

Local indicators of child poverty – explanatory note

Summary of an improved method for estimating levels of child poverty in small areas of Great Britain

Juliet Stone, Francisco Azpitarte and Donald Hirsch

April 2019

Contents

1	Background.....	3
2	Explanation of improved method	4
3	Data.....	5
3.1	Survey data	5
3.2	Aggregate local data	6
4	Producing the final estimates.....	7
5	References.....	15
6	Technical appendix.....	16

1 Background

Between 2013 and 2018, CRSP produced an annual estimate of local child poverty on behalf of End Child Poverty. Estimates were based on an adjusted version of child poverty figures released by HMRC, which rely on numbers of individuals in receipt of out-of-work benefits and reported incomes for working and receiving tax credits. That method supplemented the HMRC data with information from the UK Labour Force Survey and the DWP (Households Below Average Income) to produce an updated estimate of child poverty at a local level. However, going forward this approach became less satisfactory for several reasons:

1. Known issues with the HMRC statistics (underestimation of in-work poverty and over-estimation of out-of-work poverty).
2. The total child population in each local area used to estimate percentages in the HMRC statistics is based on receipt of child benefit. Means-testing of child benefit provision from January 2013 means that higher earning families may opt out. Therefore, the denominator has become less representative of the total population of children and this effect will likely vary by region and local area.
3. Introduction of Universal Credit means that making assumptions about poverty levels about families on 'out of work' benefits is less straightforward, since those in and out of work get the same category of benefit.
4. Different employment trends by region have in practice created the scope for serious inaccuracies when comparing these trends across areas, given that the relationship between in and out of work poverty is not being accurately measured at the local level by the HMRC data themselves,

There is therefore a need for a more robust method for estimating child poverty at a local level. To produce more reliable estimates of local child poverty, CRSP has developed an approach based on small area estimation, taking advantage of a range of administrative data combined with survey data. Administrative data is data reported from administrative systems such as benefit claimant counts which gives detailed counts down to the very local level about certain specific variables; survey data is based on national samples, usually not valid below regional level, but allows detailed information about individual households not available in administrative data, including income.

In simple terms, the new method considers the relationship between the risk of poverty for households in the survey and the characteristics of the area that they live in, in terms of a range of area-level demographic and socioeconomic indicators (usually from administrative data sources). These relationships can be used to estimate child poverty rates from the administrative data that is available locally.

2 Explanation of improved method

To estimate child poverty rates at the local level we use a form of small area estimation. This term refers broadly to statistical methods used to produce estimates of the value of a measure at a smaller level of geography than available from direct estimates from a survey or other data source. Small area estimation techniques are now commonly used by national governments and international organizations such as the World Bank to produce local estimates of poverty in developing and developed countries (e.g., Lange et al. 2018; Chandra et al. 2018).

The method applied here is based on an approach developed by the Office for National Statistics (ONS) and used to produce estimates of average household income and overall poverty levels at middle layer super output area (MSOA) level (Henretty, 2017). MSOAs are statistical (not administrative) geographies designed by the ONS to improve the reporting of area-level statistics and are comprised of 2,000-6,000 households (the majority with a population of 7-9000 individuals).

There are a total of 7,201 MSOAs in England and Wales. In Scotland, there are an additional 1,279 'Intermediate Zones', which are of roughly equivalent size to MSOAs, producing a total of 8,480 small areas in Great Britain.¹ Hereafter, the term MSOA is used to refer to this combined set of MSOAs plus Intermediate Zones.

We adapt the methods proposed by the ONS to produce local child poverty rates at the MSOA level. The method requires combining survey data on characteristics of individuals in terms of household poverty with data on the social, demographic, and economic characteristics of the MSOAs in which they live.

The process of estimating child poverty for small areas proceeds in two steps. First, it estimates a relationship between a child's probability of living in poverty, defined using survey data on household income, and the characteristics of the MSOA. The relationship between child poverty and MSOA characteristics is then used to estimate the prevalence of child poverty in all MSOAs, including those not represented in the survey data. This method thus allows us to produce a complete map of local poverty rates. The precise method for doing this is set out in the technical appendix to this paper.

¹ We do not include Northern Ireland in this analysis, due to issues around the lack of equivalent small area geography and difficulties with accessing aggregate administrative data that are comparable with the rest of the UK.

3 Data

3.1 Survey data

To estimate levels of child poverty at MSOA level using the method described in the appendix, survey data that meet the following criteria are required:

1. Includes a reliable measure of household income before/after housing costs
2. Allows identification of children within households
3. Includes an MSOA identifier to allow linkage to area-level information
4. Includes respondents from a reasonable proportion of MSOAs within the sample
5. Is large enough to provide enough statistical power to produce sufficiently precise estimates
6. Can be updated annually

The survey chosen for the present work is Understanding Society, the UK Household Longitudinal Survey. Understanding Society is a panel study of 40,000 households in the UK, started in 2009/10 and repeating annually. Eight waves of the survey are currently available.

Understanding Society meets all six criteria specified above. In addition, the survey has the advantage of providing longitudinal data, with the potential for future analysis of the dynamics of child poverty. Understanding Society is already used by the Department for Work and Pensions (DWP) in their experimental statistics on income dynamics, which look at changes in household income including a measure of persistent low income (DWP, 2017).

For our initial analyses, we use wave 8 of the survey, which includes data collected over a two-year period from early 2016 – May 2018. The survey includes about 10,000 children with valid data on household income, distributed across 3,659 MSOAs.

3.2 Aggregate local data

Data reported not for individual households but totals for small local areas (aggregate data) are available at MSOA level from a number of different sources, mainly 'administrative'. Table 1 below summarizes the variables considered for inclusion in our analyses.

Table 1: Aggregate data for potential inclusion in analysis

Variable	Data source	Description	Dates
Household composition	2011 census	Households are classified as: <ul style="list-style-type: none"> • One person household (under/over 65 yrs) • All aged 65 and over • Married couple: No children • Married couple: Dependent children • Married: All children non-dependent • Cohabiting couple: No children • Cohabiting couple: Dependent children • Cohabiting couple: All children non-dependent • Lone parent: Dependent children • Lone parent: All children non-dependent 	2011
National Statistics Socioeconomic Classification (NSSEC) of household reference person (HRP)	2011 census	1.1 Large employers and higher managerial and administrative occupations 1.2 Higher professional occupations 2. Lower managerial, administrative and professional occupations 3. Intermediate occupations 4. Small employers and own account workers 5. Lower supervisory and technical occupations 6. Semi-routine occupations 7. Routine occupations 8. Never worked and long-term unemployed	2011
Tenure mix	2011 census	% of households in owner-occupied housing % of households in private rented housing % of households in social rented housing	2011
Claimant count	DWP	Proportion of individuals of working age who are claiming unemployment-related benefits: <ul style="list-style-type: none"> • Men • Women • Total 	2018
Children in Low-Income Families Local Measure	HMRC	Proportion of children defined as being in a low income family.	2015

Variable	Data source	Description	Dates
Child tax credit receipt	HMRC	Proportion of children in families in receipt of child tax credit.	2016/17
House price statistics for small areas	HM Land Registry, Land and Property Services Northern Ireland, Office for National Statistics and Registers of Scotland.	Mean house prices	2017
Indices of multiple deprivation		Income and employment domains. <ul style="list-style-type: none"> The Income Deprivation Domain measures the proportion of the population experiencing deprivation relating to low income. The definition of low income used includes both those people that are out-of-work, and those that are in work but who have low earnings (and who satisfy the respective means tests). The Employment Deprivation Domain measures the proportion of the working-age population in an area involuntarily excluded from the labour market. This includes people who would like to work but are unable to do so due to unemployment, sickness or disability, or caring responsibilities. 	2015

Table 2 (cont.)

We use automated, stepwise selection procedures to identify the sub-set of these potential predictors that provides the best estimates of child poverty.

4 Producing the final estimates

Table 2 below shows the estimation results of the model of child poverty defined in terms of household income after housing costs. The set of variables that survived the selection process include a range of demographic and socioeconomic characteristics of the MSOAs including the prevalence of benefit claimants, unemployment, and occupational structure. We can evaluate the relative importance of each of the selected variables based on the estimated percentage change in odds of child poverty per unit change in the variable. We have standardized the variables so that all are measured on the same scale, allowing a more informative comparison. Two estimates are included: first the ‘unadjusted’ relationship between each covariate and child poverty, which

show the association without taking any of the other variables into account. The ‘final model’ estimates show the relationship when all the covariates are accounted for at the same time – the relationship “all other things being equal”. Figures 1 and 2 provide an additional illustration of the relationships between the predictor variables and child poverty. Figure 1 shows the unadjusted predicted probability of child poverty (after housing costs) when the value of each covariate is low (10th percentile), average (median, 50th percentile) or high (90th percentile). The stronger the relationship between a covariate and child poverty, the bigger the difference between the three bars.

Table 2. Modelling child poverty (after housing costs)

Area-level characteristic	Estimated % change in odds of child poverty per unit (standard deviation) change in covariate			
	Unadjusted	Final model	Interpretation of adjusted relationship	Child poverty estimate given one unit increase in variable, relative to mean AHC child poverty rate (30%)
Total claimant count rate, age 16-64: one unit = 1.5%	52.2%	58.1%	For each 1.5% increase in the total claimant count rate within an MSA, the odds of a child being in poverty there increases by 58%.	40.4%
Percentage of children in child tax credit recipient families: one unit = 17.9%	80.4%	41.5%	For each 17.9% increase in the percentage of families receiving CTC within an MSA, the risk of a child being in poverty there increases by 41.5%	37.8%
Male claimant count rate age 16-64: one unit = 2.0%	47.7%	-35.7%	For each 2% increase in the male claimant count, the risk of a child being in poverty falls by 35.7%. (NB, this assumes all other variables remain equal, including overall claimant count, so a lower male claimant count implies a higher female claimant count)	21.6%
Percentage of married couples with dependent children: one unit = 4.8%	-9.5%	35.1%	For each 4.8% more households who are married couples with dependent children, the risk of a child being in poverty rises by 35.1%	36.7%

b) Variables with weaker effects

Area-level characteristic	Estimated % change in odds of child poverty per unit (standard deviation) change in covariate		
	<i>1 Unit</i>	<i>Unadjusted</i>	<i>Final model</i>
Percentage of single households aged 65+	3.5%	-29.5%	-9.2%
Percentage of lone parent households with non-dependent children only	9.2%	28.4%	-11.7%
Percentage of household respondents in lower managerial and professional occupations	6.4%	39.3%	12.6%
Percentage of children in low income families on HMRC measure	9.5%	73.3%	-0.3%
Percentage of never worked/long-term unemployed	3.9%	58.4%	-20.4%
Percentage of people aged 65+	3.9%	-43.4%	19.4%
Percentage of household respondents in higher managerial and similar occupations	1.9%	-43.4%	13.8%

Figure 1: Predicted probability of child poverty (after housing costs) when values of area-level covariates are low, average, or high. Unadjusted relationships.

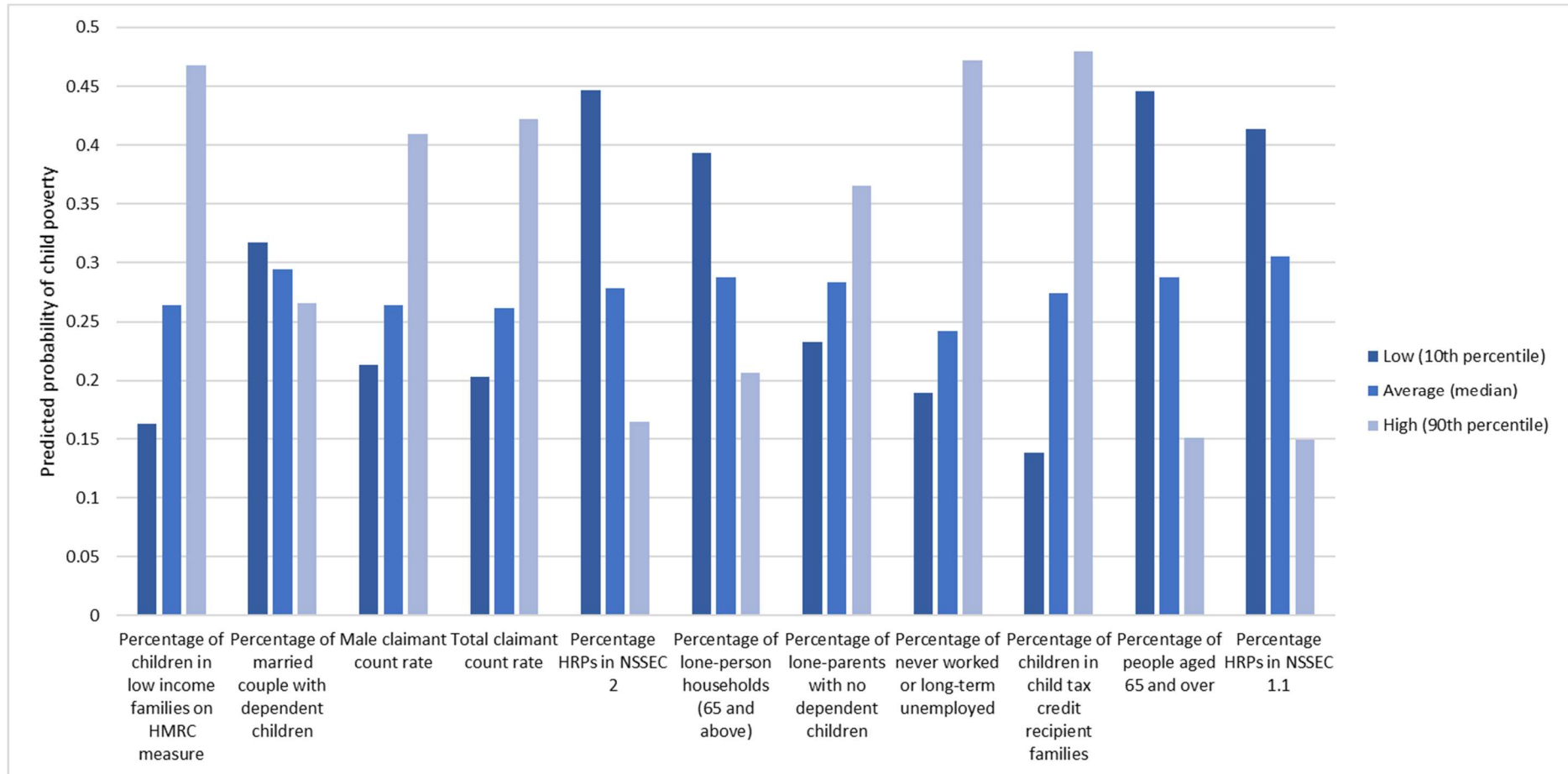
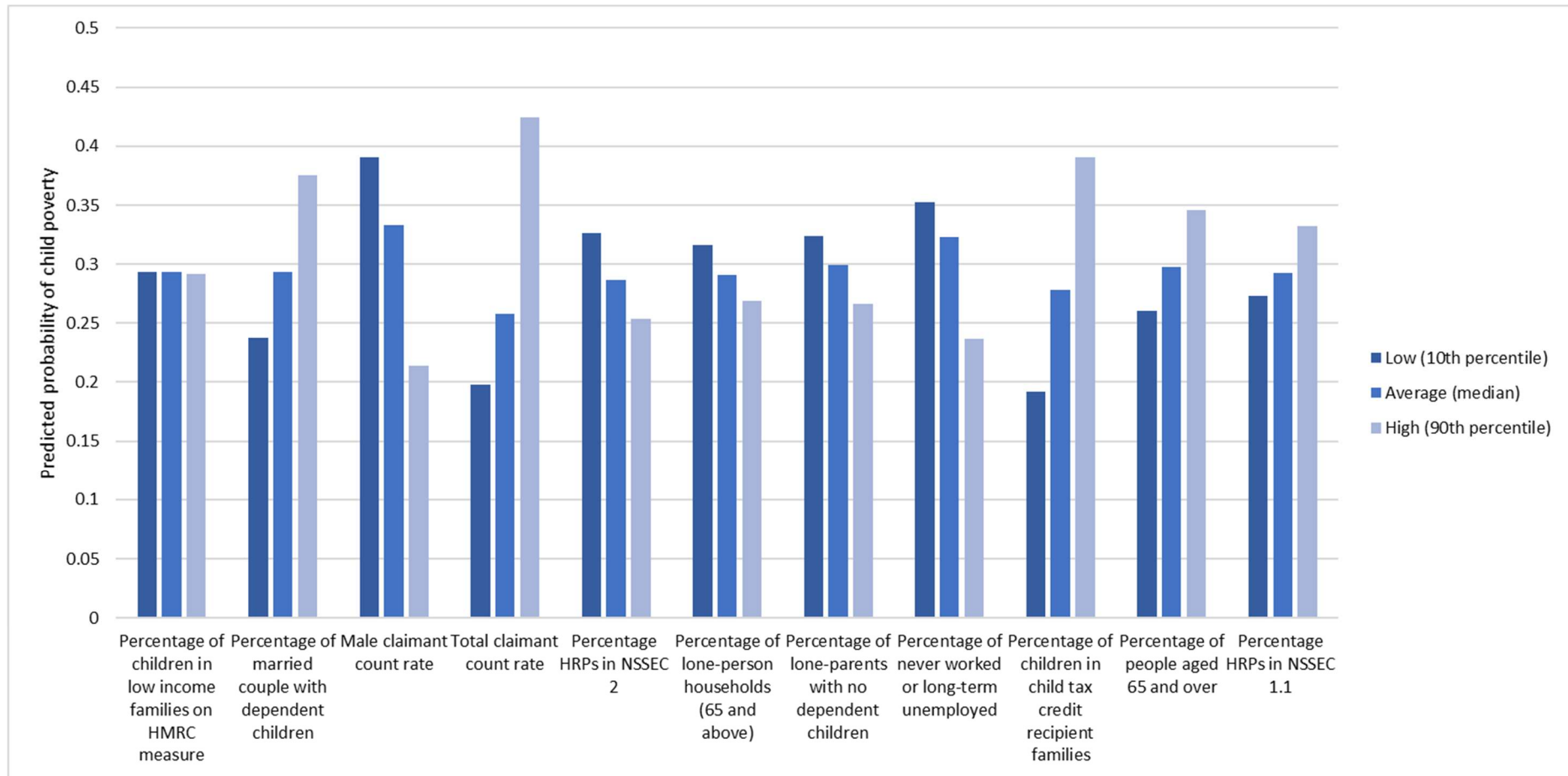


Figure 2: Predicted probability of child poverty (after housing costs) when values of area-level covariates are low, average, or high. Adjusted (final model, “all things equal”) relationships.



Our results show that unemployment-related benefit claimant rates are the strongest predictors of child poverty at the MSOA level, followed by the percentage of children in child tax credit recipient families. Some other factors have strong unadjusted effects but weak effects in the final model. In particular, the existing HMRC measure of child poverty (percentage of children in low income families) shows a strong relationship with child poverty as measured in Understanding Society, as would be expected. However, as illustrated in Figure 2, where we take all the other area-based measures into account, our measure of child poverty changes very little according to variation in the percentage of children in low income families measure – the three bars for this variable are almost equal. This indicates that the other area-level data we are using is providing enough information to estimate the variation in child poverty between MSOAs, and the HRMC measure is not in fact contributing to the explanation of this relationship in any meaningful way. This supports the use of the more nuanced approach that we are applying, rather than just relying on the HMRC measure as the only “direct” indicator of local child poverty.

Note that in Figure 2, in some cases the relationship between the area-level characteristic and child poverty has reversed, in comparison with Figure 1 where the other variables are not taken into account. In simple terms, this is because there are varying degrees of correlation between the different variables. For example, in Figure 1, the probability of child poverty is higher when the area-level, male claimant count rate is higher. This is what we would expect given that families where parents are unemployed will tend to have low incomes. However, after adjusting for all the other variables in Figure 2, the male claimant count rate shows a *negative* association with child poverty – in other words, the probability of child poverty is higher when the male claimant count rate is lower. This seems counterintuitive, but can be explained by the fact that the overall claimant count rate is also included in the model. So, we are now considering the male claimant count rate in the context of the overall rate. It helps if we recognize that female unemployment is more likely to result in child poverty than male unemployment (in large part due to the effect of lone motherhood). Therefore, if the overall claimant count rate is high, and the male claimant count rate is low, this means that the female claimant count rate (and hence child poverty rate) must be relatively high. Conversely, with a given overall claimant count, a high male claimant count signifies fewer female claimants and thus lower child poverty. Another variable where the direction of effect reverses is percentage of married couples with children in the local population. Since families with children within married couples are on average somewhat more affluent than average, having more such families is associated with a lower child poverty rate; but after taking account of other factors associated with affluence, the poverty risk of such families is higher than would have been expected.

After selecting the model that produces the most satisfactory and robust estimates of MSOA-level child poverty, we apply an additional adjustment to calibrate the results to the most recent estimates of child poverty from the Households Below Average Income series at regional level. We then use look-up tables provided by the Office for National Statistics to produce estimates at additional levels of geography:

1. Ward
2. Parliamentary constituency
3. Local authority district

All these results are presented in tabular form each year, for each of these three geographies, including estimates for the predicted percentage of children in households with income below 60% of median income before and after housing costs. For wards where there is more than one MSOA, we take the average of the estimates for these MSOAs, weighted to account for differences in population size. We also aim provide estimates of the number children in poverty in each local area. The data were first produced in 2019 calibrated to the HBAI income data for 2017/18. An estimate of change over the past year was also included, at parliamentary constituency and local authority district level, but we do not consider the data at ward level precise enough to make it useful to report change at that geography.

Below we show an example of types of map-based output we can potentially produce, based on the first year's results. Figures 3 and 4 below map child poverty rates by MSOA for England and Wales and for London, respectively.

Figure 3: Map of child poverty: England and Wales, by MSOA (2017/18)

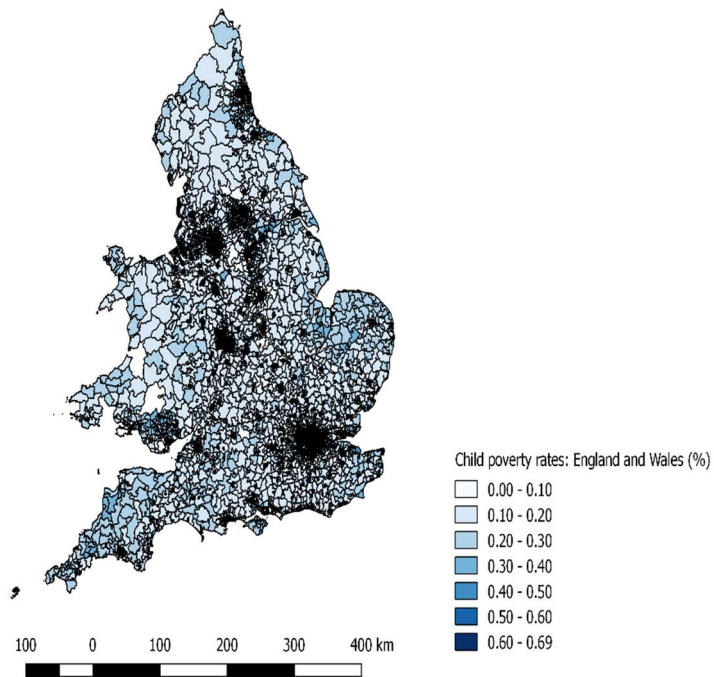
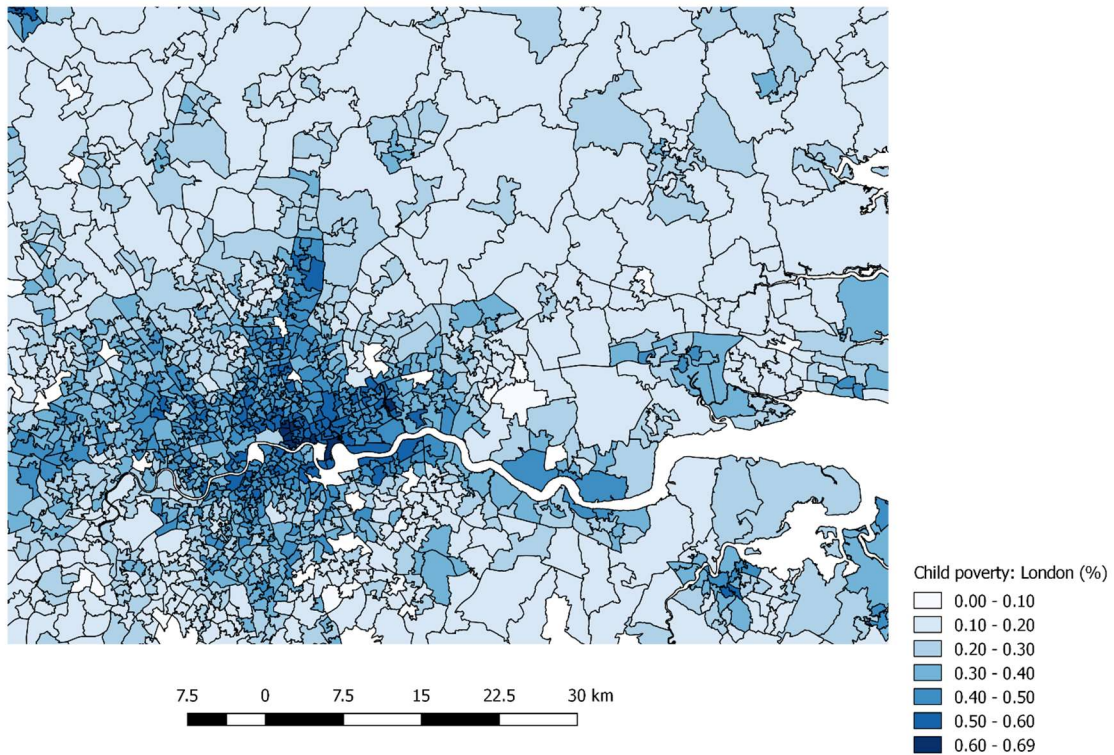


Figure 4. Child poverty rates, London area, by MSOA (2017/18)



5 References

Chandra H, Aditya K, Sud UC (2018) Localised estimates and spatial mapping of poverty incidence in the state of Bihar in India—An application of small area estimation techniques. PLoS ONE 13(6): e0198502.

<https://doi.org/10.1371/journal.pone.0198502>

DWP (2017) *Income Dynamics Background information and methodology*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/600299/income-dynamics-background-information-and-methodology.pdf (Accessed: 19 March 2019).

Henretty, N. (2017) *Small area model-based households in poverty estimates, England and Wales - Office for National Statistics*. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/smallareamodelbasedhouseholdsinpovertyestimatesenglandandwales/financialyearending2014> (Accessed: 19 March 2019).

Lange, S., Pape, U. J., & Pütz, P. (2018). *Small area estimation of poverty under structural change*. Policy Research Working Paper N. 8472, the World Bank.

6 Technical appendix

We created a dataset combining individual level data from the Understanding Society longitudinal survey with census and administrative data at the MSOA level from different sources listed above. Using this new dataset, we construct a predictor of local child poverty that takes as input the social, economic, and demographic characteristics of MSOAs. More specifically, we propose and estimate the following predictor of child poverty linking children's poverty status and MSOA characteristics:

$$\text{Ln}(y_{i,r}) = X\beta + u_r + v_{i,r},$$

where Ln is the logarithmic function and:

$y_{i,r}$ is an income poverty indicator of child i living in the MSOA r that takes values 1 if the child in the sample is living in a household with an equivalent income below the poverty line defined as 60 per cent of the median equivalent income in the population; and takes value 0 if the child is living in a household with an income greater or equal than the poverty line;

X is a matrix with data on the characteristics of the MSOA children live in and β is a vector of parameters summarising the statistical association between local characteristics and poverty;

and u_r and $v_{i,r}$ are, respectively, the area level and individual error terms capturing other factors that influence the risk of child poverty but which cannot be observed by the analyst and therefore cannot be include in the estimation.

The model for $\text{Ln}(y_{i,r})$ laid out above belongs to the family of logistic models with local area random effects. This type of models are commonly applied in social sciences for investigating categorical outcomes such as poverty.

In the first stage of the estimation the vector of parameters β is estimated using the new linked dataset containing data on the child poverty indicator ($y_{i,r}$) and the aggregate information at the MSOA level. The vector β is key to the prediction as it summarises the relationship between socioeconomic and demographic characteristics and the prevalence of child poverty. Thus, for example, areas with high local unemployment rates might have a low average household income. We can therefore estimate the extent to which, for example, area-level unemployment rates change the estimated probability that a child will be living in a household with income below 60% of median. For instance, the model might show that for every 1% increase in the MSOA-level unemployment rate, the probability that a child is living in a low-income household increases by 0.5%.

Given the large number of aggregate characteristics available for MSOAs we selected the set of variables in X finally included in the analysis using a bidirectional stepwise regression technique where potential covariates are added or deleted from the model depending on their contribution to the model fit. The set of variables that survived this process and where finally used in the regression analysis is described in Section 4 below.

Because the Understanding Society survey from which individual level data on poverty is taken is based on a sample of the total population, many MSOAs may only include a handful of cases, while other may not be represented in the survey at all. The idea of small area estimation is to use the information from those MSOAs represented in the survey to additionally predict the child poverty rate of those areas not represented in the survey data. This is precisely the what we do in the second stage of the estimation, where the estimates of the β s are combined with aggregate data from MSOAs to produce local estimates of child poverty. This is possible because aggregate data on the characteristics of MSOA are available for all MSOAs including those not represented in the survey.