A Review of occupation-based social classifications for social survey research

Roxanne Connelly¹, Vernon Gayle² and Paul S. Lambert³

Abstract
This article 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 analysis of large-scale micro-level social science datasets. This article 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 this article does not advocate the uncritical adoption of any one particular occupation-based 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 on 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

Introduction
This article is a review of issues associated with measuring occupations and using occupation-based socio-economic classifications in sociological 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: 6–7). Within sociology, there is a long-standing recognition that in industrialised societies, occupations are the most powerful single indicator of levels of material reward, social standing and life chances (Parkin, 1971). The occupational information that is routinely provided in large-scale social surveys is a key resource for studying contemporary social life, and occupation-based indicators are central to sociological investigations.

Despite the sociological consensus that occupations are central to understanding social stratification, there is no single obvious and agreed on way of measuring occupations. Debates have been exacerbated because of the complex nature of the occupational structure in contemporary labour markets. Within sociology, a wide range of measures have been advocated. These measures are often linked with mainstream sociological theories and concepts, most notably related to ‘social class’. This article is not orientated towards an evaluation of social class measures or other measures

¹The University of Warwick, Coventry, UK
²The University of Edinburgh, Edinburgh, UK
³University of Stirling, Stirling, UK

Corresponding author:
Dr Roxanne Connelly, Department of Sociology, University of Warwick, Coventry, CV4 7AL, UK.
Email: R.Connelly@warwick.ac.uk
relating to social stratification. Nor does this article document the history or development of the measures. Rather, it is orientated towards the more prosaic, but usually overlooked, practical aspects of using occupational information collected via social surveys in sociological research.

The theorisation, measurement and operationalisation of occupation-based measures has received great attention in the specialised field of social stratification research but has received far less attention in other areas of social science despite the centrality of occupations for many empirical analyses. It is our conjecture that appropriately measuring occupations and using occupational information in sociological analyses requires thought and in-depth knowledge of these measures. The aim of this article is to document issues associated with utilising occupation-based social classifications in social survey analyses and to provide some clear prescriptions for sociologists 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 because this area is seldom highlighted. We conclude with a discussion of the intersection of other key variables such as age and gender and discuss a range of issues related to the inclusion of these measures in sociological analyses.

Rationale for occupation-based social classifications in social surveys

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 social stratification researchers, 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: 39). 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 (Goldthorpe and McKnight, 2006; Rose and Pevalin, 2003).

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.

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 et al., 1979; Erikson and Goldthorpe, 1992; Goldthorpe et al., 1980, 1987; Wright, 1997), a parallel stream of literature has claimed that ‘class as a concept is ceasing to do any useful work’ (Pahl, 1989: 710) or is indeed ‘dead’ (e.g. Clark and Seymour, 1991; Holton and Turner, 1989; Joyce, 1995; Kingston, 1994; 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, the class-based divisions peaked in industrial society and have been declining ever since. Second, 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 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: 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. Breen, 2004b; Erikson and Goldthorpe, 2010; Hochstadt, 1999; Shavit and Blossfeld, 1993; Zijdeman, 2009).

While 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, they 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; Sayer, 2005; Skeggs, 1997).
Savage et al. (2013) recently proposed a new model of social class influenced by the writings of 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. The new models of class developed by Savage et al. (2013) are 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 'lowbrow culture' but that may now be more ubiquitous. Examples might include using social networking websites, going to a gym or online 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 (NS-SEC) 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 (Bradley, 2014; Lambert and Griffiths, 2013; Mills, 2013, 2014, 2015; Payne, 2013; Rose and Harrison, 2013; 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.

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 United Kingdom, it is common for social survey data to be coded into the Office for National Statistics (2010) Standardised Occupational Classification, which are known as SOC codes. 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 et al., 1993). The Computer-Aided Structured Coding Tool (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 programme 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 that operates in a similar manner, and these and other coding software are available as offline 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. Another example of a standardised occupational code is the ISCO (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. These cross-walks are usually written by researchers and/or national statistics agencies.
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 are readily available for secondary data analysis (Taylor 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 Mitchell, 2010; Treiman, 2009).

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 who 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 made this point on several previous occasions (Bechhofer, 1969; Blumer, 1956; Bulmer et al., 2010; Burgess, 1986; Lambert and Bihagen, 2014; Stacey, 1969). Most notably, Bechhofer (1969) states, ‘researchers are advised not to add to the already existing plethora of classifications without very good reason’ (p. 118).

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.

Social class schemes

Social class schemes are very widely used in sociological research and can generally be regarded as socio-economic measures that divide the population into unequally rewarded categories (Crompton, 2008: 49). Social class schemes are not necessarily hierarchical, although an ordinal structure is often evident (Carlsson, 1958; Glass, 1954). 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 Marxist, Weberian or Durkheimian 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 Bourdieusian socio-economic measures are not usually readily derived from occupational information alone and do not ordinarily feature in social survey datasets.

Many of the earliest published social class schemes focused on 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 United Kingdom’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 Goldthorpe has arguably generated the most influential occupation-based social classification in sociology and allied disciplines (Evans, 1992) – the Erikson–Goldthorpe–Portocarero (EGP) schema (see Erikson et al., 1979). The theoretical principles of the EGP approach led to the development of subsequent cognate schemes, including CASMIN (Erikson and Goldthorpe, 1992), the United Kingdom’s NS-SEC (Rose et al., 2005; 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: 36–45). Individuals within a social class are considered to share similar ‘market situation’ (e.g. levels of income, economic security and chances for economic advancement) and ‘work situation’
(e.g. authority and control; Goldthorpe, 1980; Lockwood, 1958). Accordingly, those individuals within a social class are thought to hold similar life chances and often lifestyles.

In its least aggregated form, the EGP schema identifies 11 classes, although a 7-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 schema could be extended to include more classes if there was good reason to do so (Erikson and Goldthorpe, 1992: 46). The use of varied forms of the EGP schema is consistent with the claim that the measure is an instrument du travail rather than a definitive representation of social class groupings in the United Kingdom (Erikson and Goldthorpe, 1992: 46).

In 1994, the UK Office for National Statistics commissioned the Economic and Social Research Council (ESRC) to facilitate 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 et al., 2005; Rose and Pevalin, 2003). Consequently, the 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 United Kingdom (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 as a consequence of their employment relations. Similarly, various reduced versions of the schema are recommended (see Table 2).

In a recent and related exercise, a ‘European Socioeconomic Classification’ (EScC) has been developed (Harrison and Rose, 2006; Rose and Harrison, 2010). This is a social class schema, based on the EGP model, which is designed to facilitate cross-nationally comparative research. EScC 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. Blossfeld and Hofmeister, 2005; Breen, 2004a; Erikson and Goldthorpe, 1992; Ganzeboom, 1996).

The EGP social class schema and its derivatives are widely used in British sociology, and several studies have provided evidence of acceptable construct and criterion validity for this measure (e.g. Evans, 1992; Evans and Mills, 1998, 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 (Bergman and Joye, 2005; Blackburn and Prandy, 1997). Penn (1981), Hout and Hauser (1992) and Blackburn and Prandy (1997) have also argued that the EGP schema’s categories downplay the key element of hierarchy in social stratification. Finally, measures from the EGP schema have a relatively 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.

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). The dimension captured by occupation-based scales can usually be described as a representation of ‘relative social advantage’ (Jonsson et al., 2009). In some traditions, all occupation-based scales are referred to as ‘status’ scales, but this terminology is ambiguous, since in some interpretations, ‘status’ refers to a very specific form of social inequality (Chan and Goldthorpe, 2007). 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.

A notable example of a social stratification scale is the Cambridge Social Interaction and Stratification Scale (CAMSIS) (Prandy, 1990; Stewart et al., 1980). 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 recent, separate 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
<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Higher grade professionals, administrators and officials; managers in large industrial establishments; large proprietors</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>
</tr>
<tr>
<td>IIIa</td>
<td>Routine non-manual employees, higher grade (administration and commerce)</td>
</tr>
<tr>
<td>IIIb</td>
<td>Routine non-manual employees, lower grade (sales and services)</td>
</tr>
<tr>
<td>IVa</td>
<td>Small proprietors, artisans and so on, with employees</td>
</tr>
<tr>
<td>IVb</td>
<td>Small proprietors, artisans and so on, without employees</td>
</tr>
<tr>
<td>IVc</td>
<td>Farmers and small holders; other self-employed workers in primary production</td>
</tr>
<tr>
<td>V</td>
<td>Lower grade technicians; supervisors of manual workers</td>
</tr>
<tr>
<td>VI</td>
<td>Skilled manual workers</td>
</tr>
<tr>
<td>VIIa</td>
<td>Semi-skilled and unskilled manual workers (not in agriculture etc.)</td>
</tr>
<tr>
<td>VIIb</td>
<td>Agricultural workers and other workers in primary production</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Full version</th>
<th>Collapsed versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>I + II</td>
<td>Service class: professionals, administrators and managers; higher grade technicians; supervisors of non-manual workers</td>
</tr>
<tr>
<td>I–III</td>
<td>White-collar workers</td>
</tr>
<tr>
<td>I–III + IVa+b</td>
<td>Non-manual workers</td>
</tr>
</tbody>
</table>

Table 1. Goldthorpe class scheme (Erikson and Goldthorpe, 1992: 38–39).
friendship or by marriage or cohabitation. CAMSIS scales are formed using statistical analysis of ‘dimensions’ within the social interaction structure\(^{14}\) (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, 2003) and the International Socio-Economic Index (ISEI; Ganzeboom et al., 1992). SIOPS is devised by taking survey information on prestige ratings given by respondents to samples of jobs and calculating averages within and across societies (Treiman, 1977). Treiman’s original analysis compared ratings from over 60 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 research (Hout and DiPrete, 2006). Coxon and Jones (1978, 1979a, 1979b) 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).

The 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. The ISEI by contrast calculates scores for occupations based on 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 on the average income of occupations (e.g. Sobek, 1995), career prospects in terms of average wage growth (e.g. Bihagen and Ohls, 2004) or job quality or desirability (e.g. Jencks et al., 1988; Mills, 2007).

Because they are 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 a 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 et al., 2010).

<table>
<thead>
<tr>
<th>Table 2. National Statistics Socio-Economic Classification (NS-SEC).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eight-class version</strong></td>
</tr>
<tr>
<td>1. Higher managerial, administrative and professional occupations</td>
</tr>
<tr>
<td>1.1. Large employers and higher managerial and administrative occupations</td>
</tr>
<tr>
<td>2. Higher professional occupations</td>
</tr>
<tr>
<td>5. Routine occupations</td>
</tr>
</tbody>
</table>
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 (e.g. a standard regression model). In many circumstances, this parsimony offers a major improvement over the inclusion of a categorical social class schema. 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.

'Microclass' approaches

Grusky and colleagues have provided a critique of traditional social class schemes which has led to their development of the ‘microclass’ approach (Grusky and Sørensen, 2001, 1998; 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’ approaches such as the EGP schema. ‘Big class’ schemas 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 88 occupations measured by SOC90 and 60 occupations measured by SOC2000 in the semi-routine occupations category (NS-SEC 6) of the UK NS-SEC. By contrast, the microclass approach defines a much larger number of classes based on institutionalised occupational divisions. Microclass approaches typically feature around 80–100 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 unidimensional hierarchical scales.

Details of the microclass scheme can be found in Grusky and Sørensen (2001). Theoretically, microclasses are defined by the social and/or technical institutionalisation of occupations (e.g. plumber, baker and doctor) rather than aggregate 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 (Gayle and Lambert, 2011; Jonsson et al., 2009). Despite the appeal of the microclass approach, it has been subject to theoretical critiques (see Erikson 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 variable 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.

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 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.

**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 (Lambert et al., 2008; Prandy, 1986). Lambert et al. (2008) explored this issue focussing on time periods, countries and gender. They concluded that while 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 Gabriel and Schmitz, 2007; Prandy, 1986).

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 approach recently developed using data from the Great British Class Survey (Savage et al., 2013), also appear to be strongly linked to age differences.

A suitable analytical response may be to include measures of age and gender as control variables within analysis, also investigating possible interaction effects between the age and gender measures and the occupational classification. In practice, these controls may reduce the parsimony of the model or introduce further challenges to interpretation, so researchers may have to make a careful judgement on the best way to address correlations between occupation-based classifications and age or gender (or other comparable factors).

An alternative approach to providing increased control for these relationships 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 recommended by Martin and Roberts (1984). With regard to age patterns, 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 (p. 51). 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 occupation-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).
Conclusion

We began this article by stating that despite occupations being central to a broad portfolio of substantive social science research questions, there is no single agreed on way of measuring occupations. An aim of this article 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 on 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, 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, 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.

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 occupation-based socio-economic 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 article 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.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

Notes

1. Atkinson and Rosenlund (2014) have also recently developed a social class model based on Bourdieusian theory.
2. CASCOT can be accessed here: http://www2.warwick.ac.uk/fac/soc/ier/software/cascot/.
4. 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://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume1-structure-and-descriptions-of-unit-groups/index.html
7. For a summary of several occupation-based measures, see Lambert and Bihagen (2012: 17–19).
8. Readers might be interested to know that Marxist class schemes generally consider social relations of economic production as the basis on which class groups can be defined, rather than the ‘technical’ divisions of labour (i.e. occupations; see Wright, 1989, 1997, 2005; Wright and Cho, 1992; Wright et al., 1982; Wright and Martin, 1987; Wright and Perrone, 1977; Western and Wright, 1994).
10. Full details of the ESeC scheme are available here: https://www.iser.essex.ac.uk/archives/esec/user-guide.
11. Construct validity is based on the assessment of whether a measure reflects the underlying construct of interest (Cronbach and Meehl, 1955).
12. 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).
13. 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’ (p. 6).
14. 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/.

15. It has been argued that occupation-based social classifications measure different aspects of social stratification dependent on their conceptual basis (see Bukodi et al., 2011; Chan and Goldthorpe, 2007; Goldthorpe, 2010; Marshall et al., 2005; Rose and Harrison, 2010). If this is the case, a sensitivity analysis will provide an indication of the robustness of social survey analyses across different occupation-based measures. In recent work, Lambert and Bihagen (2014) have demonstrated that, generally, occupation-based social classifications are highly correlated regardless of their theoretical or conceptual basis.

References


Rose D and Harrison E (2013) Letters: Little solidarity over the question of social class. The Guardian. Available at: http://www.theguardian.com/society/2013/apr/05/solidarity-question-social-class


Wright EO (1989) *The Debate on Classes*. Verso, London:


**Author biographies**

Roxanne Connelly is an Assistant Professor of Sociology at the University of Warwick. Her main research interests include social stratification, the sociology of education and the analysis of complex social science datasets.

Vernon Gayle is a Professor of Sociology and Social Statistics at the University of Edinburgh. His main research interests include social stratification, the sociology of youth and the sociology of education. He also has research interests in migration, subjective wellbeing, populations, fertility, digital social research and the sociology of sport. His methodological interests include the analysis of large-scale and complex social science datasets, statistical modelling, longitudinal data, administrative social science data, missing data methods and social networks.

Paul S. Lambert is a Professor of Sociology at the University of Stirling. He is engaged in research on social stratification, social distance, and statistical methodology, and teaches courses on social inequality and social research methods.