence in each context. It is well-known in stratification research that similar comparisons are much harder to conduct when categorical occupation-based measures are used. This is because there are often substantial changes in the distribution of cases to occupation-based categories over time or between countries.

7. ‘MICROCLASS’ APPROACHES
Grusky and colleagues have provided a very powerful critique of traditional social class schemes which has led to their development of the 'microclass' approach (Grusky and Sørensen 1998, Grusky and Sørensen 2001, Grusky and Weeden 2006). This novel perspective suggests that the categorical approach of class schemes is desirable, but that there are many other important divisions than are conventionally demarked in ‘big class’ schemes such as the EGP scheme. ‘Big class’ schemes generally feature a modest number of social class categories (e.g. nine in EGP). These ‘big classes’ contain a large number of occupations, for example there are eighty eight occupations (measured by SOC90) in the Semi-Routine Occupations Category (NS-SEC 6) of the UK National Statistics Socio-Economic Classification. By contrast, the
The microclass approach defines a much larger number of classes based upon institutionalised occupational divisions. Microclass schemes typically feature around eighty to one hundred different classes. Grusky and Sørensen (1998) contend that traditional social class schemes fail to represent detailed social structures within big classes, and that the social structure is not adequately represented by uni-dimensional hierarchical scales.

Details of the microclass scheme can be found on the microclass project homepage13 (see also Grusky, Jonsson et al. 2012). Theoretically, microclasses are defined by the social and/or technical institutionalisation of occupations (e.g. plumber, baker, doctor) rather than agglomerate classes (e.g. skilled manual workers, or professionals). In practice however, the empirical operationalisation of microclasses usually results in some groups being more homogeneous and more clearly institutionalised than others. Nevertheless a major attraction of the microclass approach is that it facilitates the investigation of potentially important substantive differences at the detailed occupational level that may be hidden within the large categories of ‘big’ social class schemes.

The microclass approach is still comparatively new, but it has been successfully employed to study both social mobility and educational inequalities (Jonsson, Grusky et al. 2009, Gayle and Lambert 2011). Despite the appeal of the microclass approach it has been subject to theoretical critique (see Erikson, Goldthorpe et al. 2012). Erikson et al. (2012) argue that the disaggregation of categories mean that patterns linked to microclasses cannot be clearly interpreted within the theoretical framework that is useful to a ‘big class’ measure. At a practical level the inclusion of a large multiple category explanatory variables tends to decrease parsimony in standard statistical models. In addition, to ensure adequate statistical power when working with survey data covering the large number of microclass categories, large sample sizes are required. We have stated above that compared with categorical measures, scales better facilitate analyses that include interaction effects, and we therefore note that this issue will be amplified with a categorical variable with a large number of categories.

13See: http://www.classmobility.org/.
8. SELECTING AN APPROPRIATE SOCIO-ECONOMIC MEASURE
With so many different socio-economic measures available, selecting an appropriate measure may at first appear to be a daunting task, especially for researchers who are not social stratification specialists. We argue that a sensible and defensible solution is to proceed by selecting several different operationalisations of the measures. We stress that operationalising an occupation-based measure is not necessarily a simple case of selecting one superlative measure. Therefore a good solution is to construct a number of measures and evaluate them through a ‘sensitivity analysis’.

We use the term ‘sensitivity analysis’ to describe the process of investigating the influence which small alterations to a statistical analysis, for example the use of different operationalisations of an explanatory variable in a statistical model, have on substantive results. In most circumstances a new sensitivity analysis is probably required for each new analysis. This is because the particular features of an occupation-based measure are likely to be varied in different analyses and, more importantly, cannot be predicted \textit{a priori}. The process of conducting a sensitivity analysis can seem burdensome and even uninspiring, however modern software capabilities mean that at least in principle it is now quite easy to re-run analyses using different candidate measures. We contend that undertaking sensitivity analyses is of considerable benefit to social science more generally as it increases rigour and therefore adds confidence to results. In the same way as good analysts put effort into comparing the results of different forms of statistical analysis, similar effort should be put into comparisons of measures based on alternative key social science variables.

9. SOCIO-ECONOMIC CLASSIFICATIONS AND RELATIONSHIPS WITH DEMOGRAPHIC STRUCTURES
An important consideration when analysing occupation-based socio-economic measures is their relationship with other key social science variables (e.g. age, gender, and time period). In many analyses occupation-based measures will show moderate correlations with other key variables. In some extreme cases if this is ignored there is a danger that this may result in misleading interpretations (Prandy 1986, Lambert, Tan
et al. 2008). Lambert et al. (2008) explored this issue focussing on time periods, countries and gender. They concluded that whilst temporal changes in the meaning of occupations are slight, gender differences in occupational distributions are so entrenched that they should be considered fundamental to the evaluation of occupation-based measures (see also Prandy 1986, Gabriel and Schmitz 2007).

The role of age differences in occupational classifications has received far less consideration than the effects of gender. However, several recently advocated occupation-based measures are known to have strong associations with age (see Kunst and Roskam 2010). Multidimensional measures of stratification, for example the scheme recently developed using data from the Great British Class Survey (Savage, Devine et al. 2013), also appear to be strongly linked to age differences.

In standard statistical analyses (e.g. regression models), it is reasonably straightforward to incorporate main effects for key variables (e.g. age and gender), and in many applications this will provide increased control for any correlation with occupation-based measures. Following from this we strongly advise that secondary data analysts consider including as many key variables in their analysis as appropriate, whilst being mindful of the requirement of parsimony. Similarly, analysts should consider exploring interaction effects between key variables. An alternative approach to providing increased control for relationships between key variables is to use occupation-based measures that have been derived for specific groups. Examples of such measures include the gender-specific CAMSIS scales, or the social class scheme for women’s jobs as recommended by Martin and Roberts (1984).

In the case of age, a general solution may simply be to control for age and age related interactions when including occupation-based measures in statistical models. This solution can easily be extended to incorporate nonlinear age effects. On this topic, the concept of occupational maturity is also important. An argument expressed by Goldthorpe et al. (1987) is that most adults reach a point of ‘occupational maturity’, around about the age of 35, after which it is relatively unlikely they will experience major changes in their occupational position. Analyses, particularly in the area of social mobility for example, have often been restricted to samples of adults over this age (see Erikson and Goldthorpe 1992). More recently, Tampubolon and
Savage (2012) have suggested that the appropriate age of occupational maturity may have risen slightly over time. Secondary survey data analysts using occupational-based socio-economic classifications should be cognisant of the concept of occupational maturity and consider adjusting their analyses whenever it is required.

A further issue when using occupation-based measures relates to how they can be used to represent households rather than individuals. This has been expressed in stratification research as a debate between ‘individual’, ‘dominance’ and ‘conventional’ approaches to social classification (e.g. Erikson 1984). The individual approach uses the individual’s current or last occupation. The ‘dominance’ approach measures all jobs in the household and typically assigns a measure based on the economically dominant occupation within the household. This is usually the occupation which contributes the most to household circumstances and is generally the one with the longest hours of work. The ‘conventional’ approach assigns a position on the basis of the occupation of the ‘conventional head of the household’. This is most commonly the oldest employed male living within the household.

Another alternative is simply to incorporate the necessary individual variables that relate to occupational characteristics within the household. For example within a statistical model of a child’s educational attainment information on both their mother’s and their father’s occupation might be included. In practice in some analyses such measures may be highly correlated and suitable thought must therefore be put into the precise interpretation of these effects. We advocate that secondary analysts of survey datasets should explore alternative model formulations, thoughtfully consider their effects, and then suitably document alternative results.

Finally, an enduring problem when using occupational measures is the complexity of making comparisons over time when the underlying structure of the labour market has changed. In some secondary analyses of large-scale social surveys where occupation-based measures are included as explanatory variables this will not be an issue due to the restricted time-frame of the analysis. In more specialised analyses, for example in research on inter- and intra-generational mobility, more thought will be required regarding structural changes in the labour market. We suggest that in such analyses specialist statistical approaches that are directed towards
providing increased control to help to combat this problem should be considered (see especially Erikson and Goldthorpe 1992).

10. CONCLUSION
We began this paper by stating that despite occupations being central to a broad portfolio of substantive social science research questions there is no single agreed upon way of measuring occupations. An aim of this paper has been to improve the awareness that there are a number of varied occupation-based measures which can be used in social research. We have highlighted that there are a series of issues which require thought when including occupation-based measures in any substantive analysis. Most importantly we strongly advise researchers not to develop their own measures without strong justification, or to use existing measures in an un-prescribed or *ad hoc* manner. Our clear recommendation is that researchers should always use existing occupation-based measures that have agreed upon and well documented standards.

We are careful to emphasise that we are not advocating the uncritical adoption of any one particular socio-economic measure over and above other alternatives. Rather we are advocating that researchers chose from the portfolio of existing socio-economic measures in an informed and empirically defensible way. As Lambert and Bihagen (2014) conclude, measures are often selected on theoretical grounds, which rest on the claim that a given measure captures a specific aspect of the occupational structure. Recent empirical endeavours which have attempted to provide sensitivity analyses, indicate that the theoretical orientations usually ascribed to specific occupation-based measures do not necessarily exert appreciable influence on substantive results (see Gayle and Lambert 2011, Lambert and Bihagen 2014). We conclude that rather than adopting a particular socio-economic measure on theoretical grounds, the secondary data analyst should focus more attention on the analytical benefits of competing measures. These considerations should include operational issues such as the specific form of the socio-economic measures, how to best maximise model parsimony, how to effectively specify models with suitable additional key variables, and where appropriate, how best to include relevant interaction effects.
We advocate that secondary data analysts evaluate the widest possible number of socio-economic measures and make the results of these explorations available in the form of a ‘sensitivity analysis’, for example in a data appendix. In many well curated large-scale surveys a number of alternative measures will be deposited with the data. Therefore sensitivity analyses can easily be undertaken. In datasets where alternative measures are not readily available we advocate that secondary data analysts place effort into constructing as many alternative measures as possible using detailed standardised occupational unit group codes. The secondary data analyst must always be cautious not to use these measures in an *ad hoc* manner. An obvious example of this is combining categories of a social class measure in an unstandardised or un-prescribed manner. We hope that this paper provides succinct information on the foundations of existing occupation-based measures. In addition we have attempted to provide practical advice that will make a positive contribution to how existing socio-economic measures can be better incorporated into social science analyses.
REFERENCES


