Can class and status really be disentangled?

Professor Erik Bihagen (corresponding author).

Swedish Institute for Social Research (SOFI)

Stockholm University

erik.bihagen@sofi.su.se

Professor Paul Lambert

Faculty of Social Sciences

University of Stirling

paul.lambert@stirling.ac.uk

Can class and status really be disentangled?

Abstract:

Tak Wing Chan and John Goldthorpe (CG) have argued that it makes theoretical and empirical sense to use indicators of both class and status in analyses of cultural consumption, political attitudes and labor market outcomes in order to disentangle different mechanisms of stratification. However, we argue that class and status measured by occupationally based stratification variables are too strongly mutually associated for this to be a reliable approach. We provide empirical analyses, using secondary survey data from the UK's BHPS, that indicate that the measures of class and status largely tap the same form of stratification. It turns out that class accounts for around 75 percent and more of the variation in status and even more if excluding outliers. Moreover, class and status are similarly associated with earnings, have similar experience-earnings curves, and patterns in relevant model residuals are not consistent with the theoretical differences between class and status. In conclusion we point out alternative and more accurate usages of Weber's concepts of status and also suggest a more realistic and pragmatic view on occupationally based stratification variables.

Key words: stratification, social class, social status, Max Weber, employment relationship theory

1. Introduction

In a number of publications Tak Wing Chan and John Goldthorpe (CG) have argued that it makes theoretical sense to use indicators of both class and status in analyses of cultural consumption, political attitudes and labour market outcomes in order to disentangle different mechanisms of stratification (Chan, Goldthorpe 2004, 2007a, 2007b, 2007c, 2007d, 2007e, 2007f, 2010). The idea is that Max Weber's classical distinction between class and status can straightforwardly be implemented in empirical research with the use of different occupationally based measures. Hence, consumption and lifestyle-oriented outcomes are expected to be more associated with 'status' while labour market related outcomes should be more associated with 'class'. Following this literature several empirical studies have sought to disentangle social mechanisms in the manner suggested by CG (Alderson, Junisbai et al. 2007, Bukodi 2007, Bukodi, Goldthorpe 2012, Chan, Birkelund et al. 2011, Erikson 2016, Torche 2007, Katz-Gerro, Raz et al. 2007, Torssander, Erikson 2009).

However, in our view the strong association between the occupationally based variables of class and status, that CG use, indicate that the two measures largely tap the same dimension of stratification, and that the differences between them cannot be presumed to neatly disentangle the differences between the theoretical concepts of 'class' and 'status'. We start by describing CG's standpoint and then by pointing out some weak parts of the argument. We then present analyses of data from the UK's British Household Panel Study (BHPS, see University of Essex, Institute for Social and Economic Research. 2010). We explore the correlation between measures of class and status, both measures' relationships to earnings and to experience-earnings curves, and the relative differences between how the measures perform in more extensively parameterized statistical models. In general, our findings are that the empirical qualities of class and status measures are not successful in disentangling different aspects of the theoretical concepts of class and status. After the results, in the concluding section, we suggest another concept of status that is closer to Weber's original notion of status. We also suggest that stratification measures should be viewed pragmatically because they are unlikely to provide robust tools for disentangling different theoretical concepts.

2. Bringing class and status back in – the argument of CG

For decades, Goldthorpe and colleagues have argued that there is a clear definition of the concept of class, and that this definition is well captured in empirical class schemes such as the Erikson, Goldthorpe and Portocarero or 'EGP' (Erikson, Goldthorpe 1992, Goldthorpe 1996) scheme, and contemporary 'SEC' schemes (e.g. NS-SeC, the UK's National Statistics Socio-economic Classification, and ESeC, the European Socio-economic Classification) that have been derived according to its logic (e.g. Rose, Harrison 2010, Rose, Pevalin 2003). For employees, class is suggested to be related to employment relationships, where it is rational for employers to give long term advantages to employees in occupations with relatively complex work tasks (since such employees are difficult to monitor and the asset specificity of their work tasks makes them costly to replace). Hence, those with such a 'service relationship' have better job security and stronger earnings growth than those with less complex work tasks with a 'labour contract'. More recently, Chan and Goldthorpe (2007a, p. 514) have argued that the concept of status captures other processes of social stratification and should also be used in research. Social status is a "structure of relations of perceived, and in some degree accepted, social superiority, equality, and inferiority among individuals" . Status is described as being expressed by "'commensality'": "who eats with whom, who sleeps with whom: and further in lifestyles…" (Chan, Goldthorpe 2010, p. 12). Also, it is described as the *preferences* for whom to interact with (Chan 2010, p. 30).

When it comes to the derivation of the measures, EGP has a rather complex genealogy (Tåhlin 2007). It is largely constructed *a priori*, by an assessment of employment conditions of occupations and combinations of occupations and employment statuses. On the basis of such assessments and with the aid of empirical data on employment relationships (in the case of SEC: Rose, Harrison 2010, Rose, Pevalin 2003), occupations and employment statuses are sorted into classes. Status, on the other hand, is constructed *a posteriori*, i.e. it is empirically derived from patterns of friendship or marriage data. The empirical derivation (see Chan 2010, chpt 2; Chan, Goldthorpe 2004) is similar to an approach that has been used for other measures of 'social interaction distance' ('SID') and hereafter we use the generic label of 'SID' to refer both to the CG status measure and to other similar SID scales (cf. Bakker 1993, Laumann, Guttman 1966, Levine, Spadaro 1988, Prandy 1990, Stewart, Prandy et al. 1980). Briefly, SID analyses are based on cross tabulations of own occupation and occupation of the closest friend (or occupations of the spouse). Dimensional structures are then calculated which predict the pattern of cases in the tables (using tools such as multidimensional scaling,

correspondence analysis, or loglinear association models). Results can identify multiple dimensions to social interaction patterns, but it is usually the case that a single most important dimension can be identified (Chan 2010: chpt 2), which in CG's analysis is interpreted as an indicator of social status, realized through 'status scores' that are assigned to each occupation. As a sign of the validity of the status scale CG emphasize the historical 'echoes' of the scale where non-manual occupations (those 'working with symbols') have historically been regarded as having higher status than manual occupations (those 'working with things'), and professionals have been regarded as higher than managers (Chan, Goldthorpe 2004), which is also what the scale indicates. According to CG such findings strengthen the idea that the scale does not only represent homophily – i.e. the tendency for people to make friends with others like themselves - but also status.

CG argue that their distinction between empirical measures of class and status offers an opportunity for the 'multidimensional' analysis of social stratification. Whilst CG do not maintain that status and class are orthogonal, they argue that there are a range of circumstances in which the distinctive forces of each mechanism can be empirically disentangled, leading to new insight about social inequalities. Moreover, CG provide a series of studies that identify differences in empirical outcomes between measures of class and status. In particular, the measure of class appears to be a stronger empirical predictor of labor market outcomes, and the measure of status is more associated with cultural consumption, while both are important for different aspects of political attitudes (e.g. Chan, Goldthorpe 2007b, 2007c, 2007e). From such evidence, the conclusion is drawn that social processes can be usefully understood by evaluating whether they are more strongly linked to the measure of class or that of status – if the former, then it must be class processes, related to employment relations, which are more central to differences in the mechanism examined, but if the latter, then the social process is more related to the social relations of honour and recognition which embody status. We are skeptical of this argument and will continue by discussing some weaknesses in this line of reasoning.

3. Weak points of the CG argument

We do not claim that Weber was wrong in suggesting that both class and status are potentially interesting concepts for stratification research, but we believe that the CG approach to distinguishing those concepts empirically is problematic. We argue that CG's measures largely tap the same underlying dimension of stratification, and that the small deviations

between positions in terms of class and of status do not correspond consistently to the theoretical differences between class and status.

3.1.   Advantages at the occupational level tend to coincide too much

As described above CG's measure of class is an aggregation of occupations (and employment statuses) to a number of classes while status is a scale of occupations. Occupations in turn – as measured by contemporary occupational classifications - are based on the type of work performed including distinctions between industries, between authority positions, and between the skills required for the occupation. Earlier research shows that stratification measures in general are strongly correlated (Lambert, Bihagen 2014). This suggest that different advantages at the occupational level tend to go hand in hand; e.g. occupations that are highly skilled tend to have highly educated incumbents, be highly paid, have good earnings growth and be prestigious. Therefore stratification measures that are based on empirical outcomes such as occupations' income and typical educational levels (like ISEI, see Ganzeboom, Treiman 1996), earnings growth (Bihagen, Ohls 2007), or popular views on their prestige (Treiman 1977) tend to correlate strongly. Previous research also shows that class (as being measured by EGP/SEC) and status (measures by SID-scales) are strongly correlated with each other as well as with other stratification measures (Lambert, Bihagen 2014).

Thus, our first argument is that advantages at the occupational level tend to strongly coincide to such a degree that any measure that aims to tap occupational advantages in one specific way will unintentionally tap other occupational advantages. As a consequence, if two stratification measures are included as independent variables in the same regression for predicting an outcome there is an obvious risk of multicollinearity and of artefactual differences. The latter might reflect differences in functional form between measures, or differences in measurement errors (such as in ambiguities in the original coding of occupational titles, or processing errors when linking occupations to positions on the occupation-based measure - cf. Ganzeboom 2009). Hence, differences in net associations of the variables should be interpreted cautiously.

3.2. The measures of class and status are not neatly defined in line with theories

Our second argument concerns whether the two measures of class and status are really neatly in line with theoretical definitions of the concepts. We argue that the links between the theoretical definitions and the measures are actually less straightforward than they appear to be in the writings of CG. The EGP schema was from the beginning aimed at measuring the work and market situation of occupations and was at least partly operationalized out of occupations' scores on the Hope and Goldthorpe scale of social standing (see Tåhlin 2007). However, more important than the genealogy of the EGP and SEC schemas, Tåhlin (2007) shows empirical evidence that indicates that EGP/SEC schemas are empirically more associated with the general skills required at work, rather than asset specificity as is theorized. Hence, it is clear both that class does not only measure what it is assumed to measure, and that class is strongly associated with general skills.

It is also not clear that we can assume *a priori* what status or SID scales measure. On empirical grounds we can observe a largely universal dimension (over time and across countries) that can be derived out of different interaction patterns (marriage or friendship), and is consistently associated with a number of empirical outcomes as well as with other stratification scales (Prandy 1999, Prandy, Lambert 2003). The same SID structure that has been interpreted by CG as status has been portrayed by others as an underlying structure of social stratification (Stewart, Prandy et al. 1980). Alternatively, since friendship patterns as well as marriage patterns are relatively strongly associated with formal education (Blossfeld 2009, Skopek, Schulz et al. 2011), the dimension may largely reflect the skill levels of occupations. We believe that both these alternative interpretations, as a general dimension of stratification or as an indicator of skill levels, are more plausible than CG's suggestion that it reflects a general preference order of occupations.

Despite alternative interpretations of SID in the literature, CG's standpoint has been described as a belief that by SID "the relative social status attaching to occupations is effectively captured, but *only* social status and not other, more 'material' aspects of the occupational hierarchy, as is claimed in the case of the Cambridge scale" (Bukodi, Dex et al. 2011, p629). It is evident that SIDs do capture a hierarchical element of interaction patterns but to us it seems implausible that this hierarchical element is *pure* status, or any other *pure* shared quality of occupations. To reiterate, advantages at the occupational level tend to go hand in

hand and status can hardly be orthogonal to such advantages (i.e. be unrelated to skill and income levels or even prestige). Hence, if we capture status by SID we most likely capture other associates of status at the occupational level, e.g. the typical education in occupations. Moreover, we believe that it is opportunities to interact that drive the hierarchical element of interaction patterns. At the individual level interaction patterns may substantially reflect the outcome of preferences (individuals interact with whom they like), but at the aggregated occupational level, we believe it is more likely that they are driven by opportunities to interact; the relative frequency by which incumbents of different occupations meet people in other occupations, rather than preferences for occupations (cf. Blossfeld 2009, Kalmijn 1998, McPherson, Smith-Lovin et al. 2001, Schwartz 2013, Skopek, Schulz et al. 2011). Indeed, it is worth considering that friendship and partner choice often precede the latest occupational transition, and educational attainment typically precedes the 'occupational career'. Hence, it is possible that interaction patterns between incumbents of different occupations are to some extent epiphenomenal and that they are largely driven by education or other properties that precede the transition to the current occupation, for instance social origin.

Our argument is that the two measures used by Chan and Goldthorpe are not distinct measures of the concepts they are supposed to measure. EGP/SEC may not be clearly defined in line with employment relationship theory (for employees) and SID may capture socioeconomic advantages in general. Alternatively, both measures may tap skills associated with occupations. Hence, if the measures are not distinct, and if they even unintentionally tap the same underlying dimension, differences in performance of the two measures should not be interpreted in terms of class and status respectively.

3.3. Other differences than the theoretical difference between class and status can explain the empirical differences between SID and class measures

A fair objection to our arguments is that the results of CG make sense intuitively: labour market outcomes are more strongly associated with 'class' and cultural consumption with 'status'. Seemingly there is a pattern in the results that cannot easily be reduced to only multicollinearity problems and measurement errors. Our third argument is then that there are several other possible reasons for such deviations in the net associations of SID and EGP/SEC than that they reflect theoretically meaningful distinctions.

First, an obvious reason for deviations between EGP/SEC and SID is the different functional form of the measures: a scale for SID and categories for EGP/SEC. For example, if elite

cultural consumption is more common at very advantaged positions, the aggregated measure of class (EGP/SEC) may not be a good tool for detecting such associations because of its functional form rather than because of the logic of its derivation. CG do test different functional forms of the SID measure and argue that this is not an important influence on the difference between 'class' and 'status' (Chan, Goldthorpe 2007e). Nonetheless, earlier research suggests that the 'functional form' of an occupation-based measure (for instance, whether an EGP class scheme is represented in its 11-category, 7-category or 3-category version) are often more influential to the performance of the measure than is the theoretical alignment of the measure (e.g. Lambert, Bihagen 2014).

A second source of systematic differences in outcomes between SID and EGP/SEC comes from the fact that the measure of class is not only based on occupations but also on employment status. Hence, for EGP/SEC more aspects of labour market relations are built into the measure, such as if the individual is self-employed or not. If employment status matters for labor market outcomes, which is plausible (cf. Chan, Goldthorpe 2007e), this is potentially picked up by EGP/SEC but not the SID measures that CG use (SIDs may also be based on both occupations and employment status, see Prandy, Lambert (2003), but CG choose not to do so). In EGP/SEC schemas, the self-employed with employees, as well as employees with management/supervisory tasks, are generally 'upgraded' in class compared to employees in the corresponding occupations. Hence, the same may be true if employment status is used in the derivation of SID. Alternatively, a SID measure that does not use employment status in its derivation might even push employment status differences into an orthogonal relationship to the first SID dimension.[1] One possibility is then that SID and EGP/SEC both basically capture the same form of stratification (general advantages or skills of occupations), but because EGP/SEC also single out a self-employed category (that will have a considerable variation in SID scores), the bulk of the empirical difference between a SID and EGP/SEC measure will come down to differences of employment status. Though employment status is arguably part of a theory of class, it could, equally, also be part of a theory of status, but it cannot influence the status measure if it is 'forced' out of the picture.

---

[1] This could happen if social interaction patterns are to some extent linked to employment status, but the empirical relationship cross-cuts occupational patterns so substantially that a dimensional reduction technique subsequently isolates employment status as a separate dimension. One example is evident in Levine and Spadaro (1988), who find a second dimension associated with employment contract arrangements – which they label 'class'.

A third possibility is that although stratification measures tend to be similar, due to a large overlap of advantages in occupations, they are still associated with some other systematic deviances in their qualities. In an interesting comparison of stratification measures, Warren, Sheridan et al. (1998) argue that socio economic indexes (SEI) typically have two main components: they measure 'occupational education' and 'occupational earnings', and different versions of SEI give different weights to these two components. Put differently, some occupations have higher skills than earnings and others higher earnings than skills. Since women have on average higher occupational education but lower occupational earnings than men, analyses of gender differences are strongly affected by the choice of SEI (Warren, Sheridan et al. 1998). Neither EGP/SEC nor SID are socio economic indexes in the sense that information on education and earnings have been used in the derivation of the measures, but still the measures may have 'biases' in different directions. One possibility is that SID leans towards occupational education and EGP/SEC to occupational earnings, but also that such 'biases' are linked with gender composition and industry. An inspection of EGP/SEC schemas (in this case ESeC) indicate several female dominated occupations with high skill levels from services industries that are located in service class II instead of service I (the most advantaged class in the ESeC classification).[2] The same occupations seem to be located relatively high on SID scales compared to class (see comparisons in Chan 2010).

Theoretically speaking the allocations in EGP/SEC of those occupations could potentially be in line with employment relationship theory if for example teaching professionals have less complex work tasks (defined in terms of 'monitoring problems' and 'asset specificity', see Rose, Harrison 2010) than other high skilled occupations. However, from the viewpoint that EGP/SEC mainly captures formal skills (Tåhlin 2007), EGP/SEC seems to downgrade female dominated occupations. Such deviations in terms of skills may in fact imply more predictive power of the EGP/SEC for at least earnings, as high skilled female dominated occupations may have lower earnings than occupations included in Service I. At the same time SID's potentially closer relationship with skills could make SID a better predictor of cultural consumption than EGP/SEC. Earlier research indicates that cultural consumption is relatively strongly associated with education (Chan, Goldthorpe 2007f), and hence a stratification measure that is more in line with skills may be a better predictor of cultural consumption if similar mechanisms are at work. It has also been suggested, although only partly empirically

---

[2] In ISCO group 2 at the one-digit level: 223, 232, 233, 234, 243, 244, 245, 246, 247: e.g. teachers and nurses.

supported, that cultural production occupations (which include some female dominated occupations that are 'downgraded' by EGP/SEC) tend to lean towards a cultural lifestyle (Bihagen, Katz-Gerro 2000, Christin 2012, Collins 1988). Hence, if SIDs put such occupations higher than SEC, this may explain why SID performs better regarding cultural outcomes.

To sum up, our basic argument is that 'class' and 'status', measured by EGP/SEC and SID respectively, to a large degree capture the same underlying dimension. Although there may be some systematic deviances between the measures that account for the differences in outcomes, we suspect that such deviances are not consistently related to the theoretical concepts of class and status.

4. Empirical strategy

First we provide a detailed description of the association between SID (status) and the most well-known SEC - EGP (class). We also analyze the residuals (from regressing EGP by SID) and pinpoint the occupational categories that contribute most to the deviance.

Second we study associations between SID and EGP on the one hand, and indicators of employment relationships on the other hand. According to employment relationship theory (ER), class is an indicator of employment relationships for employees and EGP should therefore have stronger associations with such indicators than SID for employees. There is no corresponding theory for status and, hence, no corresponding empirical indicators to use.

Third, ER-theory suggests that classes of employees have different work experience-earnings profiles due to their differences in employment relationships. Thus, classes in terms of EGP should have more distinct patterns of earnings growth by work experience than groups of SID. Moreover, EGP should be a better predictor of earnings than SID because it captures employment relationship better.

Following our arguments above our expectations are that;

1. Measures based upon SID and EGP are highly associated.
2. SID and EGP are similarly associated with employment relationship indicators.
3. SID and EGP have similar work experience – earnings profiles, and differences in their performance reflect factors other than associations with employment relationships.

If our results are in line with this we would conclude that class and status cannot be fruitfully distinguished with the CG measures.

5. Data and variables

We used the British Household Panel Study (BHPS) adult interview data (University of Essex, Institute for Social and Economic Research, 2010) as the coverage of variables is good. The BHPS is a longitudinal panel study that commenced with a sample of 10264 adults who were resident in 5505 households in Britain in 1991. In subsequent years the same respondents were re-interviewed when available, and further individuals joined the survey in later years for various reasons. Respondents from the BHPS are still being interviewed, since 2010 as part of the UK's 'Understanding Society' survey (University of Essex, Institute for Social and Economic Research, 2017), but we limit our analysis to data between 1991 and 2008, prior to a major reform in the organization of the survey. To maximize the power of the data, we often used 'pooled' panel data whereby more than one record comes from the same individual. Footnotes to some tables and graphs indicate, when relevant, if we used statistical adjustments for non-independence of records for the same people; for most outputs, sensitivity analysis showed that adjustments were not consequential and in most cases all available records are used without further adjustment. Some variables, i.e. 'occupational scores', used in analysis are based upon additional analyses using the British Labour Force Survey (LFS) and the European Social Survey (ESS). The on-line appendix gives supplementary details on the variables used but key features of the measures are as follows:

Occupations. Occupation-based social classifications and occupational-level explanatory variables are operationalized at the level of the UK's Standard Occupational Classification 1990 (OPCS 1990). Details of the procedures used are included in the on-line appendix. Class. The EGP-schema is used (Erikson, Goldthorpe 1992) in its complete version in the BHPS (11 categories, based upon an algorithm written by Elias, Halstead et al. (1993). Our analysis is often restricted to the subpopulation of employees only; when this happens, our analysis applies to the 8 remaining EGP categories. In some instances we also use reduced versions of class as are commonly used in studies that exploit EGP/SEC measures. These include a six category version recommended in the SEC literature (Rose, Harrison 2010: 21) which encompasses five categories of employees; a trichotomy for employees that corresponds to classes with a service relationship, with a labour contract and with mixed

contracts (Goldthorpe 2000); and a dichotomy for employees that contrasts the 'salariat' classes with all others.

Social interaction distance scales (SID). We used two different scales. The first one is the 'status' scale for Britain generated by CG based upon friendship patterns (status_CG, see Chan and Goldthorpe 2004). It is worth mentioning that this scale only differs between 31 occupational groups. In order to use a SID scale that is based on more occupations and occupation-by-employment status combinations (459 different categories for the data set used here), we also used the UK CAMSIS scale for Britain (1991 version), based upon marriage patternsThe CAMSIS scales are not initially presented as a measure of 'status', but are constructed in the same way as the status scales in the CG project.

Employment status. We used a measure of employment status that distinguishes between employee, supervisory, managerial and employer roles, and also takes account of the size of the organisations involved.

Sex. Most analyses are applied to cross-gender populations with dummy variables indicating female respondents, but the earnings regressions are conducted mainly for men.

Education. For respondents' education the analyses use the derived variable for the 'highest educational qualification held'. Our analysis used a recoded version with 6 categories: no qualifications; lower vocational, other, and lower school level qualifications; medium school level qualifications; higher school level qualifications; post-school academic qualification below degree level; qualifications at degree level and above.

Industry. The industry measure has the following values: private sector service, public sector service, other industries.

Work experience. In the earnings regressions we use a proxy indicator for work experience derived as age in years minus school leaving age, and square this term to allow for a curvilinear relationship (Mincer 1974). We also include interaction terms between these

variables and EGP or SID groups, to assess if earnings growth varies by 'class' or 'status' in the regressions (Goldthorpe, McKnight 2006).[3]

Employment relationship indicators. Employment relationship theory refers only to employees. Hence, indicators of employment relationships will only be used for separate analyses of employees.

The first and arguably most important aspect of employment relationships to theories of social class is the concept of the 'service relationship'. Our intention was, first, to use questions about the service relationship for employees asked in the 1996/1997 version of the UK Labour Force Survey (LFS), which were used when NSeC was constructed (Rose, Harrison 2010: p59-60). However, these particular questions from the LFS are not available to secondary analysts. Instead we used BHPS data which has coverage of indicators of many but not all of the elements of the service relationship that are highlighted by Rose and Pevalin (2003). Using BHPS data, we extracted a service relationship score based on whether individual respondents reported that their job featured standard pay rises, promotion opportunities, payment by salary rather than hourly rates, opportunities to work flexibly, opportunities to control working hours, a defined career path, and whether the job was held permanently.

We also derived several other occupational scores on complexity of work tasks from the European Social Survey. These measures aim to tap the preconditions of differences in employment relationships: asset specificity and monitoring problems. We then assume that occupations are homogeneous in terms of such scores across a wide range of countries, which may or not may be realistic. Since EGP and other occupationally based stratification measures (such as ISEI and SIOPS) are used across European countries this assumption is frequently made implicitly. We used questions on how much time it takes to learn the job if the employee has the right qualifications and if the job requires learning new things. These questions are aimed to measure asset specificity and monitoring problems respectively as metric variables (Bihagen, Nermo et al. 2010).

---

[3]The longitudinal nature of the BHPS means that a more sophisticated measure of work experience could be computed for most (but not all) respondents, by drawing upon life history data and subtracting information on durations post-school that are not in employment. However it is common practice to use an age-based measure in this way and our expectation is that this provides a reasonable proxy.

Education/skills of occupations. From ESS we also construct a measure of median actual length of schooling for different occupational units, and of median skill requirements. The latter is based on a question about how many years of schooling is required for doing the respondent's job well. Since educational systems vary across countries these two measures are standardized within countries in accordance with the percentiles (the ESS-based measures above are standardized according to the percentiles for all countries). Hence, they can be thought of as relative measures to the distribution of education/skills in each country. As we use the median value for each occupational unit (as for the employment relationship scores above) the difference between these two measures is not large (mean values would be more affected by under- and over-education within occupations), and the correlation at the occupational level is in the magnitude of .90.

Sex composition of occupations. The percent of women in occupational units is derived from the ESS (Waves 1 to 5).

Occupational mean earnings. Using BHPS data we calculate occupational level mean earnings based upon the empirical Bayes residuals from a random effects null model.

Earnings. The BHPS derived variable for monthly pre-tax earnings was used in earnings analyses (variable 'Wfimnl'). This provides an estimated monthly earnings for employees, which we analysed as a log function. Our earnings analyses are mainly applied to adult men for whom in the UK part-time work is very rare; accordingly monthly earnings provides a compelling indicator of an occupation's average earnings without the need for further standardization by the hours of work undertaken.

In Table 1 descriptive statistics of the variables can be found. Some of the analyses below are reported separately for employees and, hence, so are the descriptives.

--Table 1 about here -

6. Results

In Figure 1 we report the explained variance from regressing EGP (class) on SID measures when the total population is used (i.e. when 0 percent of the population is excluded on the x-axis), and successively when those individuals/occupations with the largest residuals (predicted values – observed values) are excluded. As is already well-known, EGP accounts for the lion's share of the variation in SID (almost 75% of the variation). What is new here is that this part increases rather quickly when a relatively small proportion of the population is excluded. Hence, the lack of fit between measures is largely driven by a smaller proportion of the sample. If three percent of the population is excluded EGP explains as much as 80 % of the variation in status_ CG. When 10% of the population is excluded class explains 90% of the variation in status_CG. The explained variance for the CAMSIS scale is somewhat lower but there is still a steep increase in the beginning when excluding outliers. The figure also shows that the corresponding explained variance is higher for employees only, and is higher still when we focus only on male employees. However, an interaction with gender is revealed, insofar as the correlation between measures is less when we focus upon female employees only. This suggests that mechanisms linked to gender may be relatively important aspects of the empirical difference between measures of SID and EGP. Systematic differences in how male and female occupational distributions are treated by SID and EGP measures might arise for operational reasons (such as coarseness in coding for small numbers of populous, female-dominated jobs). However there is not an obvious theoretical explanation, in terms of the difference between concepts of class and status, that would explain the gender interaction in the SID-EGP correlation. Our over-all conclusion from Figure 1 is in line with our expectation: measures based upon SID and EGP are highly associated.

--Figure 1 about here -

In the on-line appendix the mean values of SID by class are shown (Figure A1). To make values comparable the values are presented as standardized coefficients. As would be expected, Service I occupations have the highest SID scores and there is a clear class gradient in SID-scores. It is interesting to see that the class categories for which employment status are of crucial importance for the allocation (self-employed IVa in particular), have higher values on CAMSIS than status_CG. This is expected since CAMSIS is based on a combination of occupation and employment status, and we would argue that the lower values for status_CG

for these categories is an artefact of deriving status_CG without distinguishing between employment statuses.

In Table 2 we list the occupational categories that contribute most to the residuals (when predicting SID scores by EGP) as a way to explore deviations between SIDs and EGP. In total eleven occupations are presented that contribute to 21 and 27 percent of the residuals in total for CAMSIS and status_CG respectively. Some patterns are discernible from these outliers. Managerial positions (with exceptions) tend to have negative residuals, which means that they have lower SID-scores than the average of the EGP category that they are placed in. Teachers and care-workers tend to have higher SID than EGP. A clear deviance between the two SIDs concerns sales assistants that have opposite residuals. It is also worth mentioning that only a few of the occupations with the highest residuals are the same for the two SIDs. In all, although there are some similarities across SID measures in their residuals with EGP, there are many deviances. This is noteworthy because there is no theoretical reason why these two different measures of 'status' should have different empirical qualities in their relationship to class.[4]

--Table 2 about here--

Theoretically speaking from the CG point of view we would expect class to be more associated with ER-related variables than status. In Table 3 the correlation (r) from regressing EGP and SID on ER-indicators for employees are presented (first three rows). It turns out that SIDs are about as good as EGP in predicting ER which is in line with our expectations but goes contrary to what would be expected from CG's ideas.

--Table 3 about here--

We elaborate Table 3 with a range of functional forms both for EGP and SID. It is valuable to consider different functional forms for EGP since each variation is often used in practice: Table 3 shows that there can be substantial variations in the properties of an EGP/SEC measure depending upon the functional form used. We consider different functional forms of

---

[4] A more elaborate illustration of the lack of concordance between the CAMSIS and status_CG realisations of SID in their relationship to EGP is shown in A2 in the on-line-appendix.)

SID to provide fuller comparisons with class measures.[5] Looking at the ER-indicators (three first rows) it is apparent that the linear functional forms (M1 and M2) for SIDs in most cases are almost as strong as the more complex functional forms. The corresponding results for EGP indicate that the correlations are stronger the more categories are used. Taken together this pattern suggests that ER-indicators have quite a strong association with (any) stratification measures, but that crude measures of EGP (with few categories) can have a weaker association than might be predicted in theory. A similar pattern is seen for 'education years of job' and 'own educational score', while for the other variables SEC measures often perform better with more complex functional forms, suggesting some relevant non-linear patterns.

Table 3 replicates the findings of Tåhlin (2007) as it is indicated that measures of skills/education are more associated with class than ER-measures are. Moreover, it turns out that this is also true for SID-measures. Above we discussed 'occupational education' and 'occupational income' and suggested that SIDs may be closer to occupational education, and EGP to occupational income.[6] When comparing approximations of occupational education and income in Table 3 we see that both EGP and SIDs are somewhat stronger associated with occupational education rather than occupational income; however EGP does lean towards a relatively higher association with occupational income compared with SIDs. Hence, there is some support for this assumption, but at the same time EGP and SECs are very similarly associated with both occupational education and skill as well with individual education. In Table 3 we also report the association with individual earnings. Rows 8-10 summarize bivariate correlations. The last row of data follows the human capital/Mincer tradition and first controls for work experience (also as a squared term) and also education. In the next step

---

[5] SID is first treated as linear with a single metric variable (m1), second as curvilinear by adding a squared term (m2), third rescaled to its percentiles (including a squared term) which reduces the effects of extreme values (m3), fourth as 'quasi-classes' by transforming SID to the corresponding sizes of EGP and include them as dummies.

[6] To offer a rich description we have two measures of occupational education (which indicate similar results) and also a measure of individual educational level (for which correlations are smaller).

we include occupational stratification variables and their interaction terms with work experience, then the improvement in R^2 is reported for the different stratification variables. As with Goldthorpe and McKnight's (2006) analyses our results are potentially biased by cohort and period effects, and do not only reflect earnings growth, but we assume that this bias will be similar for different stratification variables. The results indicate that all stratification variables are similarly associated with earnings although EGP has a stronger association. However, the functional forms of the measures matter for the association with earnings; for all measures, more disaggregated non-linear functional forms produce better predictions, although the differences are not that great between model 3 and 4 for the SIDs.

In the on-line appendix (Figure A3) results based on predicted values from these earnings equations for men are condensed by reporting peak incomes across work experience for the different classes/SID groups (with constant values on all other covariates).[7] The same basic pattern is found as for (Goldthorpe, McKnight 2006), with a steeper progression of earnings for the Salariat (Service I and II, and especially for Service II). These variations were presented by Goldthorpe and McKnight as criterion validation of the EGP measure, yet the interesting contribution here is that a similar pattern is found for SID pseudo-classes (only CAMSIS is reported in the figure). However, the results from Figure 3 are somewhat unexpected in the sense that some less advantaged groups (using either EGP or SID) reach peak income at a relatively high level of work experience.

The findings so far indicate a large degree of similarity between measures of EGP and SID, including in suggesting similar work experience – earning profiles. However, the measures are not identical, and as we saw above EGP is somewhat more strongly associated with earnings. Hence, a key question concerns what drives the empirical differences between EGP and SID measures. One way to explore this in a multivariate context is to analyze the differences in the residuals (deviances between predicted and observed values) between regressions that predict other outcomes, based upon *either* EGP *or* SID. We do this below for the prediction of earnings, using models specified in the same way as for the last row in Table 3). Following theory, EGP should capture complexity of work tasks and related monitoring problems and asset specificity, as well as other aspects of employment relationships, better

---

[7] In the on-line appendix (Figure A4) the predicted earnings by EGP/status_CG and work experience are shown based on the regressions. To better understand the scale of the differences involved, the predictions in the on-line appendix are based on actual values in the data, taking account of group differences in the distribution of other covariates.

than SID, and this should lead to better earnings predictions using EGP that SID. If this is happening, the empirical reason for this pattern ought to be evident in that the differences in model residuals, between a model based upon EGP and one based upon SID, ought themselves to be related to measures associated with employment relationships.

--Table 4 about here --

In Table 4 we show selected results that summarize the extent to which the actual differences in performance between the EGP and SID measures are correlated with direct measures of things that they might, in theory, be related to. The values in Table 4 show an ad hoc partial-r2 statistic that is intended to show the relative scale of influence of different measures upon the magnitude of difference between the EGP and SID measure in accurately predicting earnings. It is clear from Table 4 that the relative influence of the different factors that we allow for, in shaping the differences in performance, are not well aligned to theory. The variables that are related to the employment relationships of employees (asset specificity, monitoring problems and service relationship scores) only explain a very small amount of the difference in residuals. There is some impact for experience, but it still hardly accounts for any of the differences in residuals. Moreover, in general, the bulk of the difference in performance between EGP and SID measures is not explained by the measured factors included in these models. In short, as we expected, the differences in predicting earnings by SIDs and by EGP do not seem to be substantially driven by differences in the SID and EGP measures' association with employment relationships.

For male employees, Table 4 shows that it is instead the gender composition of the occupation that is the most important source of difference between the performance of SID and EGP measures. Men located in female dominated occupations generally get lower predicted incomes when their occupation is summarized through EGP than through SID measures, and there are interactions between the gender composition of occupations and skill and industry. When we widen the population to all men (including the self-employed), the gender composition of the job still matters, but employment status becomes the strongest predictor of the difference in performance between EGP and a SID measure (employees have lower predicted incomes for EGP than SIDs, which also means that self-employed have higher predicted incomes for EGP). Since both versions of SID deviate from EGP in similar ways, it also does not seem like this is just an artefact that the CAMSIS measure uses employment

status in its derivation (whilst CG-Status does not). The importance of employment status in predicting residuals rather suggests that, for the full workforce, EGP deviates from SIDs by letting the self-employed to be a group of its own.

Table 4 shows selected results from a wider range of similar analyses that we have undertaken. Other analyses allowed for different functional forms for the SID and residual measures, for different population coverage, and for different model specifications in the earnings regression and in the models predicting residuals. Consistently however, across many permutations of results, we observed a similar pattern in this multivariate context as was seen in the bivariate comparisons: the bulk of the empirical difference between measures of EGP and SID does not seem to reflect the factors that should, in theory, account for it. Instead, the bulk of the difference is unexplained, but the fraction that can be explained seems mainly to reflect operational considerations and aspects of occupational differences, such as gender segregation in occupations, that are not - in theory - a part of the difference between concepts of class and status.

7. Concluding discussion

Our results suggest that the measures of class and status recommended by Chan and Goldthorpe largely reflect the same form of stratification, and do not reliably disentangle the theoretical concepts of class and status. As has been shown previously class accounts for the lion's share of the variation in status, and we can add that this is even more so if outliers are removed from the analyses. Moreover, the measures of class and status have very similar associations with employment relationship indicators as well with measures related to skills and education. It turned out that although the work experience – earnings profiles are similar for class and status measures, class did predict earnings somewhat better. However, when examined in detail, the small deviances in performance between the measures of class and status in predicting earnings do not seem to be due to theoretically meaningful differences, at least not for employees. For the full workforce the differences are to some extent due to differences in handling employment status where class schemas denote a category of self-employed. However, it is not the case that 'class' is a better predictor of earnings since it captures employment relations of occupations for employees while 'status' does not. In our view, therefore, the attempt to disentangle the influences of class and status by using the measures proposed by Chan and Goldthorpe has been mistaken.

So is it a bad idea to bring the concepts of class and status back in to stratification research? This is not necessarily the case. Weber's writings on class and status are arguably not in line with CG's interpretations. First of all, Weber used status in the sense of 'status groups' and used the German word 'Stände', which means estates. Hence, status groups are described as rather strong communities, who share an 'honor', and they are not necessarily easily ranked from low to high status. Second, the 'quality' that is the base for the status group can be any quality and not only occupation (Weber 1968, p.932). However, even if we restrict our interests to occupations as the 'quality' there are two sociological applications of status that may be more promising and closer to the Weberian sources. First, 'closure theory' (Murphy 1988, Parkin 1979) draws our attention to rather established status groups with clear boundaries, e.g. professions who have been successful in creating monopolies in the labor market, and writers in that tradition have presented Weber's concept of status in terms of such groups. Second, the idea of constructing 'micro-classes' which distinguish occupations that can be socially organized in different ways could also help identify status groups (Grusky, Weeden 2001, Weeden 2002). Obviously, some occupational groups do have communitarian features, as the incumbents of them collectively limit entrance by requiring certain diplomas and organizing themselves in unions and other guild type associations. But this does not mean that all individuals or occupational groups are part of status groups – some may be and some may not. In these ways, Weber's ideas on class and status might be usefully measured, but this describes a very different empirical realization of the division of the concepts to those proposed by Chan and Goldthorpe.

Finally, we think that the findings of this paper strengthen the view that stratification measures should be viewed as pragmatic tools. We can rarely escape the empirical reality that advantages at the occupational level tend to go hand in hand, and therefore stratification measures with different origins are nevertheless highly correlated. Whether we use long-established measures like EGP, ISEI, SIOPS and CAMSIS, or newer measures that might be designed to capture more recent divisions and/or national peculiarities in occupational or social structure, it will usually be the case that the more advantaged positions will still be associated with high skills/education, high incomes and accumulation of economic resources (see also Lambert, Bihagen 2014). They will also be associated with more advantaged social origins and other advantageous conditions. Certainly different measures may lean toward being more associated with a certain advantage, e.g. education/skills or incomes, and some may be more or less effective in capturing advantages in a certain setting (e.g. over time or

across countries). However, overconfidence in the influence of the theoretical rationale behind different stratification measures as an explanation for differences in their performance may lead to unsound conclusions. Consider for example a published claim that since EGP "has a clear theoretical basis", associations between EGP and health should be interpreted as being related to employment relations for employees (Shaw, Galobardes et al. 2007:78). A more realistic view is rather that such an association may be the result of several factors of socio-economic disadvantage that tend to cluster at the occupational level: low earnings, accumulation in advantages/disadvantages across the life course, relative deprivation, etc. Put differently: despite theoretically elegant conceptualisations of stratification, advantages will still be clustered at the occupational level, so that different stratification measures will still tend to pick up similar forms of stratification.

References

Alderson, A. S., Junisbai, A., & Heacock, I. (2007). Social status and cultural consumption in the United States. *Poetics, 35*(2–3), 191-212.

Bakker, B. F. M. (1993). A new measure of social status for men and women: the social distance scale. *Netherlands Journal of Social Sciences, 29*, 113-129

Bihagen, E., & Katz-Gerro, T. (2000). Culture consumption in Sweden: The stability of gender differences. *Poetics, 27*(5-6), 327-349.

Bihagen, E., Nermo, M., & Erikson, R. (2010). Social class and employment relations: Comparisons between the ESeC and EGP class schemas using European data. In D. Rose, & E. Harrison (Eds.), *Social class in Europe: An introduction to the European socio-economic classification*. Abingdon(Oxon): Routledge.

Bihagen, E., & Ohls, M. (2007). Are women over-represented in dead-end jobs? A Swedish study using empirically derived measures of dead-end jobs. *Social Indicators Research, 84*(2), 159-177.

Blossfeld, H. P. (2009). Educational assortative marriage in comparative perspective. *Annual Review of Sociology, 35*, 513-530.

Bukodi, E. (2007). Social stratification and cultural consumption in Hungary: Book readership. *Poetics, 35*(2–3), 112-131.

Bukodi, E., Dex, S., & Goldthorpe, J. H. (2011). The conceptualisation and measurement of occupational hierarchies: A review, a proposal and some illustrative analyses. *Quality and Quantity, 45*(3), 623-639.

Bukodi, E., & Goldthorpe, J. H. (2012). Decomposing 'Social origins': The effects of parents'

    class, status, and education on the educational attainment of their children. *European*

    Sociological Review, *29*(5), 1024–1039.

Chan, T. W. (2010). The social status scale: Its construction and properties. In T. W. Chan

    (Ed.), *Social status and cultural consumption* (pp. 28-56) Cambridge University Press.

Chan, T. W., Birkelund, G. E., Aas, A. K., & Wiborg, Ø. (2011). Social status in Norway.

    European Sociological Review, 27(4), 451-468. doi:10.1093/esr/jcq019

Chan, T. W., & Goldthorpe, J. H. (2004). Is there a status order in contemporary British

    society? Evidence from the occupational structure of friendship. *European Sociological*

    Review, 20(5), 383-401.

Chan, T. W., & Goldthorpe, J. H. (2007a). Class and status: The conceptual distinction and its

    empirical relevance. *American Sociological Review, 72*(4), 512-532.

Chan, T. W., & Goldthorpe, J. H. (2007b). Social status and newspaper readership. *American*

    Journal of Sociology, 112(4), 1095-1134.

Chan, T. W., & Goldthorpe, J. H. (2007c). Social stratification and cultural consumption:

    Music in England. *European Sociological Review, 23*(1), 1-19.

Chan, T. W., & Goldthorpe, J. H. (2007d). Social stratification and cultural consumption: The

    visual arts in England. *Poetics, 35*(2-3), 168-190.

Chan, T. W., & Goldthorpe, J. H. (2007e). Class and status: The conceptual distinction and its

    empirical relevance. *American Sociological Review, 72*(4), 512-532.

Chan, T. W, & Goldthorpe, J. H. (2007f). The social stratification of cultural consumption: Some policy implications of a research project. *Cultural Trends, 16*(4), 373-384.

Chan, T. W., & Goldthorpe, J. H. (2010). Social status and cultural consumption. In T. W. Chan (Ed.), *Social status and cultural consumption* (pp. 1-27) Cambridge University Press.

Christin, A. (2012). Gender and highbrow cultural participation in the united states. *Poetics, 40*(5), 423-443.

Collins, R. (1988). Women and men in the class structure. *Journal of Family Issues, 9*(1), 27-50.

Elias, P., Halstead, K., & Prandy, K. (1993). *Computer assisted standard occupational classification*. London: HMSO.

Erikson, R., & Goldthorpe, J. H. (1992). *The constant flux: A study of class mobility in industrial societies*. Oxford: Clarendon.

Erikson, R. (2016). Is it enough to be bright? Parental background, cognitive ability and educational attainment. *European Societies, 18*(2), 117-135.

Ganzeboom, H. B. G. (2009). Multiple indicator measurement of occupational background. *Paper Presented at the European Survey Research Assocation, Warsaw, 1 July 2009.*

Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research, 25*(3), 201-239.

Goldthorpe, J. H. (2000). *On sociology: Numbers, narratives, and the integration of research and theory*. Oxford: Oxford University Press.

Goldthorpe, J. H., & McKnight, A. (2006). The economic basis of social class. In S. Morgan, D.B. Grusky and G.S. Fields (Eds.), *Mobility and inequality: Frontiers of research from sociology and economics* (pp. 109-136). Stanford, CA: Stanford University Press.

Goldthorpe, J. H. (1996). Class analysis and the reorientation of class theory: The case of persisting differentials in educational attainment. *British Journal of Sociology, 47*(3), 481-505.

Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology, 24*, 395-421.

Katz-Gerro, T., Raz, S., & Yaish, M. (2007). Class, status, and the intergenerational transmission of musical tastes in israel. *Poetics, 35*(2–3), 152-167.

Lambert, P. S., & Bihagen, E. (2014). Using occupation-based social classifications. *Work, Employment & Society, 28*(3), 481-494.

Laumann, E. O., & Guttman, L. (1966). The relative associational contiguity of occupations in an urban setting. *American Sociological Review, 31*, 169-178.

Levine, J. H., & Spadaro, J. (1988). Occupational mobility: a structural model. In B. Wellman & S. D. Berkowitz (Eds.), *Social Structures: A Network Approach* (pp. 452-475). Cambridge: Cambridge University Press.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*, 415-444.

Mincer, J. (1974). *Schooling, experience, and earnings*. New York: National Bureau of
Economic Research.

Prandy, K. (1999). The social interaction approach to the measurement and analysis of social
stratification. *The International Journal of Sociology and Social Policy, 19*(9-10-11),
215-249.

Prandy, K. (1990). The Revised Cambridge Scale of Occupations. *Sociology, 24*(4), 629-655.

Prandy, K., & Lambert, P. (2003). Marriage, social distance and the social space: An
alternative derivation and validation of the Cambridge scale. *Sociology, 37*(3), 397-411.

Rose, D., & Harrison, E. (Eds.) (2010). *Social class in Europe: An introduction to the
European socio-economic classification*. Abingdon(Oxon): Routledge

Rose, D., & Pevalin, J. (Eds.). (2003). *A researcher's guide to the national statistics socio-
economic classification*. London: Sage.

Schwartz, C. R. (2013). Trends and variation in assortative mating: Causes and consequences.
Annual Review of Sociology, 39, 451-470.

Shaw, M., Galobardes, B., Lawlor, D. A., Lynch, J., Wheeler, B., & Smith, G. D. (2007). *The
handbook of inequality and socioeconomic position: Concepts and measures*. Bristol:
Policy Press.

Skopek, J., Schulz, F., & Blossfeld, H. P. (2011). Who contacts whom? Educational
homophily in online mate selection. *European Sociological Review, 27*(2), 180-195.

Stewart, A., Prandy, K., & Blackburn, R. M. (1980). *Social stratification & occupations*. New
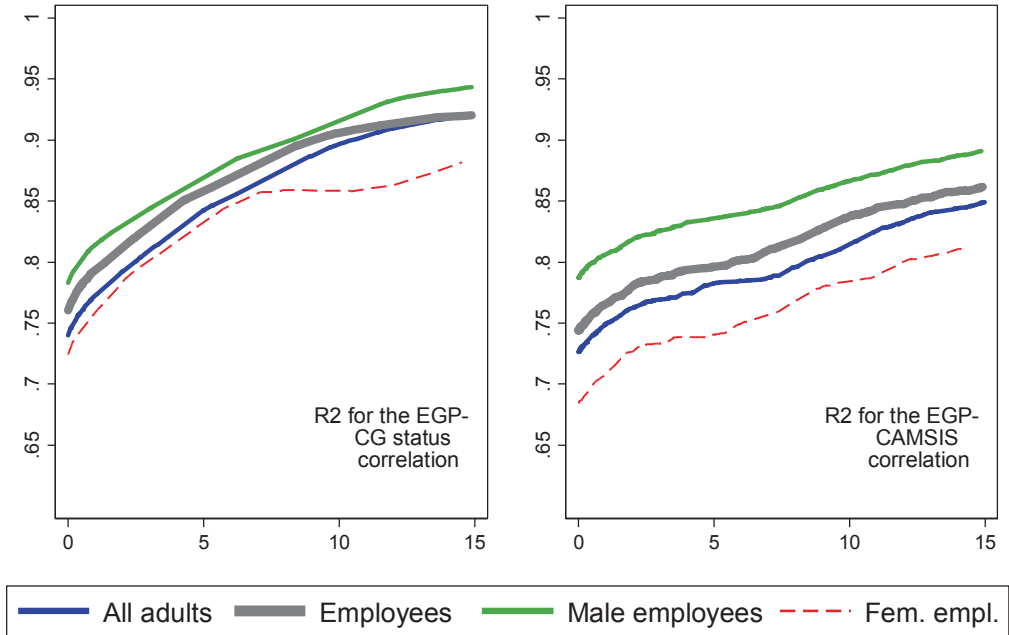York: Holmes & Meier.

Tåhlin, M. (2007). Class clues. *European Sociological Review, 23*(5), 557-572.

Torche, F. (2007). Social status and cultural consumption: The case of reading in Chile. *Poetics, 35*(2–3), 70-92.

Torssander, J., & Erikson, R. (2009). Marital partner and mortality: The effects of the social positions of both spouses. *Journal of Epidemiology and Community Health, 63*(12), 992-998.

Treiman, D. J. (1977). *Occupational prestige in comparative perspective*. New York: Academic Press, Inc.

University of Essex, Institute for Social and Economic Research. (2010). *British Household Panel Survey: Waves 1-18, 1991-2009 [computer file], 7th Edition*. Colchester, Essex: UK Data Archive [distributor], July 2010, SN: 5151.

University of Essex, Institute for Social and Economic Research, NatCen Social Research, Kantar Public. (2017). *Understanding Society: Waves 1-7, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009*. [data collection]. 9th Edition. UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-10

Warren, J. R., Sheridan, J. T., & Hauser, R. M. (1998). Choosing a measure of occupational standing: How useful are composite measures in analyses of gender inequality in occupational attainment? *Sociological Methods & Research, 27*(1), 3-76.

Tables and Figures for 'Can class and status really be disentangled?'

Table 1: Descriptives

	All adults (aged 25-65)				Employees only (aged 25-65)			
	mean	sd	min	max	mean	sd	min	max
Occupation-based social classifications:								
EGP: I	0.18		0	1	0.18		0	1
EGP: II	0.23		0	1	0.25		0	1
EGP: IIIa	0.13		0	1	0.15		0	1
EGP: IIIb	0.07		0	1	0.08		0	1
EGP: IVa	0.02		0	1	0		0	0
EGP: IVb	0.06		0	1	0		0	0
EGP: IVc	0.01		0	1	0		0	0
EGP: V	0.07		0	1	0.08		0	1
EGP: VI	0.06		0	1	0.07		0	1
EGP: VIIa	0.15		0	1	0.18		0	1
EGP: VIIb	0.01		0	1	0.01		0	1
SID: CG Status	0.01	0.37	-0.60	0.56	0.02	0.36	-0.60	0.56
SID: CAMSIS	53.1	14.0	10.2	95.7	52.4	14.0	10.2	95.7
Individual level measures								
Employment status: Employee	0.86		0	1	1		1	1
Gender: Female	0.48		0	1	0.51		0	1
Education: No quals.	0.13		0	1	0.13		0	1
Education: Low school level	0.07		0	1	0.07		0	1
Education: Medium school level	0.19		0	1	0.19		0	1
Education: Advanced schl. Level	0.11		0	1	0.11		0	1
Education: Post-schl. sub-degree	0.32		0	1	0.32		0	1
Education: Degree level	0.17		0	1	0.17		0	1
Work experience (years)	25.4	10.7	1	54	25.0	10.6	1	54
Work experience squared	761	585	1	2916	736	574	1	2916
Log of monthly earnings	7.02	0.80	3.91	11.19	7.05	0.77	3.91	11.19
Occupation-level measures								
Industry: Non-services	0.34		0	1	0.33		0	1
Industry: Private sector services	0.42		0	1	0.39		0	1
Industry: Public sector services	0.24		0	1	0.28		0	1
Service Relationship score	0.10	0.64	-1.35	2.12	0.12	0.65	-1.35	2.12
Occupational earnings	-0.11	0.49	-1.28	1.16	-0.13	0.50	-1.28	1.16
Asset specificity score	36.5	14.4	3	75.5	36.1	14.6	3	75.5
Monitoring problems score	51.0	16.3	17	86	50.6	15.9	17	86
Years of education	49.1	21.5	8	93	49.1	21.3	8	93
Skill requirements	34.3	19.4	1	91	34.1	19.5	1	91
Women in job: < 35%	0.30		0	1	0.29		0	1
Women in job: 35-65%	0.31		0	1	0.28		0	1
Women in job: > 65%	0.39		0	1	0.43		0	1
N records	107716				92884			
N unique respondents	15364				14083			

Figure 1: Explained variance (R2) for EGP on SID measures after dropping occupations with the highest residuals (% of sample dropped is on the X axis)

Source: BHPS pooled records for 25-65yrs in employment 1991-2008, N=108304k. Lines show R2
values for relation between the measures and % of sample used (regressions use full sample
then successively remove up to 15% of cases from those with the highest magnitude residuals).

Table 2: Occupational categories that contribute the most to the residuals for the regression of SID by class (EGP-11 categories)
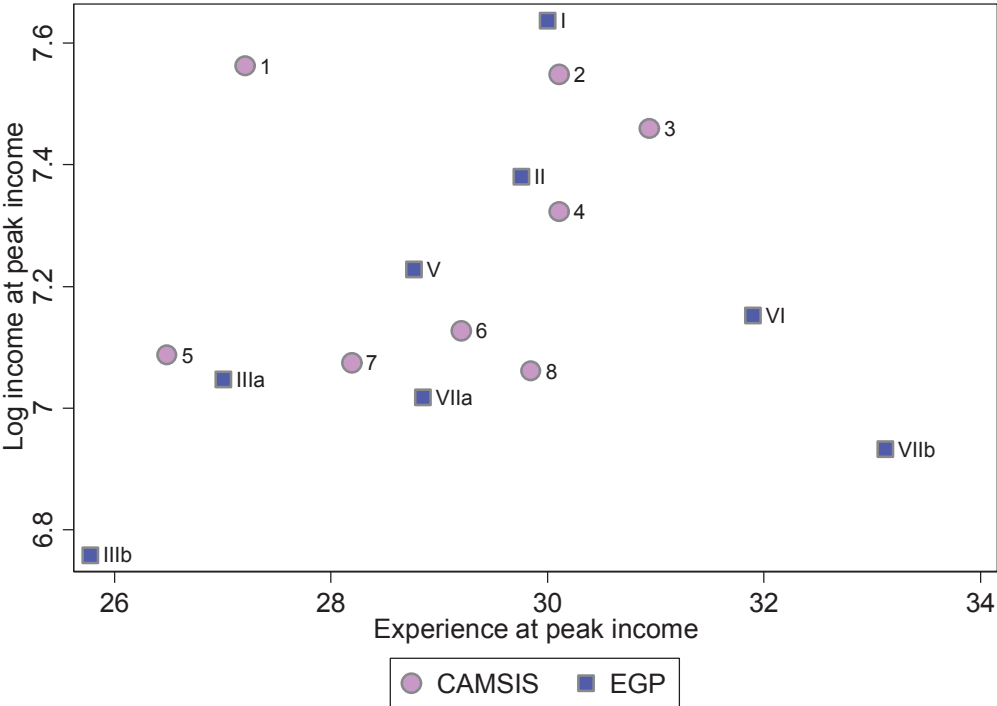
Occupation	Modal emp. status	Modal EGP	SID zscore	% of residuals	% cases	Actual residual
CAMSIS-EGP						
340. Nurses	Supervisor	II	0.33	3.2	2.3	-0.56
720. Sales assistants	Employee	IIIc	-0.12	2.3	3.2	0.62
430. Clerks	Employee	IIIa	0.12	2.0	2.8	-0.97
644. Care assistants	Employee	VIIa	-0.69	1.9	1.6	1.10
459. Other secretaries n.e.c.	Employee	IIIa	0.68	1.9	1.5	0.11
410. Accounts clerks	Employee	IIIa	0.35	1.8	2.4	0.44
233. Secondary teachers	Employee	II	1.28	1.7	1.6	1.24
371. Welfare workers	Employee	IIIa	0.70	1.6	1.3	1.09
179. Managers in services	Manager	II	0.44	1.6	2.4	-0.36
320. Computer analysts	Employee	II	0.32	1.5	1.1	-1.60
110. Production managers	Manager	I	0.49	1.5	1.2	-0.26
CG-status-EGP						
179. Managers in services	Manager	II	-0.14	4.3	2.4	-1.60
110. Production managers	Manager	I	-0.19	3.4	1.2	-1.70
720. Sales assistants	Employee	IIIc	-0.33	2.8	3.2	-0.47
610. Police officers	Employee	II	-0.64	2.5	0.7	-2.60
234. Primary teachers	Employee	II	1.36	2.5	0.7	1.30
233. Secondary teachers	Employee	II	1.36	2.2	1.6	1.30
199. Other Managers n.e.c.	Manager	I	0.63	1.9	1.8	2.90
644. Care assistants	Employee	VIIa	-0.60	1.9	1.6	0.96
659. Other childcare n.e.c.	Employee	IIIb	0.28	1.9	1.1	0.74
121. Sales managers	Manager	I	1.40	1.9	1.5	1.40
111. Buildings managers	Manager	I	-0.19	1.7	0.6	-2.40

'% of residuals' refers to the sum of magnitude of residuals for the occupation, as a total of the magnitude of all the residuals across the sample. The actual residual equates to the difference between the EGP mean and the mean for the individual occupation: a positive value implies higher SID score than average, and a negative value implies lower SID score than average.

Table 3: Correlations*100 for relationship between occupation-level and individual level measures and the class and SID scale measures, allowing various functional forms for the SID scale (employees only)

	EGP (#classes)				CG				CAMSIS			
	11(8)	6(5)	5(3)	2	M1	M2	M3	M4	M1	M2	M3	M4
SR-score*	80	75	75	66	67	67	70	73	75	76	77	79
Asset specificity*	75	74	69	65	59	64	62	74	68	68	70	71
Monitoring problems*	77	75	74	70	70	70	72	81	73	74	75	77
Education years of job*	87	84	83	79	84	84	84	86	87	87	88	87
Skill requirements*	83	81	80	79	74	77	77	79	79	81	81	82
Occupational earnings*	75	73	65	64	37	56	57	67	50	54	58	61
Own educational score	50	49	49	47	47	51	49	50	51	52	53	54
Own log earnings	57	54	48	47	25	40	41	45	37	40	43	45
Own log earnings (men)	51	47	45	42	35	37	37	42	43	43	44	45
Own log earnings (women)	61	59	59	54	47	51	50	53	51	52	53	55
Log earnings increment**	12	11	10	8	3	5	5	7	6	6	7	8

* = indicator at the occupational level (value is correlation weighted by individual level data)

Value shown is calculated as square root of R2 statistic from a bivariate regression with the relevant score as outcome, except (**) which shows the increase in R2 between baseline model with controls for gender, education level and work experience, gender-experience and experience-education interactions, and the increased value after adding the occupation-based measure and its interaction with experience. Baseline r2 =0.35.

EGP classes are those of employees for different EGP versions with the indicated number of different categories.

M1: SID scale in linear functional form; M2: SID in quadratic functional form; M3: SID percentile value in a quadratic functional form; M4: SID in 8 quantile categories, allowing a categorical functional form. Source: as Table 1.

Figure 3: Peak log incomes by experience, by EGP category or CAMSIS octile for men (Experience at peak income based on predicted values from income regression)

Source: As Figure 1.

Table 4: Partial R2 statistics in four models predicting the absolute difference between SID and EGP residuals in predicting income

	Male employees		All adult males	
	CG status	CAMSIS	CG status	CAMSIS
Individual level influences				
Experience and experience-squared	0.2*	0.1*	0.1	0.1
Education (6 category measure)	0.7*	0.1*	0.5*	0.4*
Employee status			-12.2*	-12.2*
Occupational level influences				
Service relationship (z-score)	+0.2*	+0.2*		
Asset specificity (z-score)	-0.1	-0.1*		
Monitoring problems (z-score)	+0.1	-0.0*		
Job skill level (z-score)	-1.8*	-0.1*	-1.1*	-1.0*
Gender of job (3 categories: % female (0-35; 35-65; 65+)	-2.9*	-2.2*	-0.8*	-0.4*
Industry of job (3 categories: Private sector services; public sector services; all others)	0.7*	0.0*	0.5*	0.2*
Gender (>65% female)*Job skill score	+0.7*	+0.0*	+0.2*	+0.0*
(>65% female)*Job skill*services sector	-0.1*	-0.1*	-0.1*	-0.0
N records	41622	41622	50107	50107
N respondents	6471	6471	7268	7268
Total model r2*100	6.2	4.2	14.2	13.7

Notes:

Outcome variable in each model is the absolute difference in residuals for two regressions predicting log income, one including EGP as a predictor and the other including SID. Income prediction regressions use occupation-based social classification, and controls for education, experience, experience squared, and interactions between education and experience, and the stratification measure and experience. SID measures used in these models were in 8-category functional form.

For each model, the values shown are approximated partial r2 values*100 for the relevant variable. * indicates the variable is significant at 95% criteria. Partial r2 is approximated as the difference between the model r2 (Efron's r2) with and without the other explanatory variables. For linear/ordinal measures, sign indicates direction of association between the measure and the absolute difference in residuals. All models are based upon BHPS records for adults interviewed 1991-2008 with relevant valid data. Models use linear regression (random effects for panel-level clustering were tested, but had no important impact on results).

More detailed descriptions of some of the variables used are found here (i.e. occupations, social interaction distance scales, employment status, education, employment relationship indicators, and education/skills of occupations):

Occupations. For the BHPS data occupation-based measures and scores are derived for the 3-digit form of the UK Standard Occupational Classification 1990 (OPCS 1990). This scheme differentiates between 371 different occupational unit groups. Occupation-level information from the European Social Survey was developed at the level of ISCO-88 3-digit 'minor' groups (ILO 1990), which in this dataset differentiates 111 different occupational units. The BHPS UK occupational codes are translated to 4-digit ISCO-88 units by the data producers, based upon a cross-walk algorithm provided by the UK's Office for National Statistics, and the ESS occupational data is linked to the BHPS through these ISCO-88 units.

Social interaction distance scales (SID). The UK CAMSIS scale for Britain (1991 version, based upon marriage patterns, was accessed from www.camsis.stir.ac.uk [1/8/16]). When matched to the BHPS, this scale differentiates between 459 different occupation/employment status categories for the 'male scale' and between 350 different units for the 'female scale'. The CAMSIS scale calculates separate scales for male and female incumbents of occupations. Used independently, both are said to represent the relative position of the incumbents of the occupations within the respective male/female occupational distributions. Following common practice, however, all our analyses focus on the male scale values applied to all relevant respondents.

Employment status. We used a measure of employment status that is widely used in UK datasets. The seven category measure distinguishes between employee, supervisory, managerial and employer roles, and also takes account of the size of the organisations involved (see description at http://www.geode.stir.ac.uk/ougs.html#ukempst).

Education. For respondents' education the analyses use the derived variable for the 'highest educational qualification held'. This variable ('qfedhi') has 12 valid categories (University of Essex, Institute for Social and Economic Research, 2010). Our analysis used a recoded version with 6 categories: no qualifications (qfedhi=12); lower vocational, other, and lower school level qualifications (qfedhi=8,9,10,11); medium school level qualifications (qfedhi=7); higher school level qualifications (qfedhi=6); post-school academic qualification below degree level (qfedhi=3, 4,5); qualifications at degree level and above (qfedhi=1,2).

Employment relationship indicators. Using BHPS data, we extracted measures of whether individual respondents reported that their job featured standard pay rises, promotion opportunities, payment by salary rather than hourly rates, opportunities to work flexibly, opportunities to control working hours, a defined career path, and whether the job was held permanently. For each measure we calculated a mean standardized occupational level average (these were calculated using the 'empirical Bayes' residuals in a random effects model clustered by occupations, since the 'shrinkage' property of these residuals reduces the risks of distorted estimates for those occupations with few incumbents – see e.g. Mills 2007). A derived occupation-level service relationship score was then obtained from the average of these item-level residuals.

The measures of asset specificity and monitoring problems based on ESS-data used three variables from waves 1 to 5: how much time it takes to learn the job if the employee has the right qualifications and if the job requires learning new things (only wave 2 and 5: variable names "Jblrn" and "Jbrqlrn", see http://www.europeansocialsurvey.org/data/), and a question on the respondent's autonomy at work ("Wkdcorga"). Using the ESS-data ISCO-88 3-digit occupational units account for respectively 19 and 21 percent of the variation in these variables (OLS regression). We also considered using an additional measure of asset specificity based on an assessment of how replaceable the respondent is ("Rpljbde"), and a more direct measure of monitoring problems which is based on asking how difficult it is for the immediate boss to know how much effort (you) put into work ("Bseftwk"). However, it turns out that occupational units discriminate very poorly on these variables (r2: .06 and .01, OLS) and, hence, valid occupational scores cannot be derived from these variables. It goes beyond the scope of this paper to judge if the low discriminatory power of occupational units for these two variables is a consequence of low validity of the measures, or of a genuine lack of relationship.

Education/skills of occupations. These two ESS-based measures used the variable Eduyrs from wave 1 to 5 respectively Jbedyrs from wave 2 and 5. At the individual level in ESS, occupational units account for 31 and 34 percent of the variation in these two variables (OLS).
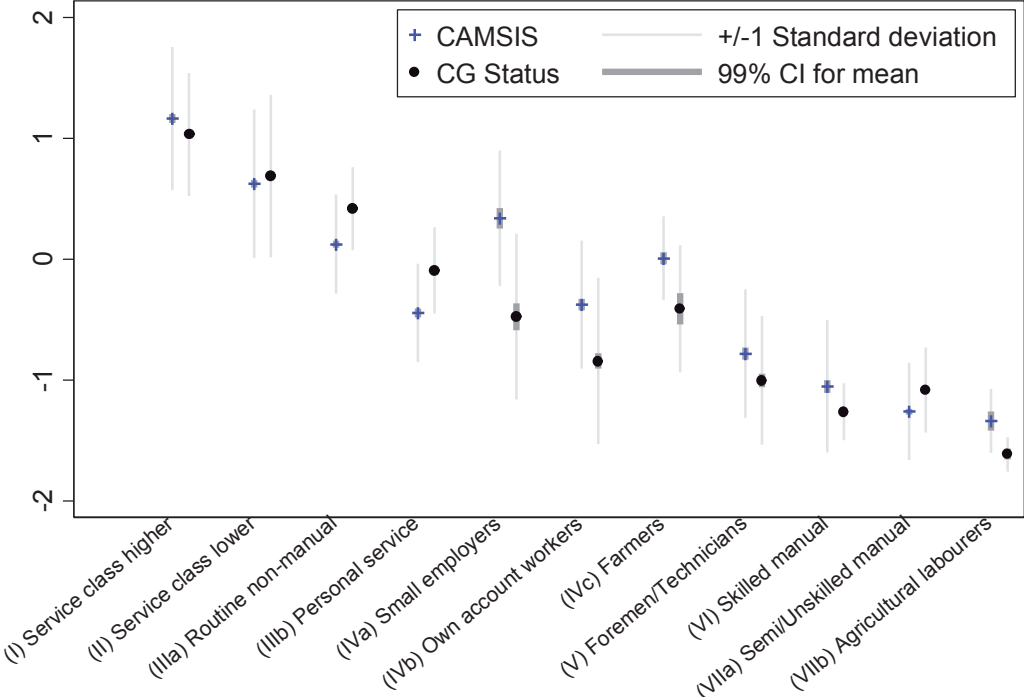
References:

ILO. (1990). *ISCO-88: International Standard Classification of Occupations*. New York: International Labour Office.

Mills, C. (2007). Unequal jobs: Job quality and job satisfaction. In S. Hill, C. Mills & M. White (Eds.), *Market, class and employment* (pp. 233-282). Oxford: Oxford University Press.
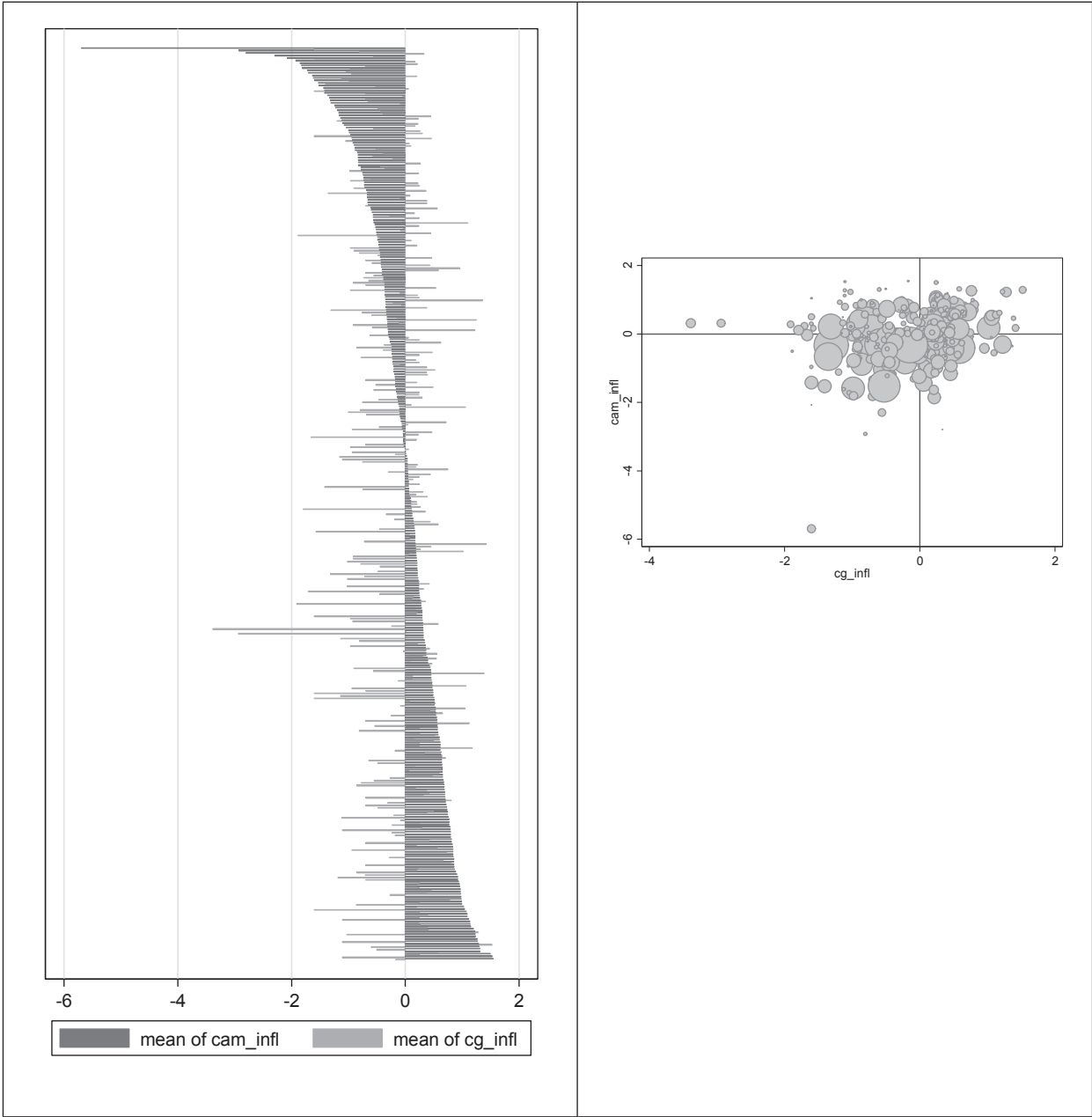
OPCS (Office for Population Censuses and Surveys). (1990). *Standard Occupational Classification, Volume 1: Structure and definition of major, minor and unit groups*. London: OPCS.

Figure A1: Mean and standard deviation for SID by EGP (SID scales in Z-score units)
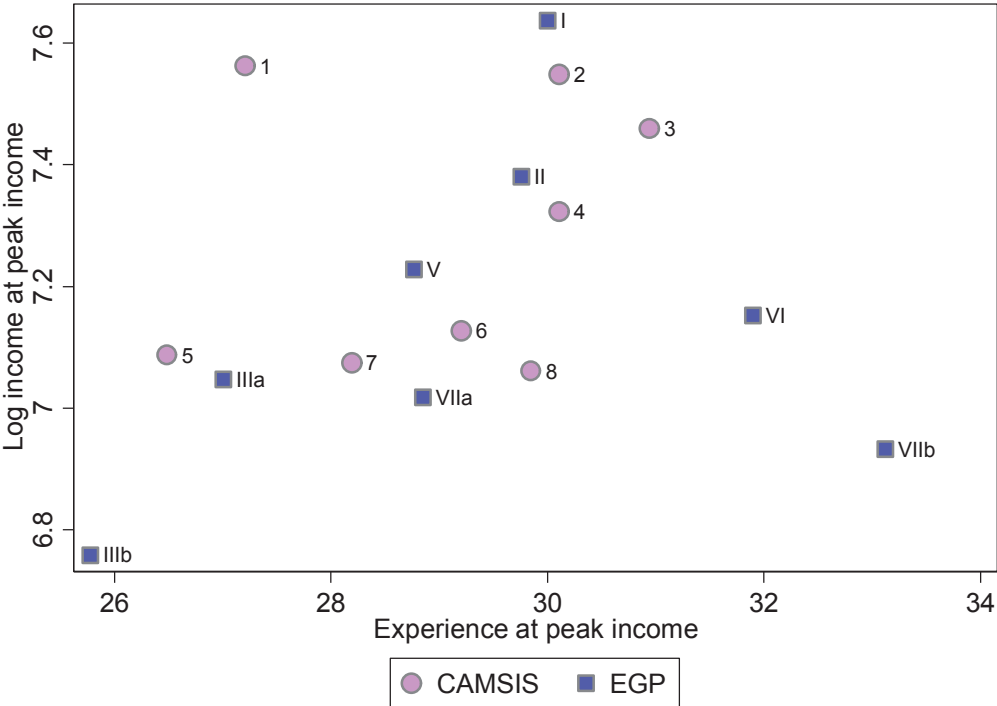

Source: As Figure 1.

Figure A2: Relationship between occupation level influence statistics for the CAMSIS-EGP and CG-Status-EGP regressions respectively
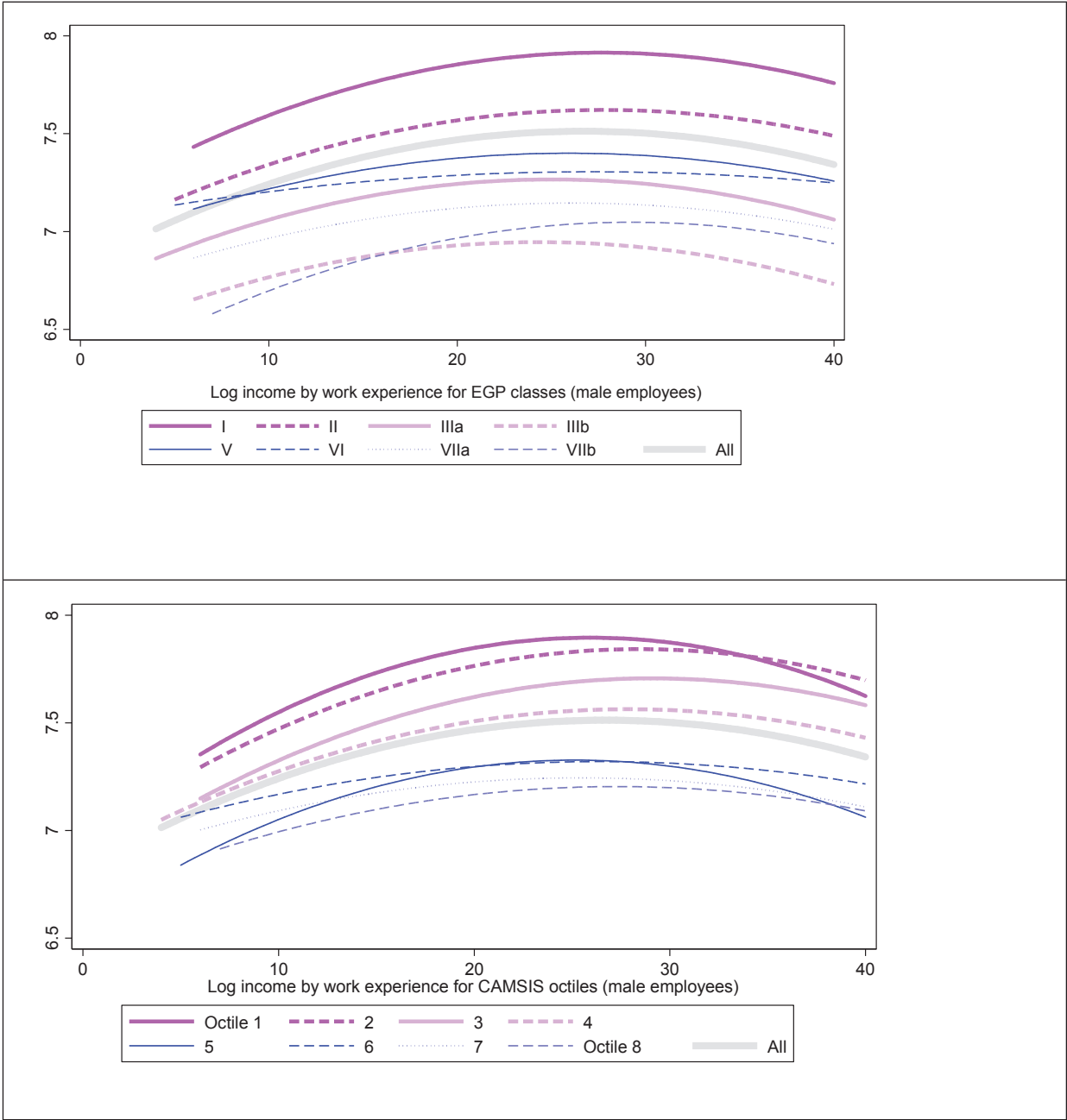

Notes: Analysis based upon UK BHPS (as Figure 1). Figures show two alternative depictions for the same occupations for the two sets of relationships (residuals for CAMSIS-EGP, and CG-status-EGP), in both cases adjusted for their relative influence as a proportion of the total volume of residuals (a value above 0 means more influence than average, and a value below 0 means less influence). Panel A shows the relationship in rank order of the CAMSIS-EGP residuals, and Panel B shows them with weightings in which the size of the point is proportional to the number of cases in the occupation. Both figures suggest substantially random differences between the nature of how the two SID measures, CAMSIS and CG-Status, relate to EGP.

Figure A3: Peak log incomes by experience, by EGP category or CAMSIS octile for men (Experience at peak income based on predicted values from income regression)

Source: As Figure A1.

Figure A4: Experience-earnings curves for EGP categories (panel 1) and for CAMSIS octiles (panel 2) for male employees (based on predicted values for earnings regression)

Source: As Figure A1