Appl. Spatial Analysis (2016) 9:463–489
DOI 10.1007/s12061-015-9163-1

Small-Area Incomes: Their Spatial Variability and the Relative Efficacy of Proxy, Geodemographic, Imputed and Model-Based Estimates

Paul Williamson [1]

Received: 30 July 2014 / Accepted: 22 July 2015 /
Published online: 4 August 2015

Abstract This paper uses data from a UK Census Rehearsal to explore the problem of small-area income estimation. First, the nature of the problem is revisited through an examination of the way in which incomes vary spatially. Residential rather than labour market sorting is found to be the dominant driver; and the rich are found to exhibit greater spatial segregation than the poor. Even so, location is shown to capture only a small fraction of the overall variation in incomes. Second, the performance of competing small-area estimation strategies is assessed, uniquely comparing proxy, geodemographic, imputation and model-based estimates; and validating all of these against directly observed values. An area-level model, ecological regression, performs best. Unit-record imputation approaches capture similar levels of spatial variation in mean income, but have higher variances and greater systematic biases. The same can be said of a simple univariate proxy (% professionals), which even so proves surprisingly effective.

Keywords Income distribution · Income segregation · Imputation · Synthetic estimate · Ecological regression

Introduction

It is becoming increasingly apparent that inequalities in income underpin a wide range of social phenomena, from voting and leisure habits, through to long term health prospects and, ultimately, life expectancy (Dorling 1999; Wilkinson and Pickett 2009). For this reason, and notwithstanding debates about the mismatch between measures of income, deprivation and wealth (Gordon et al. 2000), there remains strong

✉ Paul Williamson
P.Williamson@liv.ac.uk

[1] Department of Geography, Roxby Building, University of Liverpool, Liverpool L69 7ZT, UK

user demand for information on the spatial distribution of incomes. This was evidenced in the UK during the run-up to both the 2001 and 2011 UK Censuses, when consultation revealed overwhelming user support for the addition of a census question on income (Moss 1999; Rees 1998; ONS 2006). This demand came from commercial companies, national and local government agencies, campaigning groups and academia. Potential uses cited included the allocation of public resources; equality monitoring; identification of areas where housing affordability or fuel poverty is an issue; and the location of commercial retail outlets. Even so, following pre-testing income was dropped from the final version of the Census form used, both in 2001 and 2011, in the main due to concerns about the potential negative impacts on response rates, and in part due to concerns about the quality and bias of the responses that would be obtained (Moss 1999; Collins, undated).

However, lack of small-area income data is not just a UK issue. Government-sponsored surveys, censuses and population registers can fail to satisfy user needs in one or more ways – partial coverage of the population (e.g., earners only), insufficient sample size, lack of detail due to disclosure control measures, lack of spatial coding, and lack of timeliness, especially for inter-censal periods. At the same time alternative commercial sources of spatially detailed income data suffer from problems of access (notably cost barriers) and sampling bias arising from to the way in which data are collected (Birkin and Clarke 1995; Longhurst et al. 2004b). As a result many researchers and organisations around the world are actively involved in the production of small-area estimates of average (mean) incomes. These include academics (Lee et al. 1995; Jargowsky 1996; Cloutier 1997); commercial 'lifestyle' companies (CACI 1999; Experian 2005) and national statistical organisations (Bond and Campos 2010; The EURAREA Consortium 2004). In this context, a 'small area' is defined as any spatial unit lacking a timely and reliable survey- (or register-) based estimate.

The first hurdle to be overcome in producing effective small-area estimates of income is an understanding of the nature of estimation problem. Therefore the first question this paper explores is the scale(s) at which income appears to concentrate. The paper extends the work of others by examining this question for both a wider range of spatial scales and a greater variety of measures of income (individual, household and equivalised). The results presented add to the mounting evidence that in neoliberal market economies the poor are less spatially segregated than the rich (Dorling et al. 2007; Berthoud 2008), and that residential rather than labour-market sorting is the key driver of spatial variations in income. It is further shown that location captures only a small fraction of the overall variation in incomes; and that even this 'area effect' disappears once sufficient account is taken of local population composition. This provides hope that an adequately specified small-area estimation strategy should be able to capture a significant element of the observed spatial variation in incomes.

The second focus of this paper is upon the efficacy of alternative strategies for estimating average (mean) small area incomes. These include univariate and multivariate proxies (indices), and a range of synthetic estimates, including geodemographic classification (mean income given area type), imputation (unit-level models) and ecological regression (area-based models). This coverage contrasts with previous studies addressing the issue of estimate validation, all of which have been restricted to considering only one of the above four approaches.

Further, data limitations have caused many previous evaluations to compare estimates only against other measures that are not quite equivalent – such the socio-economic composition of the local population, or the incomes of those known not to be fully representative of the whole population (e.g., bank account holders and tax-payers). Nor has any previous evaluation considered the performance of income estimates in relation to arguably the most common policy-relevant goal – the correct ranking of areas by income. The evaluation reported in this paper addresses all of these shortcomings. In a UK context, the findings of the evaluation suggest clear ways forward for producing census-based estimates of small-area incomes using results from the 2011 Census; and subsequently as these Census data begin to age. In a wider international context the precise way in which incomes segregate spatially will clearly be country-specific. Even so the relative efficacy of alternative estimation approaches identified by this paper is likely to be generalisable, in particular with regards to which strategies are most and least sensitive to changes in spatial scale and type of income unit (person or household).

The structure of this paper is as follows. Second section introduces the Census Rehearsal dataset and some of the methodological issues associated with its analysis. "The Spatial Variability of Incomes" section uses the Rehearsal data to explore the ways in which incomes vary spatially. "Small-Area Income Estimation Strategies" section reviews the nature and limitation of the four main approaches to small-area income estimation: proxy indicators; geodemographic classification; imputation and model-based estimates. This section also details how each strategy was implemented to provide a series of rival small-area estimates. "The Relative Efficacy of Alternative Estimation Strategies" section evaluates the relative efficacy of these competing estimates, and considers the extent to which they are 'fit-for-purpose' by examining their relative accuracy at ranking areas. "Conclusion" section concludes by summarising the main findings of the paper and reviewing their relevance to current UK and wider international debates.

The Census Rehearsal Dataset and Associated Methodological Issues

To undertake this research a unique data resource was exploited: the 1999 UK Census Rehearsal – a survey of nearly 150,000 households sampled from parts of Angus, Bournemouth, Ceredigion, Dundee City, Gwynedd, and Leeds. The rehearsal supplemented the standard set of census questions with a question on gross income from all sources. This question was addressed separately to each household member and used a closed set of income bands. Questions of this type can be expected to provide results broadly in line with those obtained from a more detailed investigation into unbanded income (Micklewright and Schnepf 2010).

Three aspects of the Census Rehearsal make it uniquely suitable for the task of exploring the spatial variability of income and the relative efficacy of alternative small-area estimation strategies. First, the available geo-coding allows respondents to be located to the nearest unit postcode (blocks of c. 15 households). Second, this information was collected from spatially contiguous households spread across

relatively large areas (wards or groups of wards). Third, the Rehearsal has a high achieved sample size – some 53 % of the total population living in the areas surveyed. As a result Rehearsal data can be used to calculate rates and averages for a wide range of nested spatial units. This contrasts with, for example, the Family Resources Survey (FRS) used by Berthoud (2008) and Guangquan et al. (2011), which locates respondents only to the nearest postcode sector (blocks of c. 2600 households), lacks spatial contiguity and has a low achieved sample size (15 households per postcode sector).

As a check against non-response bias, selected indicators from the 1999 Rehearsal were compared with their 1991 Census equivalents. Making due allowance for the passage of time, the two datasets were found to be strongly correlated. To further guard against non-response bias, all analyses in this paper have been repeated for sub-sets of the survey data ranging from all individuals with any valid responses, through to a sub-set comprising only households supplying complete information for all household members. No differences of substance were found, so the results presented in this paper are based on analysis of the full set of rehearsal responses.

To enable the calculation of mean incomes by area, the income bands recorded in the survey had to be replaced with a set of imputed values. Analysis of the 1998 FRS revealed that variation in income-band means was minimal between population sub-groups (social strata), except for the bottom (non-zero) and top (open) income bands. Consequently the income value associated with each income band was imputed using the national mean value for that band observed in the FRS. To check the sensitivity of results to this imputation strategy, the income value associated with the bottom and top bands were also imputed using a log-normal modelling approach which adjusted the imputed mid-point to reflect the shape of the local income distribution. Although this alternative imputation strategy led to more pronounced between-area differences, it did not lead to any change in the overall conclusions reported in this paper. Given mixed evidence about the relative merits of the two imputation strategies, results throughout this paper are based upon imputation using the more conservative strategy of national band means.

For the smallest spatial units, and for certain non-rehearsal geographies, the number of rehearsal respondents is so small that any results obtained must be treated with caution. For this reason all spatial units with responses from fewer than 10 households or 25 residents have been excluded from the analyses reported in this paper. When appropriate, areas containing fewer than 25 adults with known income bands have also been excluded. Re-running the analyses including all areas, regardless of sample size, leaves the conclusions drawn in this paper unchanged.

The ‘Modifiable Areal Unit Problem’ was tackled by repeating analyses of the Census Rehearsal using the 1991 Census and 1999 Census Rehearsal geographies, which differed at the district, ward and enumeration district levels. It was found that the change of geographies made no substantive difference in results. For the sake of clarity, therefore, all of the results reported in this paper are based solely on the 1999 Rehearsal geography. The full set of spatial units analysed are reported in Table 1, alongside an indication of their number and average size.

Table 1 Census Rehearsal respondents

		Mean respondents per area with known incomes	
Spatial unit	Units	Adults	Households
District	7	14,071	7853
Ward	40	2462	1374
Postcode Sector	73	1348	723
ED	648	152	84
Postcode Unit	1269	37	20

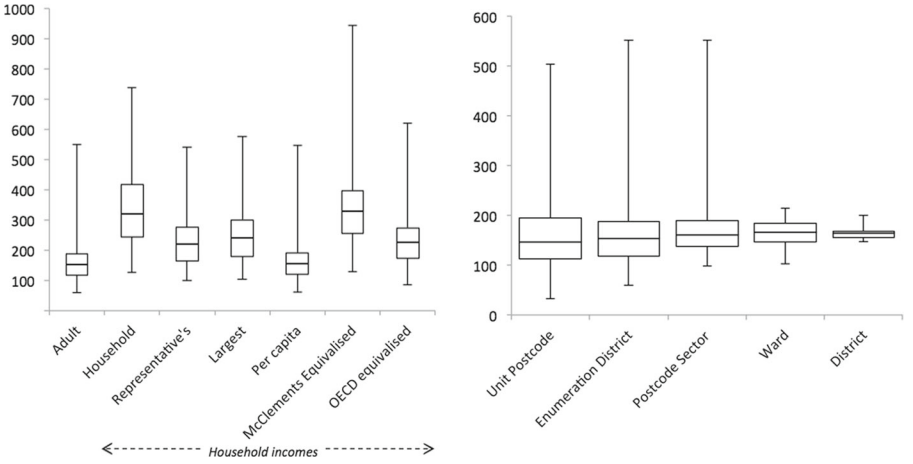
Based on 1999 Rehearsal Geography. Excludes areas with less than 25 known adult incomes, hence total respondents varies by spatial unit. Household counts exclude households with one or more unknown adult incomes

The Spatial Variability of Incomes

The Census Rehearsal dataset allows the spatial variability of incomes to be investigated from four different angles: changes in variability with scale and with unit of measurement; the extent to which variability in incomes is explained by variation between areas; and the extent to which spatial segregation varies with income.

Measurement Effects

Analytical interest may lie in the spatial distribution of individual, household or equivalised household incomes. For this reason Fig. 1a shows the distribution of mean incomes per Enumeration District (ED) for a variety of income measures. Of the numerous household equivalence scales in circulation, only two were chosen. These were the McClements scale, the de facto scale of choice within the Office for National



a) By income measure (ED-level) b) By spatial unit (adult incomes)

Fig. 1 The spatial variability of mean area incomes (£ per week, gross)

Statistics until c. 2007, and the modified OECD scale, subsequently adopted by the ONS and already used widely by analysts around the world. Similar spatial variability is found for all income measures except Total and McClements-equivalised household income, both of which show considerably greater variability, the latter strikingly so. This accords with Berthoud's finding, based upon an analysis of the 1994 and 1995 FRS, that McClements-equivalised net household income displayed a wider range of variation between Postcode Sectors than any other income measure (Berthoud 2008).

Berthoud's finding serves as a reminder that the Rehearsal measured gross rather than net incomes. Given that gross incomes include state welfare transfers, the spatial variability in incomes reported in Fig. 1, and throughout the remainder of this paper, could be viewed as an upper limit, likely to be reduced following the taxation of high income individuals – assuming that they are spatially clustered. That said, Berthoud reported finding greater spatial variability between net rather than gross household incomes, perhaps due to temporal lags between current incomes and earlier housing choices. The banded nature of the income question used by the Rehearsal will also act to dampen observed between-area variability.

For all income measures the average (mean) income per ED is positively skewed. As a result EDs with the very highest average incomes lie much further from the national average than EDs with the lowest incomes. This would appear to reflect in part the safety net offered by the British welfare state (a floor to how low incomes can go); and in part the structure of labour market opportunities (no income ceiling). Whether this positive skew is also reflective of the rich showing more spatial clustering than the poor is a question deferred until "Disentangling population composition and spatial context" section. Household incomes also show greater spatial variability than adult incomes – wider overall and inter-quartile ranges – suggesting that spatial sorting is driven more by household than by individual incomes, with households pooling resources to meet their housing costs.

Between-Area Variation

Figure 1b explores the impact of spatial scale on the observed variability in mean adult incomes. As might be expected, the smaller the spatial unit, the greater the observed diversity in mean incomes between areas. Even ignoring outliers the inter-quartile range of observed small-area incomes increases as the size (population) of the unit of analysis decreases. A similar pattern (not shown) was found for all of the measures of income considered in Fig. 1a. At its most extreme, the mean adult income for the most affluent Unit Postcode was 15 times that of the least affluent; whilst the mean income of the most affluent district (Bournemouth) was one-third higher than that of the least affluent (Leeds). But even at district level the apparently narrow range of between-district differences masks considerable within-district variability. Leeds, for example, was home to the Rehearsal EDs with the highest and lowest mean incomes. The distribution of mean ED incomes also varies. Leeds, for example, followed the national pattern with a clear positive skew to its distribution of mean ED incomes, whilst Gwynedd, a mostly rural area in Wales, showed a marked negative skew.

Following Berthoud (2008) it is also possible to assess the spatial sorting of incomes more formally, via an analysis of variance. The outcomes of this analysis are presented in Table 2. This extends Berthoud's work (which was based upon an analysis of the FRS) by presenting equivalent results for adult as well as household incomes, for

Table 2 The spatial component of income variability

	Adult incomes				Household incomes			
	Between-area variation (%)		% Contribution to between-postcode variation		Between-area variation (%)		% Contribution to between-postcode variation	
	Empirical	Non-structural	Empirical	Non-structural	Empirical	Non-structural	Empirical	Non-structural
District	0.9	0.9	6.2	7.6	0.8	0.8	3.4	4.2
Ward	2.2	2.2	9.6	11.6	3.9	3.8	12.6	15.6
Sector	4.7	4.6	18.0	22.1	6.8	6.7	11.9	14.8
ED	9.6	8.9	34.8	37.8	15.6	14.4	35.5	39.4
Postcode	13.9	11.2	31.3	20.9	24.6	19.5	36.7	26.1
			100.0	100.0			100.0	100.0

(1) Sector=Postcode Sector; ED=Enumeration District; Postcode=Postcode unit; (2) Based on 1999 Rehearsal Geographies; (3) Empirical=variance measured in Census Rehearsal; Non-structural=Empirical less variance to be expected by chance alone

spatial units smaller than postal sectors, and by taking account of the extent to which between-area variations in income are a structurally inherent by-product of the interaction between income banding and shrinking sample size.[1]

The first point to draw out from Table 2 is that the structural component of observed between-area variations in income (adult or household) is negligible for spatial units of ED size and above; but accounts for fully one-fifth of the empirically observed between-postcode variations in income.

The second message of Table 2 is that the spatial segregation of incomes is dominated by the smallest geographies. This finding supports the view that residential sorting, rather than larger-scale structural differences between labour markets, dominates the spatial distribution of personal incomes. This suggests that as spatial units decrease in size they increase in population homogeneity, leading to increasing spatial segregation of incomes and, therefore, increasing diversity of mean incomes between areas. But homogeneity is a relative term. Even at postcode-unit level location accounts for only 11 % of the non-structural variability in individual incomes and 20 % of the variability in household incomes. In short, the vast majority of the variation in incomes is attributable not to spatial location, but to other causes. This finding is line with Berthoud's observation that only a small minority (10 %) of the total variance in net household incomes is explained by their spatial sorting across postcode sectors.

Finally, Table 2 reports the relative contributions of each spatial scale to the overall level of spatial sorting observed at postcode-unit level. To do so it is assumed,

[1] The fewer responses per spatial unit, and the fewer the income bands used to classify their incomes, the fewer available degrees of freedom there are, and the greater the 'structural' component of between-area income variation will be. This structural component can be estimated by randomly assigning responses (adult and household incomes) across spatial units, subject to matching the marginal constraints captured in the Census Rehearsal of the number of respondents per income band and the number of respondents per spatial unit, and then measuring the resulting between-area variation. The 'structural' variation is taken to be the 100-run average of such measures.

following Berthoud, that the spatial units under consideration are nested, such that their spatial contributions to the overall variation in incomes are, for all practical purposes, independent of one another. The outcome is that for both adult and household incomes, once inherent structural variability is taken into account, the key scale for the spatial concentration of incomes is the enumeration district, which accounts for around two-fifths of the overall spatial concentration of incomes. A further two-fifths is split between postcode units and postcode sectors, whilst districts (a reasonable surrogate for any labour market effect) account for only one-twentieth.

Disentangling Population Composition and Spatial Context

The results presented in Table 2, if read naively, suggest that incomes are, at least to some extent, 'spatially determined'. But, as Mitchell (2001) points out, people within an area do not all live their lives within the same fixed spatial hierarchy. This brings into question precisely what any 'area' effect is really capturing. In addition, it is clear that observed area effects might simply be an artefact of failing to control sufficiently for the influence of local population composition. Therefore, to further evaluate the relative importance of individual and ecological factors on income levels, a range of multi-level models were fitted to the Census Rehearsal data, based on a set of individual-level regression models introduced in "Small-area income estimation strategies" section of this paper, but with random intercepts for each of person, household, enumeration district and ward level. Reduced versions of the basic regression model were fitted to sub-sets of the Rehearsal data, split along the dimensions of persons/earner/non-earner and persons/adults/household representatives. House prices were estimated, for England and Wales only, using data supplied by Experian based on 3-year average postal sector house prices, disaggregated by household type (detached, semi, terraced, flat). The relative importance of house price effects was investigated by adding and removing measures of house price at the household level (mean house price for houses of the same type in that postal sector) and enumeration district level (mean price of houses in that ED). For households in Leeds council-tax band data were also available, in the form of mean council-tax band by postcode unit. Treating council-tax band as an ordinal variable, council tax was added at household and enumeration district level as for house price.

The vast majority of unexplained between-adult variation in incomes was found to arise at the individual (85–90 %) or household (10–15 %) level, leaving less than 1 % attributable to location by district, ward or enumeration-district. For this reason the inclusion of house value and area affluence (mean income), although statistically significant at the enumeration district level, made very little overall impact on levels of unexplained variation. Estimated house prices were found to be statistically non-significant at every level, although this might possibly be an artefact of the way prices had to be estimated from ward-level averages. These findings accord with research by McCulloch (2001) who found that ecological associations of various indicators of individual adversity with a census-based indicator of deprivation were largely, if not entirely, accounted for by household and individual characteristics; and with Gibbons et al. (2010) who found when analysing UK earnings data that labour market effects make a very small contribution (<1 %) to the overall variation in observed incomes. The implication is that any small-area estimation strategy that adequately controls for

the local population composition should be able to provide reasonable estimates of local incomes.

Distributional Aspects of Income Segregation

One final aspect of the spatial variability of incomes remains to be considered – the extent to which spatial segregation varies across the income distribution. In other words, is it the rich or the poor who are most spatially segregated? The Census Rehearsal directly captured the distribution of individuals by income band. This information was supplemented by placing households into income bands, arbitrarily defined as deciles of the household income distribution. To do so, imputed respondent incomes were summed at household level. Using the percentage of respondents (or households) in each income band it is possible to compute a measure of segregation – the index of dissimilarity (D) – for each income band, for each type of spatial unit. For this index, a value of 1 indicates complete segregation (all of the persons associated with that band live in the same area), whilst a value of 0 indicates complete mixing (persons associated with that income band are equally spread across all areas). The outcomes are illustrated in Fig. 2, for a range of spatial scales. Several points of note emerge. First, regardless of income band, the level of spatial segregation increases as the size (population) of the spatial unit decreases. Second, at all spatial scales the vast bulk of segregation of incomes is associated with those persons (and households) in the highest income bands, most notably those in the top (open) income band. In contrast the pattern of segregation for those in paid employment (not shown) is more clearly U-shaped, with high observed levels of segregation for those earners reporting themselves as being in one of the two lowest income bands, decreasing for those in middle income bands before rising again for those in the top income bands. Similar patterns of segregation are found within each district, although the greatest levels of segregation amongst top-band incomes were found in Leeds and Bournemouth.

To further disentangle the influence of spatial scale on income segregation it is possible, following Voas and Williamson (2000), to calculate the contributions to dissimilarity associated with each spatial scale (again assuming spatial nesting). The result, as Fig. 2 shows, is that the vast majority of segregation by income takes place at the level of the enumeration district or postcode unit. However, the role of district of residence becomes more significant in influencing the spatial segregation of those with the highest incomes. Similar findings apply for the distribution of earner incomes. Had the Rehearsal included coverage of parts of London or the South East it is possible that the role of district-level geographies would have become even more influential, at least for the highest incomes; but as two-thirds of the segregation of those on the highest incomes is explained by social sorting across postcode units and enumeration districts, it is unlikely that even this ‘capital region’ effect would dominate the observed residential sorting effects.

Small-Area Income Estimation Strategies

Analysis of the Census Rehearsal has allowed us to establish that average incomes are spatially highly variable; that this spatial variability is most pronounced for the smallest

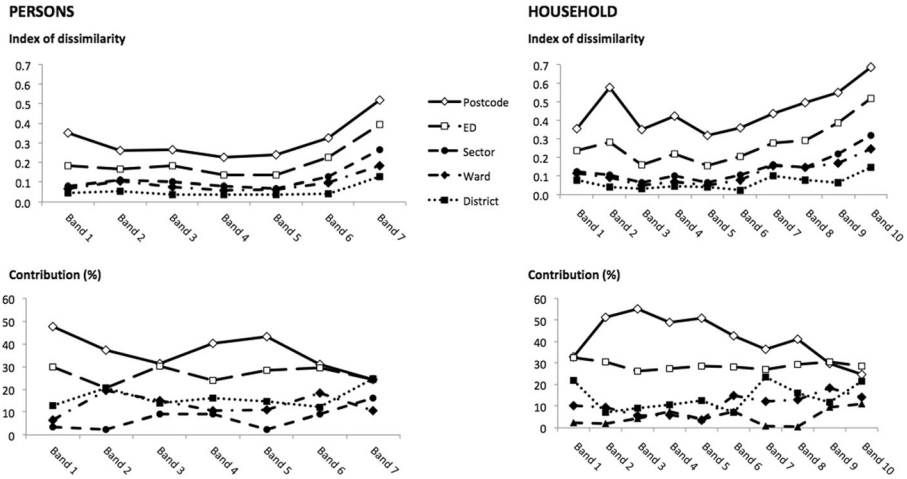


Fig. 2 Income segregation. NOTES: Income bands for persons defined as per Fig. 1; Income bands for households defined as deciles of the household income distribution

spatial units; and that it is those persons and households with the highest incomes that are most spatially segregated. It is now time to turn our attention to the most effective methods of capturing this spatial variability. Four main approaches to small-area income estimation are considered: the use of proxy indicators; geodemographic classification; unit-level imputation and ecological regression. In this section the nature and limitations of these competing strategies are reviewed. Details are also provided of how each strategy has been implemented for this paper in order to provide a series of rival small-area estimates. Subsequently ("The relative efficacy of alternative estimation strategies" section) data from the Census Rehearsal will be used to evaluate these estimates.

Proxy Indicators

Lee et al. (1995) provide a comprehensive review of the 'indicators' of area deprivation used in the UK, ranging from car-ownership (univariate) through to the well known Townsend index (multivariate). Of course, as Gordon et al. (2000) have ably shown, lack of income does not directly equate to material deprivation, in part because of lag effects, and in part due to life-course effects (the retired typically have lower incomes but a lifetime of asset accumulation). Even so the spatial patterns of low income and material deprivation are strongly linked (Gordon and Forrest 1995), meaning that deprivation indices still offer potentially useful proxies for the relative rankings of small-area incomes. At the other end of the income spectrum, a multivariate census-based 'wealth index' has been proposed by Green (1994), whilst Dorling et al. (2007) provide wealth estimates based on the type, tenure and price of local housing.

As Tunstall (2005) has observed, one problem with all proxies for income is that interactions between variables may well vary over space (e.g., the link between car ownership and income differs between central London and rural Scotland). This applies equally to a more recent proxy indicator of income deprivation (DCLG 2011), based upon specially commissioned welfare claimant (administrative) data, since there are

known socio-demographic and spatial variations in the uptake of benefits (c.f. Bramley et al. 2000). This proxy has the added disadvantage of focusing only on the lower tail of the income distribution, when high incomes may be the main drivers of between-place differences in mean income. More directly relevant proxy administrative data do exist (e.g., tax records), but they are not yet readily available for use in the construction of small-area estimates, at least within a UK context; often lack full population coverage; and often have their own data quality issues. Similar issues apply to proxy data held by commercial companies, such as credit scores, supermarket spend or bank account turnover.

Drawing upon this literature a wide-range of potential and previously adopted univariate and multivariate surrogates for income were identified for evaluation using Census Rehearsal data. A full list of those considered is provided in Tables 3 and 4. It is important to reiterate that a number of these potential surrogates were originally devised for other purposes. Unfortunately lack of claimant data for the year of the Census Rehearsal debarred the evaluation of claimant-based proxies.

Ecological Regression

Small-area income estimation has been the recent focus of three major European research projects (EURAREA; AMELI and SAMPLE) as well as a US National Research Council report (NRC 2000). The EURAREA project compared performance of model- and design-based (survey-based) estimates, and concluded that model-based estimates were better, most noticeably so for the smallest spatial units (The EURAREA Consortium 2004). A subsequent European project, AMELI, extended the work of EURAREA to consider the robustness of these findings to outliers and alternative survey sampling schemes, and to outcomes other than mean incomes, such as the poverty rate and Gini coefficient (Lehtonen et al. 2011). At the same time, building upon the work of Tzavidis et al. (2008; 2010), another European project (SAMPLE) explored the role of M-Quantile regression in estimating distributional aspects of small-area incomes (Casarosa et al. 2011). Unfortunately both AMELI and SAMPLE confined themselves to considering scenarios in which survey samples are available from all of the areas being estimated. Estimates for small areas unrepresented in the sample survey necessarily have to be derived using what Heady and Ralphs (2004) describe as model-based (synthetic) estimates. In this context a wide range of model-based approaches are possible (c.f. Rao 2003), the simplest of which is ‘ecological regression’, also known as the Fay-Herriot model, an area-level model in which the relationship between survey and area aggregates for sampled areas are used to estimate incomes for areas lacking sample data.

Ecological regression has been found to provide effective small-area estimates for six different European countries (The EURAREA Consortium 2004); and is the approach currently adopted by the UK Office for National Statistics in producing a series of small-area income estimates (Longhurst et al. 2004a; White et al. 2009; Bond and Campos 2010), most recently updated to provide estimates not only of mean incomes but also of poverty (Fry 2011). Separate estimates are produced for four types of mean household income: gross; net; net equivalised before and after housing costs. Regional dummy variables are used to relax the assumption that nationally observed relationships between covariates apply locally. A priori this method is likely to work

Table 3 Efficacy of alternative small-area estimation strategies for the 1999 Census Rehearsal enumeration districts

		Mean income per															
		Resident		Adult		Household		Household Head		Main wage earner		Equivalised (per capita)		(McClem.)		(OECD)	
		Coefficient of Determination (Coefficient of Variation) (%)															
Univariate surrogates																	
% 16–74 year olds in NS-SEC 1 or 2	PNSSEC12	81	(36)	83	(21)	65	(22)	76	(17)	75	(16)	80	(34)	83	(19)	82	(16)
% 16–74 year olds in Social Class I or II	PHICLASS	77	(46)	79	(29)	64	(22)	72	(19)	72	(19)	77	(44)	79	(26)	79	(22)
% 16–74 year olds in Social Class IV or V	PLOCLASS	54	(114)	51	(86)	53	(67)	55	(72)	55	(70)	55	(110)	57	(86)	58	(82)
% 16–74 year olds in NS-SEC 6 or 7	PNSSEC67	41	(118)	35	(89)	42	(68)	39	(74)	41	(72)	43	(115)	42	(89)	43	(84)
% households socially rented	PSOCIAL	35	(111)	36	(90)	49	(80)	46	(81)	48	(80)	36	(109)	41	(91)	43	(89)
% households with no car	PNOCARS	33	(121)	42	(94)	66	(75)	60	(79)	62	(78)	33	(117)	45	(93)	49	(89)
% economically active unemployed	PUNEMPEA	31	(76)	33	(64)	39	(64)	38	(62)	39	(62)	31	(74)	36	(66)	37	(66)
% owner-occupied households	POWNOCC	31	(131)	38	(94)	47	(53)	49	(63)	50	(61)	30	(126)	40	(91)	42	(83)
% 16–74 year olds in NS-SEC 8	PNSSEC8	24	(77)	29	(63)	18	(55)	27	(57)	25	(56)	21	(74)	25	(62)	25	(61)
% working age unemployed	PUNEMPAG	22	(67)	24	(57)	34	(61)	30	(58)	32	(58)	23	(66)	27	(60)	29	(60)
% households with 2+ cars	P2CARP	19	(56)	27	(41)	56	(30)	45	(30)	49	(30)	19	(54)	31	(40)	35	(37)
% detached households	PDETACHD	10	(73)	16	(57)	31	(48)	29	(46)	30	(47)	10	(71)	16	(58)	19	(56)
% housholds with 3+ cars	P3CARP	9	(55)	11	(48)	32	(47)	22	(44)	25	(44)	9	(54)	15	(48)	17	(48)
% detached or semi-detached households	PDETSEM	2	(119)	7	(87)	24	(53)	18	(61)	20	(59)	2	(115)	8	(85)	10	(79)
% households in flats	PFLAT	5	(92)	1	(75)	5	(69)	1	(68)	2	(68)	4	(89)	0	(76)	0	(74)
Multivariate surrogates																	

Table 3 (continued)

		Mean income per															
		Resident		Adult		Household		Household Head		Main wage earner		Equivalised (per capita)		(McClem.)		(OECD)	
		Coefficient of Determination (Coefficient of Variation) (%)															
Wealth Index (after Green)	GREENWTH	55	(39)	59	(28)	68	(27)	64	(23)	67	(23)	56	(37)	63	(26)	65	(25)
Poverty Index (after Green)	GREENPOV	44	(110)	54	(85)	60	(67)	63	(72)	63	(70)	44	(106)	54	(84)	56	(80)
Townsend Index	TOWNSEND	43	(129)	48	(98)	58	(74)	58	(81)	59	(78)	43	(124)	51	(97)	54	(92)
Carstairs Index	CARSTAIR	42	(84)	45	(67)	59	(63)	55	(63)	57	(62)	43	(82)	51	(69)	54	(67)
Breadline Index	BREADLINE	21	(79)	22	(63)	27	(57)	25	(58)	26	(57)	22	(77)	23	(64)	24	(62)
Synthetic estimates: unit-level																	
Occupational means (after Dale)	DALINCM	88	(15)	83	(25)	58	(60)	71	(38)	70	(43)	87	(15)	84	(57)	82	(37)
Occupational means (after Lee)	LEINCM	86	(15)	80	(26)	59	(60)	71	(39)	70	(43)	86	(16)	82	(58)	80	(38)
Individual-level regression (project team)	R_IND	86	(32)	72	(46)	41	(72)	55	(56)	53	(59)	85	(34)	74	(70)	70	(55)
Household-level regression (project team)	R_HH	62	(41)	74	(18)	81	(42)	86	(15)	86	(20)	63	(38)	75	(41)	77	(17)
Household-level regression (after Davies et al.)	DAVIESHH	35	(100)	47	(62)	74	(21)	65	(39)	68	(30)	36	(97)	50	(24)	54	(40)
Synthetic estimates: area-level																	
Three-factor ecological regression	ECOLOGICAL	88	(12)	88	(10)	85	(13)	86	(13)	87	(12)	87	(12)	87	(12)	87	(12)
Principal components (96 categories)	P96INC	85	(13)	–	–	82	(14)	82	(14)	83	(14)	85	(13)	83	(13)	83	(13)
Principal components (40 categories)	P40INC	79	(16)	–	–	76	(16)	76	(16)	77	(16)	78	(16)	76	(15)	76	(15)
Principal components (10 categories)	P10INC	66	(20)	–	–	66	(20)	66	(20)	67	(19)	65	(20)	65	(18)	66	(18)
Percentile splits (96 categories)	M96INC	57	(22)	–	–	68	(19)	64	(20)	66	(19)	57	(22)	60	(20)	61	(20)

Note: − =Not calculated

Table 4 Multivariate income surrogates (indices)

CARSTAIR	Combination of: % male unemployment, % residents in overcrowded households, % residents in households with no car and % households with household head in Social Class IV or V
TOWNSEND	Combination of: % economically active unemployed, % overcrowded households, % households with no car and % of households not owner-occupied
BREADLINE	Weighted combination of: no. of households with no car, households not owner-occupied, lone parent households, economically active residents in Social Class IV or V, households with 1 or more persons suffering from limiting long-term illness and number of unemployed
GREENPOV	Combination of: % economically active unemployed, % households with no car, % rented households, % of working age population ecomically inactive
GREENWTH	Combination of: % households with 2+ cars, % persons aged 16–74 in NS-SEC 1 and % adults with high educational qualifications

All variables used in multivariate surrogates standardised and/or normalised as appropriate

well, given that ONS reports that their models capture around 90 % of the spatial variation in mean incomes. From a user perspective the main problems with this approach are that the method is not publically replicable due to data access issues, restricting estimates to only those years and geographies for which ONS choose to produce estimates, and that the model for each inter-censal update uses a different set of covariates, rendering time-series analysis problematic.

An interesting alternative to the standard ecological regression has been adopted by Dorling et al. (2007) and Famhy et al. (2011). In this mixed-level variant survey-based household-level logit-regression models provide model coefficients used to convert small-area aggregate counts into estimated counts of the poor (and rich). As implemented the focus has been the creation of small-area estimates of material deprivation, but the approach is readily extendable to provide small-area estimates of the % of the population falling below a given percentile of the national income distribution. The outcomes appear plausible. However, Heady and Ralphs (2004) considered and dismissed this approach as suffering from unacceptable bias due to the ecological fallacy involved in assuming that relationships observed between households hold for their area aggregate counterparts. In any case the performance of estimates based upon a standard ecological regression is likely to provide the limiting case of how well such a mixed-level modelling approach can perform.

Lacking appropriate external data, an ecological regression model, predicting small-area incomes given area aggregates, was constructed directly from Census Rehearsal data. To mitigate against the dangers of over-fitting, variants of this model were evaluated using partitioned data and limited sets of predictor variables. Little difference was found. However, to err on the side of caution the results presented in this paper relate to ecological regressions restricted to only three area-level covariates, albeit fitted to the entire dataset and with the precise combination of factors used tailored to suit the measure of income being predicted (see Table 5 for details).

Table 5 Synthetic estimators

Estimation strategy	Definition
Geodemographic classification[a]	
M96INC	Weighted Geodemographic classification with areas divided into subgroups after classification into above or below median on the following measures: % households with children, % households with female head, % lone parent households, % economically inactive females and % persons of pensionable age; and into thirds on the basis of % households with no car [96 categories]
P10INC	Geodemographic classification based on principal components analysis of normalised: % economically active unemployed, % households with>1 persons per room, % households with no car; % persons non-white; % persons of pensionable age; % lone parent households; % socially rented households; % privately rented households; % overcrowded households; % households with 2+ cars; % persons aged 16–74 in Social Class I or II; % persons aged 16–74 in Social Class IV or V; % households with<0.5 persons per room; % of economically active working full-time; % working-age population aged 18 and over registered in full-time education, % population aged<16, % households with 4+ residents; % households with 7+ rooms; % households with 1 or 2 rooms; % detached households, % terraced households; % households in flats [10 clusters]
P40INC	As P10INC, but with 40 clusters
P96INC	As P10INC, but with 96 clusters
Ecological regression	
ECOLOGICAL	Area-level income predicted given: the best performing three area-level co-variates. Models fitted separately for each spatial unit and income measure. All models include % of 16–74 year olds in NS-SEC 1 or 2. All except the models for mean income per Household and Main Wage Earner include the poverty index GREENPOV (see Table 4). The models for mean Resident, Adult, Per Capita and McClements-equivalised Household incomes include the % of households in flats. Other predictors used by one or more models are the wealth index GREENWTH (see Table 4), the % economically active unemployed and the % of households with no car.
Sub-group mean	
LEEINC	Income imputed given mean income for population sub-group defined by: SOC 2000 minor group (81 categories); Economic activity (Child, not applicable, employed full-time, employed part-time, self-employed, unemployed, retired, other inactive) [maximum of 649 valid sub-groups]
LEEINC2	Income imputed given mean income for population sub-group as for LEEINC, but with greater disaggregation of economic activity (Missing, child, not applicable, employed full-time, employed part-time, self-employed, unemployed, student, retired, permanently sick, other inactive) [Maximum of 731 valid sub-groups]
DALINCM	Income imputed given mean income for population sub-group defined by: Sex (male, female); SOC 2000 minor group (81 categories); Economic activity (Missing, employed full-time, employed part-time, self-employed, other); Age (Missing, 0–15, 16–19, 20–29, 30–49, 50+) [Maximum of 4860 valid sub-groups]
Unit-level Regression	
R_IND	Person-level income ($INCOME^{0.5}$) predicted given:

Table 5 (continued)

Estimation strategy	Definition
	mean income by SOC2000 unit; mean income by Industry category, age, age^2, residents, residents^2, rooms and cars plus dummy variables for sex, white, full-time student, married, Single/Widowed/Divorced, Long-term ill, No qualifications, GCSE or equivalent, A levels or equivalent, Undergraduate degree or equivalent, employed full-time, employed part-time, self-employed, unemployed, retired, permanently sick, other economically inactive excluding pensioners and students, Semi-detached, terrace, flat, caravan, privately rented, social rented, employed manager or supervisor and district of residence. Children (<1 16 years old) assumed to have no income.
DAVIESHH	Household-level income (HHINC$^{0.5}$) predicted given: number of children (<16 years old) and dummies for female, age group (0–19, 20–29, 30–39, 40–49, 60–64 or 65+), owner-occupied household with no car, rented household with no care, rented household with 1+ car, unemployed head, economically inactive head, head in Social Class IV or V,, one person household, 1+ dependent children at home, lone parent household and district of residence
R_HH	Household-level income (HHINC$^{0.5}$) predicted given same set of predictors as R_IND, but based only upon head of household's characteristics

[a] For use as an income surrogate, average income for each category of the geodemographic classification is calculated using Census Rehearsal data aggregated to appropriate spatial scale. This process was repeated separately for each income measure

Geodemographic

Geodemographic classification offers an alternative area-level synthetic estimator. Areas are grouped into clusters based upon their similarity across a range of census (or other) variables (Vickers and Rees 2006). If survey data, although stripped of much geographical information, is released tagged with the type of cluster that each survey respondent comes from, then it is possible to calculate the mean income of each cluster type; and to assume that this holds for all areas of that cluster type. For example Daniel and Bright (2011) examine the variation in differing aspects of wealth, including various elements of income, by geodemographic area-type (ONS Output Area Classification), although they do not take the next step of converting this into a set of small-area income estimates. The geodemographic approach is an example of a classic area-level synthetic estimator (Gonzalez and Hoza 1978). The main shortcoming of this approach is the diversity of areas represented by a given area 'type' (Voas and Williamson 2001), meaning that two areas with broadly similar incomes can end up allocated to different clusters, and hence estimated to have markedly different incomes.

Commercially available geodemographic classifications were too costly to procure for this project, and did not in any case cover all of the geographical levels of interest. Nor was direct replication of these commercial classifications possible using Census Rehearsal data alone, as they invariably draw upon additional non-census 'lifestyle' data that are not publically available. Potentially the rehearsal data could have been aggregated to 2001 Census Output Areas, in order to utilise the ONS 2001 Census Output Area Classification (OAC), but the required postcode to Output Area lookup

table was unavailable. Instead, four new geodemographic classifications were produced (see Table 5). A first classification was based on sub-dividing areas into above and below median groups along a range of demographic and socio-economic dimensions, as suggested by Voas and Williamson (2001). The remaining three classifications, following Openshaw and Wymer (1995), were based upon a Principal Components Analysis of a wide range of Rehearsal variables, clustered (using k-means) into 10, 40 and 96 area types to reflect the number of categories typically available in alternative classifications – mirroring the process used to produce the ONS 2001 Census OAC.

Unit-Level Imputation

The term 'imputation' is used here as short-hand to describe the set of synthetic estimation techniques which involve the assignment (imputation) of income to individuals and households on the basis of their known characteristics, typically via a unit-level regression model, followed by the summation of these imputed incomes to area level. In practical terms, the regression model, once fitted, can be applied either directly to unit-level records with unknown incomes – e.g., respondents to the UK 2011 Census; or, when the unit-level regression model is relatively simple, indirectly using published tabulations of respondent characteristics. Not included here are synthetic estimators which, in the absence of unit level records for an area, reduce to a simple area-level regression model.

Both Lee et al. (1995) and Dale et al. (1995) have used the mean income recorded in a national survey, given occupation and other relevant factors, to impute the income of each person of equivalent type captured locally by the census. Birkin and Clarke (1989) have gone further, acknowledging the problem that, for a given occupation, inner city workers are likely to be earning less than comparable workers in more affluent suburbs. For this reason they have proposed adjusting the imputed income to reflect the nature of the local occupational mix relative to the national one. For example, if an area contains more than the usual national share of high earning occupations, imputed incomes for all occupations in that area are inflated. Alternatively Davies et al. (1997) mooted the possibility of imputing census respondent incomes on the basis of a unit-level regression model. A further variant on imputation has been adopted by Anderson (2013), who produced small-area household poverty estimates by weighting government survey data to fit univariate local area constraints given by the census. In all cases the accuracy of these small-area estimates rests heavily upon the assumption that the factors taken into account during the imputation or reweighting process capture all of the processes driving spatial variations in incomes; and that the relationship between these factors does not itself vary spatially.

Table 5 provides details of the various imputation strategies evaluated for the purposes of this paper. Ideally, mean occupational incomes would have been derived from an external source. However, occupation in the Census Rehearsal was coded using SOC2000, whereas all other government social survey data available for the same time period were coded occupation by SOC90. As the two coding schemes do not map on to each other, the only solution was to estimate SOC2000 means from the Census Rehearsal itself, having first replaced reported income bands with imputed income values. This, in combination with the large number of sub-group means required (approximately 650 for Lee and 4800 for Dale), gave rise to a substantial danger of

'over-fitting'. As a precaution, therefore, the Census Rehearsal was partitioned. Sub-group means calculated for one half of the dataset were then used to impute individual incomes in the other half. Partitioning was found to make no significant difference to the outcomes obtained. Therefore this paper presents only the results for models based on unpartitioned data.

Three regression models were evaluated for imputation purposes (detailed in Table 5). An adult and a household model were devised on the basis of extensive experimentation with data from the 1998 FRS. Worthy of note in passing, the impact of house price, as measured by council-tax band, was found to be statistically non-significant. An innovative element in both models is the use of mean occupational and industry incomes in place of more conventional dummy variables. At the household level the model published by Davies et al. (1997) was also replicated. The use of occupational mean incomes as predictor variables, allied with the possibility of non-response bias in the Census Rehearsal, meant all regression models had to be fitted using Rehearsal data. The possibility of over-fitting was once again addressed by data partitioning, with the same outcome (no difference of note).

The Relative Efficacy of Alternative Estimation Strategies

To date, a lack of small-area income data has limited evaluation of the various estimation methods outlined in "Small-Area Income Estimation Strategies" section. In the main this has led to estimates being validated against income proxies or alternative estimates of limited comparability (c.f. Longhurst et al. 2004b). The one notable exception to this is EURAREA's benchmarking of alternative model-based estimates against directly equivalent empirical data (The EURAREA Consortium 2004). This served to show that ecological regression models perform well, but left unanswered the question of how they fare relative to proxy, geodemographic and imputation-based estimates. In addition no published validation has paid more than the scantest attention to the impact of scale, or to the measure of income used. Yet the spatial segregation of incomes varies with scale and type of income (c.f. "The spatial variability of incomes" section), suggesting that the optimal estimation strategy may vary according to the areal unit and income measure for which small-area estimates are required.

Efficacy in Capturing Spatial Variation in Mean Incomes

Table 3 reports how well the various small-area estimation strategies capture the mean incomes of enumeration districts in the Census Rehearsal. Two measures are used. The coefficient of determination (r^2) reports the percentage of the spatial variation in mean income captured by each estimate. In contrast, the coefficient of variation (CV) reports the root mean square error (standard deviation) of the estimate error, expressed as a % of the value being estimated. Since surrogate measures such as the % unemployed do not provide direct estimates of area incomes, indicative CVs were calculated once the surrogate estimates had been scaled to the observed range of area incomes using range standardisation. A perfect estimate would have an r^2 of 100 % and a CV of 0 %. ONS (2013) deems estimates with a CV of less than 20 % to be 'fit for purpose'.

The results presented have been found to be insensitive to (i) Geography (1991 or 1999 boundaries); (ii) Method of imputing top-band mean income; (iii) degree of 'missingness' in the Rehearsal sub-set being analysed; (iv) use of a full or partitioned Rehearsal dataset to estimate model parameters. As a result, concerns over the possible impacts of the Modifiable Areal Unit Problem, the actual value of income for individuals within income bands, question-specific non-response bias and data over-fitting can be dismissed. Full discussion of the influence of scale is deferred for the moment, but in summary similar findings apply regardless of spatial scale.

As can be seen from Table 3, by far the most effective univariate surrogate for mean ED income, across the full range of income measures, is PNSSEC12, the % of the economically active population in NS-SEC categories 1+2 (Managerial and Professional occupations). This surrogate enjoys the lowest CVs and captures 65 to 83 % of the observed spatial variation in mean enumeration district incomes, dipping below 75 % only for total household income. Indeed, total household income is the only income measure for which a univariate surrogate was found with a higher r^2 (% no car households); and even here PNSSEC12 had a much lower CV. Of the multivariate surrogates considered the Wealth Index proposed by Green (1994) [GREENWTH] offers the best performance across all measures of income; but even this multivariate index is out-performed on all occasions by PNSSEC12, both in terms of r^2 and CV.

In general the synthetic estimation strategies considered are more successful at competing with PNSSEC12. Both estimates based upon the imputation of occupational means performed well, generally matching and sometimes marginally out-performing PNSSEC12. The same is true of the individual and household level regression-based approaches, in terms of both r^2 and CV, provided that performance is considered only in relation to the relevant individual or household income measures. Most notably, a household-level regression model (DAVIESHH) was found to provide a better small-area estimate of mean household incomes than any other surrogate or unit-level synthetic estimation strategy considered.

Turning to area-level synthetic estimates, the 96-fold geodemographic classification derived via principal component analysis showed only marginal gains compared to its 40-fold equivalent, but significantly out-performed the 96-fold classification based upon quantile splits. It also out-performed PNSSEC12, particularly with regards to CV and household incomes. However, the single best performing estimation strategy, across all types of income measure, proved to be a three-factor ecological regression model, in terms of both r^2 and CV. This estimation strategy matches whichever of the other unit- and area-level synthetic estimation strategies perform best for a given type of income measure, and significantly out-performs the various univariate or multivariate surrogates across all income types.

Of course, it is possible that an estimation strategy may offer good overall performance, but exhibit systematic patterns of local failure. To test for this, those estimation strategies identified as most successful were examined further to see if their performance varied between districts, repeating the analysis summarised in Table 3, but separately for the set of EDs located within each district. This further analysis reconfirmed the overall picture, with the performance of most estimates varying little between districts. Those strategies that proved most sensitive to district effects were the multivariate index, GREENPOV, and the geodemographic classification, P10INC. Both provided notably poorer estimates for the mean income of EDs in Bournemouth

(a seaside town with a relatively elderly population) and for some of the more rural districts.

The Influence of Scale

It has already been indicated that the findings from Table 4 are reasonably robust to changes in the spatial unit for which estimates are required. The basis for this claim is summarised in Fig. 3, which both confirms and qualifies this observation. Figure 3 plots r^2 and CV for a selection of the best performing small-area estimation strategies, at a variety of spatial scales. The larger the spatial unit, the greater the spatial variation in income that is captured, and the lower the CV, with the notable exception of only one strategy (GREENWTH). The rank order of the estimate performances also proves to be remarkably resilient to changes in spatial scale. What is influenced by the change in spatial scale is the relative performance advantage of the differing strategies, with the performance gap narrowing as the spatial units increase in size. It is also noticeable that the performance of synthetic estimation strategies involving regression or occupational means are far less sensitive to changes in scale than the other approaches considered. Finally, it should be noted that for the very smallest spatial units (unit postcodes), the performance of the ecological regression approach degrades, falling below that achieved by the most successful scale invariant strategies.

Efficacy at Accurately Ranking Areas

If it is assumed that a linear relationship exists between observed and estimated small-area incomes, then the percentage of spatial variation captured by the various strategies should be a good guide to their overall performance. This assumption is examined, in Fig. 4, for the best performing income surrogate (PNSSEC12), the two best performing unit-level synthetic estimation strategies (R_IND and LEINCM) and the best area-level synthetic estimation strategy (ecological regression). This reveals that the general relationship between PNSSEC12 and mean individual income is broadly linear across the range £0-£300 per week, arguably breaking down only for the six enumeration districts (out of 651) with mean incomes above £300. Estimated mean ED incomes synthetically estimated using individual-level regression or sub-group means are slightly less linear. In both cases predicted incomes rise less slowly than observed incomes. As a result, divergence from mean income is again greatest for the most affluent EDs.

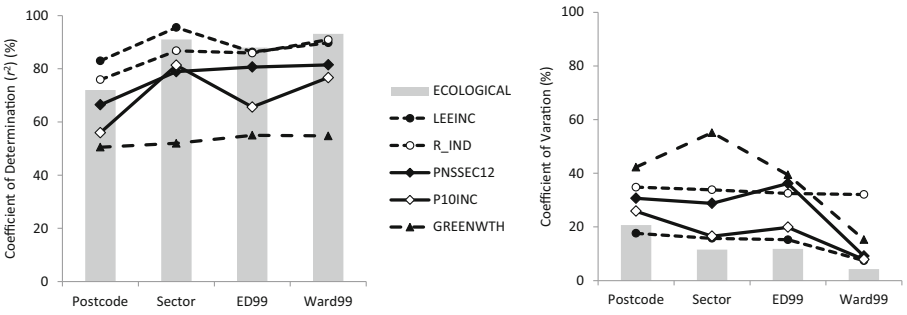


Fig. 3 Variation in the efficacy of estimation strategy with scale

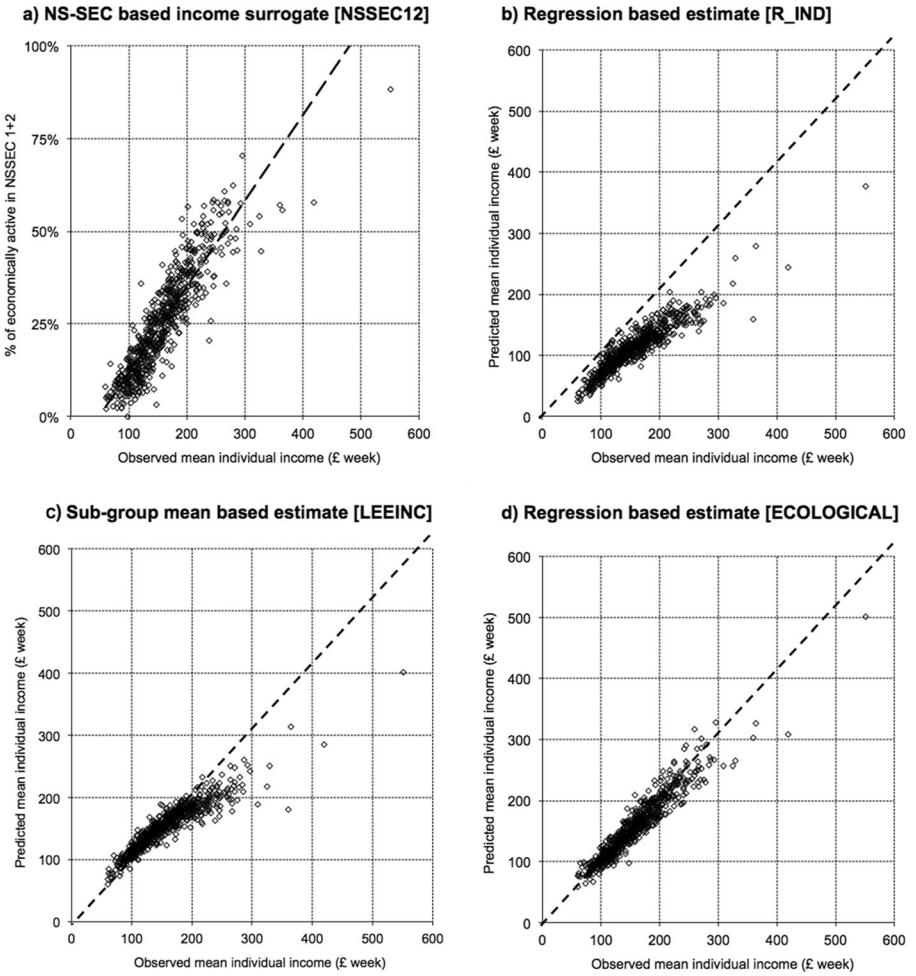


Fig. 4 Performance of selected small-area income estimates

In addition, the regression-based results appear to consistently under-predict mean income. In contrast, estimates derived via ecological regression appear equally likely to under- or over-estimate mean incomes, with the exception of the same few high-income EDs under-predicted by PNNSSEC12. In contrast to PNSSEC12, however, the estimates from ecological regression are more tightly clustered around the true (observed) values, reflecting its lower CV.

Collectively these findings suggest that, especially for the more affluent enumeration districts, differential selection by income appears to persist, even within narrowly defined population sub-groups. This effect remains and is further exaggerated if modelled income means are used to impute top income-band values in the place of the FRS-based means. Even so, the relatively linear relationship found between observed and predicted income for the imputation and ecological regression estimation strategies suggests some scope for producing improved estimates by simple rescaling to control totals.

Table 6 Accuracy of ED rankings by mean individual income

		Surrogate/Estimate			
		% NSSEC 1+2 [PNSSEC12]	Individual Regression [R_IND]	Sub-group mean [LEEINC]	Ecological Regression [ECOLOGICAL]
		% ranked in same decile as income			
Decile	1	71	66	74	80
[low income]	2	46	34	40	52
	3	32	40	35	43
	4	32	26	37	40
	5	25	34	39	37
	6	17	28	45	30
	7	26	28	43	31
	8	23	35	48	46
	9	28	51	57	60
[high income]	10	55	77	82	82
Overall		36	42	50	46
Within±1 decile		82	84	89	92

Rescaling, however, will not change the relative rankings of EDs by estimated income; and it is the ranking of areas by income that is often of most policy relevance. For this reason Table 6 reports the accuracy of the ranking of enumeration districts by income decile using the same estimation strategies as Fig. 4. On average, all four approaches misclassify between 50 and 65 % of all EDs. However, this figure more than halves if the focus is restricted to correctly identifying EDs in the lowest and highest income deciles. Of the four approaches, the best-performing for ranking purposes are ecological regression and unit-level synthetic estimation using sub-group means (LEEINC). Of these two, ecological regression is the more successful at identifying those areas with the highest and lowest mean incomes, whilst sub-group means offer superior ranking mid-table and, therefore, overall. The difference, however, is marginal.

Conclusion

Researchers and National Statistical Institutes around the world continue to grapple with the problem of estimating incomes for small areas. Exploiting a unique resource, a UK Census Rehearsal that captured information on income for large spatially-contiguous blocks of population, this paper offers a number of contributions to the debate. First, the nature of the problem being tackled has been explored further, through a detailed appraisal of the way in which incomes vary spatially. Second, the performance of competing strategies for the estimation of small-area incomes has been assessed, uniquely comparing simultaneously proxy, geodemographic, imputed and ecological estimates; and validating all of these estimates against direct rather than indirect observations of the true values.

Unsurprisingly the variability in mean area incomes has been shown to vary inversely with the size of spatial unit, with the smallest spatial units displaying the greatest variability in mean incomes. Mean household incomes are also generally more spatially variable than mean individual or equivalised-household incomes. More surprisingly, the spatial variability in incomes has been found to be almost entirely due to the effects of residential sorting. Wage variations across labour markets have been shown to play at best a minimal role, with the possible exception of a London/South East effect (an area not covered by the Rehearsal data). This residential sorting is stronger for households than for individual incomes, indicative of the pooling of household resources to meet housing costs, and operates most notably at the Enumeration District level (i.e., blocks of c. 150–200 households). Most crucially, in relation to small-area estimation, this residential sorting has been found to be strongest amongst those persons and households with the highest incomes. It has also been found to be explicable almost entirely in terms of individual and household level compositional factors.

Of the various estimation strategies considered the simplest of all, a univariate proxy based on the % of the economically active in the highest social classes, proved to be surprisingly effective, performing on a near par with other far more complex approaches and capturing around 80 % of the spatial variation in mean incomes at enumeration district level. The biggest shortcoming of this univariate proxy was its failure to accurately reflect spatial variations in household incomes. Other drawbacks include the obvious lack of a direct estimate of income levels or deprivation rates; inconsistency in the way in which social classes have been classified over time; and the danger that its effectiveness as a proxy will change over time as the levels and distribution of societal prosperity changes.

Equally effective at capturing spatial variations in income were the best performing of the various unit-level synthetic estimation strategies considered. These all involved imputing the income of Census Rehearsal individuals or households, using individual- and household-level regression models or occupational mean incomes, and summing to find mean area-level incomes. An advantage of these unit-level approaches it that they deliver estimates of the full local income distribution, not just point estimates such as the mean or median; offer the potential to cross-classify income with other variables of interest, such as ethnicity; and allow estimates of deprivation (or affluence) to be derived based on user-specified income thresholds. On the other hand, imputation strategies like these require access to sizeable samples of survey units for each estimation area, if not to a full set of census records; and the resulting estimates have coefficients of variation that generally exceed 20 %, which from a statistical agency point-of-view make them ‘not fit for purpose’ (ONS 2013). Another issue to be addressed is that all of the unit-level estimates displayed systematic bias. Occupation-based estimates over-estimated incomes in poorer areas; and under-estimated incomes in more affluent areas. Regression-based approaches under-estimated incomes across the board. These systematic biases are likely to be attributable in part to the omission of relevant within-household clustering effects (e.g., adults tend to partner or house-share with others of similar earning capacity); in part to poor and affluent areas attracting, respectively, the least- and best-paid with any particular occupational grouping; and in part to omitted spatial interactions, such as the way in which the income-elasticity of car ownership varies with degree of rurality. That said, the existence of a (near) linear

relationship between the observed and estimated mean incomes suggests that scope for suitable rescaling exists; and although estimate variance might exceed the stringent standards of a national statistical agency, it is generally low enough for the estimates to be of practical use for at least some purposes.

The most successful strategies considered, however, were those based on area-level synthetic estimation. Two strategies were evaluated. The first involved imputing an area's mean income on the basis of the mean income for areas of its type (geodemographic class). The second involved establishing a relationship between mean area income and other observed area-level aggregates, based on a sample of records, and using this to predict mean area-level incomes for all areas (ecological regression). In the evaluation conducted for this paper the most detailed geodemographic classification (96 categories) explained as much of the spatial variation in incomes as the best-performing unit-level synthetic estimates, but with much lower coefficients of variation (<20 %). However, the most consistently effective strategy of all was found to be ecological regression, which performed at least as well as the best unit-level estimation strategies, but without any systematic bias, and with markedly lower variance (CVs of c. 12 %). In terms of policy relevance, ecological regression was also found to display a lesser tendency to under-estimate the mean incomes of the most affluent areas relative to the other approaches considered. An apparent drawback of ecological regression is that it delivers an estimate of only one point in the full income distribution. However, this does not preclude specifying separate models to estimate the proportion of the population falling below given income thresholds, such as the 60 % of national median income used throughout Europe when assessing poverty rates. Ecological regression also has the benefit of relative simplicity, and does not require access to the underlying microdata for all of the areas being estimated.

From a UK perspective the 2011 Census appears to provide the Office for National Statistics with a once-in-a-decade opportunity to provide improved small-area estimates of income. In-house imputation at unit level (individual or household), using the Census database, would allow the creation of distributional as well as summary small-area estimates. However, the potential benefits of adopting this strategy should be tempered by acknowledging the likely systematic bias and possibly higher than desirable variance that would ensue. This, of course, is a message of wider relevance to other National Statistical Institutes holding large-scale spatially detailed population data, be it register, census or survey based. For UK users the results presented here also provide an indication of the strategies that can be most successfully pursued in attempting to understand the small-area geography of income when faced with a Census devoid of an income question. These results lend credence to ONS's inter-censal model-based income estimates. From a wider international perspective these results should provide pointers to the most effective solution to small-area income estimation bearing in mind local data contexts, in particular demonstrating that other strategies, including simple proxies, can compete surprisingly well with the model-based estimates currently most widely adopted. Finally, and perhaps almost as importantly, this paper also provides some indication of which strategies are likely to be most and least sensitive to changes in spatial scale and type of income unit (person or household).

Acknowledgments The research reported here arose from the ESRC-funded project 'Income imputation for small areas', award no. H507255166. Grateful thanks are due to the Census Custodians of England, Wales and Scotland for granting permission to access the Census Rehearsal dataset. A debt of gratitude is owed to a number of staff at the Office for National Statistics, in particular Keith Whitfield and Philip Clarke. Thanks are also due to David Voas for undertaking some of the preparatory work for this project. The Family Resources Survey, 1998, used to derive mean values for income bands, was sponsored and collected by the Department for Social Security, and supplied for research use by the UK Data Archive. Thanks also to two anonymous reviewers for their perceptive comments. All analyses and conclusions remain the sole responsibility of the author.

References

Anderson, B. (2013). Estimating small are income deprivation: an iterative proportional fitting approach, Chapter 4. In R. Tanton & K. L. Edwards (Eds.), *Spatial microsimulation: a reference guide for users* (pp. 49–68). London: Springer.

Berthoud. (2008). Area variations in household income across Great Britain. *Cambridge Journal of Regions, Economy and Society, 1*, 37–49.

Birkin, M., & Clarke, M. (1989). The generation of individual and household incomes at the small area level using SYNTHESIS. *Regional Studies, 23*(6), 535–548.

Birkin, M., & Clarke, G. (1995). Using microsimulation methods to synthesize census data. In S. Openshaw (Ed.), *Census users' handbook* (pp. 363–387). Cambridge: GeoInformation International.

Bond, S., & Campos, C. (2010). *Understanding income at small area level, Regional Trends 42* (pp. 80–94). London: Office for National Statistics.

Bramley, G., Lancaster, S., & Gordon, D. (2000). Benefit take-up and the geography of poverty in Scotland. *Regional Studies, 34*(6), 507–519.

CACI. (1999). *Paycheck: Targeting by income*. London: CACI.

Casarosa, M., Fierravanti, S., Colucci, A. et al. (2011). European policy brief: Small area methods for poverty and living condition estimates. SAMPLE Consortium. http://www.sample-project.eu/images/stories/policy_brief.pdf

Cloutier, N. R. (1997). Metropolitan income inequality during the 1980s: the impact of urban development, industrial mix, and family structure. *Journal of Regional Science, 37*(3), 459–478.

Collins et al. (undated) 2007 Census Test: the effects of including questions on income and implications for the 2011 Census, Office for National Statistics. http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/the-2011-census-project/2007-test/income-evaluation/index.html

Dale, A., Middleton, E., & Schofield, T. (1995). New earnings survey variables added to the SARs. SARs Newsletter 6, Census Microdata Unit, University of Manchester.

Daniel, E., & Bright, G. (2011). Exploring the geographical distribution of wealth using the output area classification. *Economic and Labour Market Review* 2011, 59–78.

Davies, H., Joshi, H., & Clarke, L. (1997). Is it cash the deprived are short of. *Journal of the Royal Statistical Society A, 160*(1), 107–126.

DCLG. (2011). *The english indices of deprivation 2010*. London: Dept. for Communities and Local Government.

Dorling. (1999). Who's afraid of income inequality? *Environment and Planning A, 31*(4), 571–574.

Dorling, D., Rigby, J., Wheeler, B., et al. (2007). *Poverty, wealth and place in Britain, 1968 to 2005*. Bristol: The Policy Press.

Experian (2005). Household income – UK 2004 data profile. Experian. http://cdu.mimas.ac.uk/experian/household%20income.pdf [accessed 13/12/2011]

Famhy, E., Gordon, D., Dorling, D., et al. (2011). Poverty and place in Britain, 1968–99. *Environment and Planning A, 43*, 594–617.

Fry, R. (2011). *Understanding household income poverty at small area level, Regional Trends 43* (pp. 1–16). London: Office for National Statistics.

Gibbons, S., Overman, H. G., & Pelkonen, P. (2010). *Wage disparities in Britain: People or place?* SERC Discussion Paper 60, Spatial Economics Research Centre, Dept. of Geography and Environment, London School of Economics.

Gonzalez, M. E., & Hoza, C. (1978). Small-area estimation with application to unemployment and housing estimates. *Journal of the American Statistical Associationion, 73*, 7–15.

Gordon, D., & Forrest, R. (1995). People and places 2: social and economic distinctions in England, Bristol: School for Advanced Urban Studies (SAUS).

Gordon, D., Adelman, L., Ashworth, K., et al. (2000). *Poverty and social exclusion in Britain*. York: Joseph Rowntree Foundation.

Green, A. (1994). The geography of poverty and wealth. Institute for Employment Research, University of Warwick.

Guangquan, L., Clarke, P., Taylor, A., Richardson, S., & Best, N. (2011). Improving small area estimates of income using Bayesian hierarchical models with heteroscedastic sampling error variances. Poster presented at the Small Area Estimation conference at Trier, Germany (11–13, 2011).

Heady, P., & Ralphs, M. (2004). Some findings of the EURAREA project – and their implications for statistical policy. *Statistics in Transition, 6*(5), 641–653.

Jargowsky. (1996). Take the money and run: economic segregation in U.S. Metropolitan Areas. *American Sociological Review, 61*(6), 984–998.

Lee, P., Murie, A., & Gordon, D. (1995). Area measures of deprivation: a study of current best practices in the identification of poor areas in Great Britain, Centre for Urban and Regional Studies, University of Birmingham.

Lehtonen, R., Veijanen, A., Myrskylä, M., & Valaste, M. (2011) Small area estimation of indicators on poverty and social exclusion, Deliverable 2.2, AMELI Advanced Methodology for European Laeken Indicators.

Longhurst, J., Cruddas, M., Goldring, S., & Mitchell, B. (2004a). *Model-based estimates of income for wards, 1998/99: Technical Report*. Titchfield: Office for National Statistics.

Longhurst, J., Cruddas, M., Goldring, S., & Mitchell, B. (2004b). *Model-based estimates of income for wards, 1998/99: Validation Report*. Titchfield: Office for National Statistics.

McCulloch, A. (2001). Ward level deprivation and individual social and economic outcomes in the British Household Panel Study. *Environment and Planning A, 33*(4), 667–684.

Micklewright, J., & Schnepf, S. V. (2010). How reliable are income data collected with a single question? *Journal of the Royal Statistical Society A, 173*(2), 409–429.

Mitchell, R. (2001). Multilevel modelling might not be the answer. *Environment and Planning A, 33*(8), 1357–1360.

Moss, C. (1999). Selection of topics and questions for the 2001 Census. *Population Trends, 97*, 28–36.

National Research Council. (2000). Small-area income and poverty estimates: priorities for 2000 and beyond, Panel on estimates of poverty for small geographic areas. In C. F. Citro & G. Kalton (Eds.), *Committee on national statistics*. Washington D.C: National Academy Press.

ONS (2006). The 2011 Census: Assessment of initial user requirements on content for England and Wales – Income, Information Paper. Titchfield: Office for National Statistics. http://www.ons.gov.uk/ons/about-ons/consultations/closed-consultations/2006/2011-census—responses/response-to-consultation—2011-census.pdf

ONS (2013). Beyond 2011: Producing socio-demographic statistics 2. M12 Methods & Policies. Titchfield: Office for National Statistics.

Openshaw, S., & Wymer, C. (1995). Classifying and regionalizing census data. In S. Openshaw (Ed.), *Census users' handbook* (pp. 239–270). Cambridge: GeoInformation International.

Rao, J. N. K. (2003). *Small area estimation*. Hoboken: Wiley.

Rees, P. H. (1998). What do you want from the 2001 Census? Results of an ESRC/JISC survey of user views. *Environment and Planning A, 30*, 1775–1796.

The EURAREA Consortium (2004). Project Reference Volume, Enhancing small area estimation techniques to meet European needs, Deliverable D7.1.4. http://www.ons.gov.uk/ons/guide-method/method-quality/general-methodology/spatial-analysis-and-modelling/eurarea/downloads/project-reference-volume–volume-one–main-report.pdf

Tunstall, R. (2005). Using the US and UK Censuses for comparative research. Discussion Paper prepared for the Brookings Institution Metropolitan Policy Program, STICERD, London School of Economics. http://sticerd.lse.ac.uk/dps/case/CBIR/Using_US_and_UK_Censuses.pdf

Tzavidis, N., Salvati, N., Pratesi, M., & Chambers, R. (2008). M-quantile models with applications to poverty mapping. *Statistical Methods and Applications, 17*, 393–411.

Tzavidis, N., Marchetti, S., & Chambers, R. (2010). Robust estimatin of small-area means and quantiles. *Australian and New Zealand Journal of Statistics, 2*, 167–186.

Vickers, D., & Rees, P. H. (2006). Introducing the area classification of output areas. *Population Trends, 125*(3), 15–24.

Voas, D., & Williamson, P. (2000). The scale of dissimilarity: concepts, measurement and an application to socio-economic variation across England and Wales. *Transactions of the Institute of British Geographers, 25*, 465–481.

Voas, D., & Williamson, P. (2001). The diversity of diversity: a critique of the geodemographic classification. *Area, 33*(1), 63–76.

White, N., Dent, A., Clarke, P., Silva, D., & Naylor, J. (2009). *Model-based estimates of income for middle super output areas, 2007/08: Technical Report*. London: Office for National Statistics.

Wilkinson, R., & Pickett, K. (2009). *The spirit level: why more equal societies almost always do better*. London: Allen Lane.