Journal of Official Statistics, Vol. 38, No. 1, 2022, pp. 213–237, http://dx.doi.org/10.2478/JOS-2022-0010

Experimental UK Regional Consumer Price Inflation with Model-Based Expenditure Weights

James Dawber[1], *Nora Würz*[2], *Paul A. Smith*[1], *Tanya Flower*[3], *Heledd Thomas*[3], *Timo Schmid*[4], *and Nikos Tzavidis*[1]

Like many other countries, the United Kingdom (UK) produces a national consumer price index (CPI) to measure inflation. Presently, CPI measures are not produced for regions within the UK. It is believed that, using only available data sources, a regional CPI would not be precise or reliable enough as an official statistic, primarily because the regional partitioning of the data makes sample sizes too small. We investigate this claim by producing experimental regional CPIs using publicly available price data, and deriving expenditure weights from the Living Costs and Food survey. We detail the methods and challenges of developing a regional CPI and evaluate its reliability. We then assess whether model-based methods such as smoothing and small area estimation significantly improve the measures. We find that a regional CPI can be produced with available data sources, however it appears to be excessively volatile over time, mainly due to the weights. Smoothing and small area estimation improve the reliability of the regional CPI series to some extent but they remain too volatile for regional policy use. This research provides a valuable framework for the development of a more viable regional CPI measure for the UK in the future.

Key words: CPI conceptual framework; basket of goods and services; small area estimation; Fay-Herriot models.

1. Introduction

For a long time, users of price statistics in the UK have suggested that regional indices of consumer prices would be valuable in understanding how inflation varies across the country, and whether there are important differences in regional inflation (RPI Advisory Committee 1971; Fenwick and O'Donoghue 2003; UK Statistics Authority 2013, paragraph 3.13). The official position has been that the regional numbers of price quotes are too small to support the calculation of indices, and it has not been a sufficiently high

[1] University of Southampton, Highfield, Southampton, SO17 1BJ, UK. Emails: J.P.Dawber@soton.ac.uk, P.A.Smith@soton.ac.uk and N.TZAVIDIS@soton.ac.uk
[2] Freie Universität Berlin, Department of Business and Economics, Garystraße 21, Berlin, 14195, Germany. Email: NoraUlrike.Wuerz@fu-berlin.de
[3] Office for National Statistics, Government Buildings, Cardiff Road, Newport, NP10 8XG, UK. Emails: tanya.flower@ons.gov.uk, Heledd.Thomas@ons.gov.uk
[4] Otto-Friedrich-Universität Bamberg, Feldkirchenstraße. 21, Bamberg, 96052, Germany. Email: timo.schmid@uni-bamberg.de
Acknowledgments: James Dawber and Paul Smith were funded by the ONS under contract PU-16-0031-6.009. Nikos Tzavidis and Paul Smith were supported by EU Horizon 2020 funding as part of the MAKSWELL project, grant agreement No. 770643. Nora Würz received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 730998, InGRID-2 – Integrating Research Infrastructure for European expertise on Inclusive Growth from data to policy. The opinions expressed are those of the authors, and do not necessarily reflect those of their organisations or the ONS.

priority to invest in additional price collection for this purpose. Some limited information from the Office for National Statistics (ONS) on variation in regional prices has been made available through publications on Relative Regional Consumer Price Levels (RRCPLs) (Baran and O'Donoghue 2002; Wingfield et al. 2005; ONS 2011), which have used information from additional price collections made every six years to adjust Purchasing Power Parity (PPP) statistics. PPP prices are collected in the capital city, and a periodic exercise is undertaken to adjust indices to represent the whole country. RRCPLs are *spatial* price indices, which show the differences in price levels between regions (relative to a reference region = 100) with a fixed basket of goods, but are not *temporal* indices designed to show price change (inflation, relative to a reference time = 100), because the basket is different each time they are produced. Because of the methodology and differences in the weights on each occasion, RRCPLs are not satisfactory even for a once every six years approximation of regional inflation. The focus of this article is on a regional index that can measure the temporal differences (inflation), rather than the spatial differences. So we allow the baskets and weights to vary by region, to best reflect expenditure patterns within regions. The resulting indices are suitable for comparing inflation rates between regions, but not for comparing price levels.

The Consumer Prices Index (CPI) is used as a national measure of inflation in many countries. Some countries have also developed inflation measures at a sub-national level. In the United States (US) the construction of the CPI includes a regional aggregation phase, and therefore component indices at the regional level are part of the standard outputs (BLS 2018); Japan also produces regional indices (e.g., Statistics Bureau of Japan 2020; Nagayasu 2011). Weber and Beck (2005) compiled a database of regional prices for Europe where they found regional or city indices in Austria, Finland, Germany, Spain and Portugal. In Germany, the Federal Statistical Office publishes regional CPIs for the 16 federal states (Statistisches Bundesamt 2020) and there are also publications with regional indices for 401 German districts relying on an econometric model (Kosfeld et al. 2009). The German national weighting pattern is updated every five years using the "Einkommens und Verbraucherstichprobe" (Statistisches Bundesamt 2018), but the sample size is insufficient for the estimation of regional weights. Other countries with regional inflation measures include Poland (Gajewski 2017), Russia (Brown et al. 2018), Indonesia (Purwono et al. 2020), South Korea (Tillman 2013) and Turkey (Yesilyurt and Elhorst 2014; Duran 2016). We did not find detailed methodological descriptions in all these cases, but of those countries where information is available, almost all have a regional aggregation stage in constructing the national price index, so that regional indices are automatically available. Other methods are the econometric model for district inflation in Germany, and an experimental investigation of small area estimation in Indonesia (Fengki et al. 2020; sec. 5).

Other countries, including the UK, do not have a regional aggregation phase, so a special exercise is needed to produce regional CPIs. There has been interest in regional indices over many years in the UK. The Chancellor of the Exchequer announced work on regional prices in 2003 (Fenwick and O'Donoghue 2003), which was translated into the development of RRCPLs. Although these have been published each time PPP data collections have taken place, there has not been any substantial development of these statistics. Fenwick and O'Donoghue (2003) also discuss the potential for regional temporal indices, but conclude that they need further development; the annex to their paper lists the issues which make such a development challenging.

Economists have an interest in regional variations in price inflation (and more widely in regional differences in the cost of living, which is not so easily defined or calculated). Borooah et al. (1996), Hayes (2005) and Rienzo (2017) have all attempted to calculate regional versions of a consumer price measure for the UK with simplified methodology and based on available data sources. Regional variations in price levels measured by PPPs are also important inputs in local economic analysis. For example, Marchetti and Secondi (2017) estimate regional household consumption expenditure adjusted for differences in regional PPP in Italy, and Marchetti et al. (2019) use regional (province)-specific prices to adjust the national poverty threshold in Italy. Both of these applications use small area modelling approaches to make predictions of the variables of interest at local levels.

We assess the feasibility of producing a regional CPI measure for comparison of inflation rates between regions of the UK, that is, regional temporal indices, based on the official data collections. The UK has twelve statistical regions, including Wales, Scotland, Northern Ireland and nine regions of England; these are the Nomenclature of Territorial Units for Statistics (NUTS) 1 statistical regions. We develop a regional CPI rather than a CPIH (CPI including owner occupiers' housing costs) despite the importance of regional variation in housing and rental costs, because of the additional complications caused by the lack of regional microdata.

The derivation of the CPI requires data on price changes of certain goods, and also on consumer expenditure, used to determine both the basket composition and the weights (the proportion of spending on the goods). Sources are needed for all these components to ensure that the national CPI can be used to calculate reliable inflation rates, with suitable sampling designs and sample sizes to ensure they are nationally representative. For the development of a regional CPI, these data for prices and expenditure are partitioned into the separate regions. The reliability of a CPI at the regional level will be reduced because the smaller sample sizes lead to lower precision of the estimates. We consider this the primary limitation to the development of reliable and temporally stable regional CPIs. Although there may be bias in the regional CPI estimates as well as high variance, we believe the latter to be the larger problem and assume the bias to be ignorable.

A second limitation is the availability of regional-level data sources. At the national level, expenditure weights are calculated using many data sources. However, not all these sources are readily available for each region. For example, at the time of researching there was no regional equivalent for the National Accounts, though since then ONS (2018) has published experimental household final consumption expenditure (HFCE) at the regional level. For future research this regional HFCE data could be investigated to give balanced regional expenditure weights. Other data sources, such as administrative data are not very accessible, but the Living Costs and Food (LCF) survey data are accessible and also have region identifiers, hence can be used to estimate regional expenditure. For this reason, we use only LCF survey data to estimate the expenditure weights. Price data do not have the same issues since the data are readily available with region identifiers, and are used directly.

The aims of this research are first to assess the feasibility of calculating an experimental regional CPI series using available data sources, and second, to investigate model-based methods to overcome the primary limitation of the reduced sample sizes. We look at smoothing and small area estimation (SAE) methods that may improve the reliability of regional CPIs without having to collect additional data. SAE methods based on composite estimation have been examined in the US in a similar context (Swanson et al. 1999), however we focus on the use of Fay-Herriot models.

The structure of this article is as follows. In Section 2, we provide more background on the conceptual framework of a regional CPI as well as background on available data sources and the Classification of Individual Consumption by Purpose (COICOP). In Section 3, we present methods for constructing an experimental regional CPI with just LCF survey data and publicly available price data. We also assess the experimental regional CPI series for 2010 – 2016. In Section 4 we investigate the use of smoothing and small area estimation approaches to estimate the regional weights, intending to improve the regional CPI series. Finally, in Section 5 we discuss the results and suggest further research.

2. Structure, Data Sources and COICOP Classification

2.1. A Conceptual Framework for Regional CPIs

To have a sound basis for the methodology of regional price indices, it is important to set out the variations in the conceptual framework from the calculation of a national index. Here we set out the target concept; then we can be clear where the available data requires us to deviate from it.

The starting point for a (temporal) regional price index should be the regional basket of goods and services, to account for differences in spending patterns between regions. A threshold is used to exclude products with only minor expenditures. The maximum threshold is set at one part per thousand (ppt) according to the EU regulation 1687/98 on HICP (EU regulation 1998), but the UK implementation uses judgement for products with between 0.14 and 0.56 ppt of expenditure. For the regional estimates, we therefore take a fixed threshold of 0.5 ppt. Products below this threshold are not considered in the derivation of the index.

For items in the regional baskets, we require regional expenditure weights. For consistency, these should come from HFCE after balancing within the national accounts. But the information is not available at this level, and therefore we approximate it using the LCF survey source. LCF data points are fuzzy, especially for households residing near regional boundaries, since the expenditure is not certainly in the region of residence; and this applies more generally for non-geographic expenditure. An alternative approach would be to obtain expenditure weights from businesses, which could be better localised, but this information does not have sufficient commodity detail for use in constructing a CPI.

The final component is the prices, and these should be operative in the region of interest. Price quotes are labelled with region identifiers, but there is a need to deal appropriately with central and nongeographic prices.

2.2. Data Sources

To develop an experimental regional CPI we use price quote data for the prices and LCF survey data for the expenditure. Monthly price quote and item index data for the UK are available from the ONS website from January 2010 (ONS 2020). The price quote data provide prices for items (goods or services) with corresponding information about the shop type, region, validity and stratum weight. Not all prices are represented in the price quote data, because many have nationally defined pricing and are collected centrally – approximately 45% of the weight of the basket is comprised of these centrally collected items. These central items are reported in the item index data sets, which report the indices and weights for the national CPI, but do not include a regional breakdown. For the regional CPI, the price quote data will need to be partitioned by region to calculate the item indices for each region, and then national level indices used for those collected centrally.

LCF survey data were obtained for the years 2008–2014 (Department for Environment, Food and Rural Affairs and Office for National Statistics 2019) for estimating expenditure, which will contribute to regional CPIs for 2010–2016 (because the CPI is a Lowe index with expenditure data from an earlier period than the reference period). The household sample sizes are shown in Table 1, and vary between regions and across years. Increases to the sample size in Northern Ireland (from 2016/7) and Scotland (from 2018) have more recently been implemented. The LCF survey data provide expenditure on products purchased by each sampled household, classified according to the COICOP classification at the COICOP-plus level. An example of the COICOP classification is shown in Table 2, which also shows the labels for the different levels, including item level, which sits below the formal COICOP hierarchy. Items are chosen by the ONS to be representative within a COICOP5 category and it is item prices that are reported in the price quote data.

We use the LCF survey data and accompanying household weights to estimate the mean (or equivalently the total) household expenditure by COICOP4 level in each of the twelve regions of the UK. For the national weights, the LCF survey data is one of multiple sources used to estimate expenditure. However, for the regional expenditure weights, we calculate directly from the LCF survey data without other sources, which are only accessible within the ONS. Relying on just one source for the expenditure weights is expected to adversely affect the reliability of the regional CPI, but it remains to be seen to what extent.

Table 1. Household sample size for LCF data, 2008 to 2014.

	2008	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**Mean**
North East	235	236	258	283	262	251	255	**254.3**
North West	592	582	596	647	623	585	588	**601.9**
Yorkshire and the Humber	491	484	485	521	521	462	459	**489.0**
East Midlands	405	393	413	455	425	424	440	**422.1**
West Midlands	469	527	470	526	513	526	470	**500.1**
East	532	499	515	543	563	497	498	**521.0**
London	472	464	476	536	490	480	407	**475.0**
South East	806	701	679	761	783	681	740	**735.9**
South West	502	518	495	507	493	429	468	**487.4**
Wales	265	272	261	251	266	246	222	**254.7**
Scotland	500	544	468	500	483	412	434	**477.3**
Northern Ireland	574	602	147	161	171	151	152	**279.7**

Table 2. Example COICOP classification for UK national CPI framework

Level	COICOP2	COICOP3	COICOP4	COICOP5	COICOP-plus	Item
Label	Division	Group	Class	Expenditure code	Category	Item
Example	Food and non-alcoholic beverages	Food	Bread and cereals	Bread	Buns, crispbread and biscuits	White sliced loaf branded 750g
Code	01	1.1	1.1.1	1.1.1.2	1.1.1.2.2	–
Number*	12	47	85	303	367	731

*Numbers tend to change over time, and the UK aggregates some groups as well, so these should be taken as approximate.

3. Constructing the Experimental Regional CPI

3.1. Regional CPI Price Aggregation

Ideally, the methods used to derive a regional CPI should be kept as close as possible to the national CPI. For full details of how the national CPI is calculated, see ONS (2019). Figure 1a outlines the process – elementary aggregates are calculated as geometric means of the relative prices weighted at the shop level to reflect market share of chain shops. The elementary aggregates are then aggregated into item indices using arithmetic means weighted at the stratum level, which capture variations by region and shop type (independent vs multiple). The item indices are then weighted together in proportion to the national consumption of the item, derived from the National Accounts expenditure data, LCF survey data, market research data and other sources including administrative data. Weighted arithmetic means are then calculated at the class level and then finally, these class indices are aggregated using class expenditure weights to produce the national CPI.

In adapting the methods of the national CPI to the regional level, we use the conceptual framework for regional price indices (Subsection 2.1) and follow these concepts and the national methods wherever possible.

The first step in constructing an experimental regional CPI is to collate the publicly available price quote and item indices data. Of the 713 items with non-zero weights listed in the 2016 item indices dataset, 548 items were available in the price quote dataset, so regional versions could be calculated. This leaves 165 items (23.1%) which are collected nationally, and will therefore not contribute to differences in prices between regions (though they may have different effects in different regions if they are weighted differently). For 2016, the 713 items account for 98.3% of the total weight, and the remaining 1.7% are not represented to prevent individual retailers from being identified. We ignore the undisclosed weights, and rescale the weights accordingly. The 548 items in the price quote dataset made up 53.9% of the weight. Similar percentages were observed in the other years prior to 2016.

Just as for the national CPI, the stratum indices can be calculated for each region using the weighted geometric means of the price relatives. We use the geometric mean (Jevons

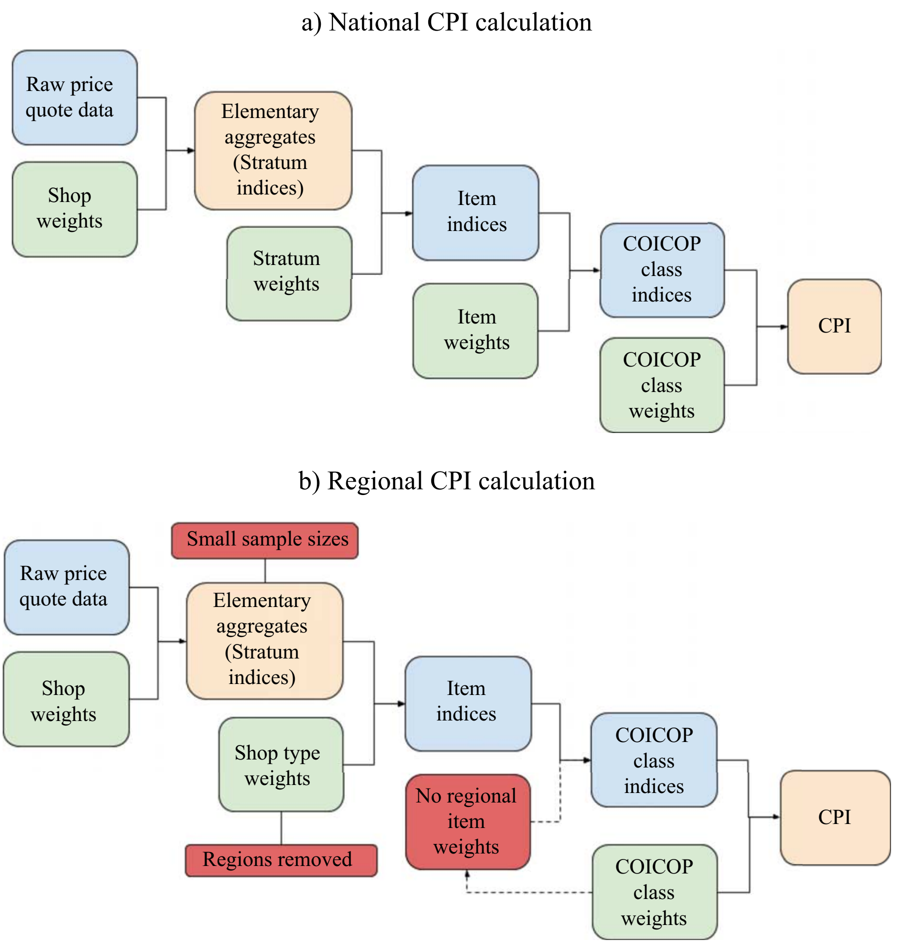

Fig. 1. Calculation of the national and regional CPI.

elementary aggregate formula) for all items, though in the national CPI the ratio of averages (Dutot formula) is sometimes used (ONS 2019). There are large numbers of strata which have extremely small sample sizes, including 5.4% with only one price measurement. The majority of the strata with small sample sizes are for independent shop prices. The geometric mean of very small samples is not very reliable, highlighting the primary limitation in constructing the regional CPI. Although it is inadvisable, for this experimental regional CPI the small sample sizes are treated as if they are satisfactory and the regional elementary aggregates are used.

With the prices aggregated into the stratum indices, the item indices can then be calculated by taking the arithmetic mean of these elementary aggregates weighted with the stratum weights. However, the stratum weights, which adjust for region and shop type in the national index, must not adjust for region here and reduce to shop type weights. Once the item indices are calculated for each region the available national item indices not represented in the price quote data can then be added to give the full set of item indices required to calculate COICOP class indices (except for the 1.7% excluded).

The next step is to aggregate item indices to give regional COICOP class indices. However, the item weights cannot be determined at the regional level, because the publicly available version of the LCF does not contain this detailed classification information (in principle the ONS could make this calculation on the full LCF data set). To overcome this problem we use a non-standard approach using the national (subscript n) item weights w_{nk} and class weights w_{nc} to first get the national proportions of item k within class c. We then multiply these national proportions by the estimated regional (subscript r) class weight w_{rc}, which gives approximations of regional item weights $\hat{w}_{rk}$:

$$\hat{w}_{rk} = w_{rc} \frac{w_{nk}}{w_{nc}}.$$

This ensures that the item weights sum up to the class weight for each region. Note that w_{rc} can be estimated using LCF data, as described in the following section.

Finally, the regional item indices and weights are used to derive the COICOP class indices and also the unchained regional CPI for each region. This process was replicated for all available years where price quote data and LCF data were readily available. The regional CPI series were then chained together using the same approach as the national CPI (ONS 2014) and indexed with January 2010 set to 100 for all regions.

The limitations encountered when constructing the regional CPI fall broadly into two categories –those due to small sample sizes, and those due to only the LCF survey being used for expenditure weights. These lead to the adjusted regional CPI framework shown in Figure 1b. These limitations require us to approximate the target concept and therefore render the regional CPIs less reliable than the national CPI. The extent of this unreliability is assessed by constructing experimental regional CPIs.

3.2. *Regional CPI Expenditure Weight Estimation*

The LCF survey data provide expenditures for a sample of households in each of the twelve regions of the UK. To get expenditure weights for the regional CPI the expenditures should be aggregated within a certain COICOP level and region. We found that the COICOP class level (COICOP4) was the most suitable, as more detailed levels had too many zero expenditures and the COICOP group level was not specific enough. Even at the class level, there were still some classes with few or no expenditures recorded at the region level, so we first inspect how well each class was represented in the LCF survey.

We use the term ‘observation’ here specifically to mean any household with non-zero expenditure recorded from a given COICOP class. Zero expenditures are still used in estimation, but for simplicity, we do not refer to these as observations. At COICOP class level, there can also be multiple observations (of more detailed expenditures) from the same household. Figure 2 shows the mean number of observations in the North East region across the available years. The North East is used as it generally has the smallest household sample in the LCF survey. A reference line at 30 observations is included in the plot, which shows that there are many classes with fewer than 30 observations. At the COICOP class level, we can see that Division 01 ‘Food and non-alcoholic beverages’ is well represented, as well as classes 5.6.1 ‘Non-durable household goods’, 11.1.1 ‘Restaurants and cafes’ and 12.1.2 ‘Appliances and products for personal care’. On the other hand, five out of 85 classes (5.9%) have no representation in the LCF data for the

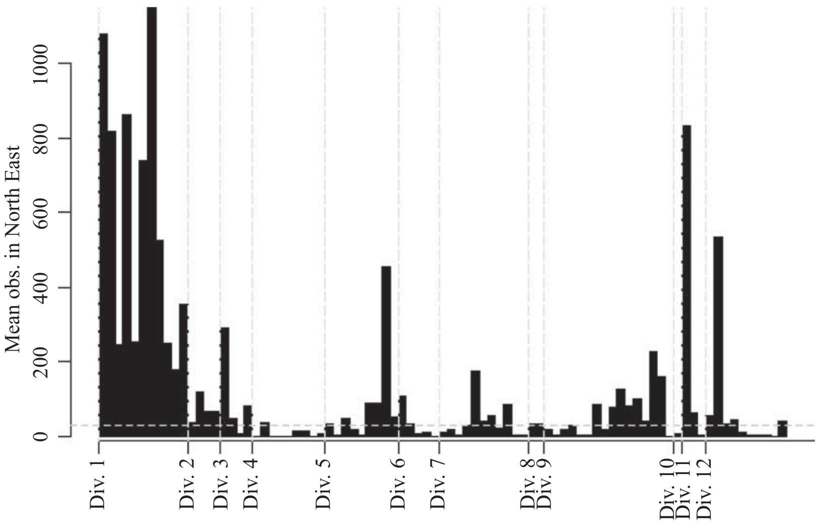

Fig. 2. North East region – mean annual number of observations from 2008–2014 for each COICOP class with a reference line at 30 observations.

available years, including water supply and sewerage, household repair services, hospital services and package holidays. These will all have zero weights in the provisional regional CPI. This is not just an issue for the North East region but for all regions, which are all reasonably comparable. This is a substantial limitation and can only be overcome through the use of additional data sources that are not publicly available at the regional level.

Although the number of observations is an important consideration, the actual expenditure must also be considered. The problem of having so few observations will be compounded when there is higher expenditure because it adds more weight to the regional CPI. The North East mean expenditures for COICOP classes are shown in Figure 3, with a

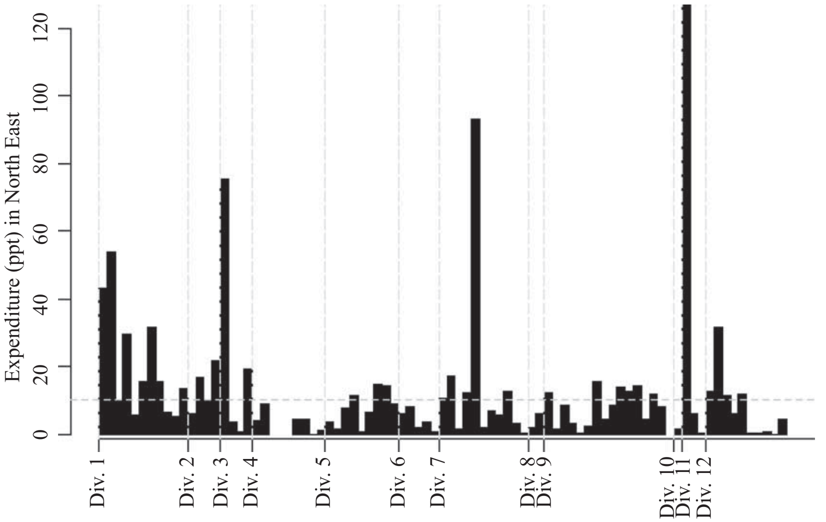

Fig. 3. North East region – estimates of relative expenditure by COICOP class with a reference line at 10 ppt.

reference line added at 10 ppt (1%). There is a lot of variation between classes, with five that make up more than 40 ppt and the remainder mostly under 10 ppt. The five largest classes, in order, are 11.1.1 'Restaurants and cafes', 7.2.2. 'Fuels and lubricants', 3.1.2. 'Garments', 1.1.2. 'Meat' and 1.1.1. 'Bread and cereals'. Note that the other regions also have highly variable relative expenditure across classes, and generally share the same largest classes.

To get direct estimates of the expenditure weights the following calculations are made. Let y_{ijct} denote the total household expenditure in pounds for COICOP class c in household j in region i and year t, and w_{ijct} be the provided household survey weight. Suppose that n_{it} is the number of sampled households in region i and year t, then the direct estimate of the mean household expenditure θ_{ict} is:

$$\hat{\theta}_{ict}^{direct} = \frac{\sum_{j=1}^{n_{it}} y_{ijct} w_{ijct}}{\sum_{j=1}^{n_{it}} w_{ijct}}. \tag{1}$$

Then the relative weights in ppt can be calculated using:

$$\hat{w}_{ict}^{direct} = 1000 \frac{\hat{\theta}_{ict}^{direct}}{\sum_c \hat{\theta}_{ict}^{direct}}. \tag{2}$$

These weights can then be used to generate the regional CPI series for 2010–2016, which is shown for the direct estimates in Figure 4b. This series can be compared to the series calculated with the national weights (but regional prices) in Figure 4a. This comparison makes it clear that the differences between regions and the general volatility of the series come primarily from the expenditure weights rather than the prices.

Due to the small sample sizes and lack of additional data, it is expected that there would be instability of the regional CPI series over time. To quantify this variability over time, we propose measuring the standard deviation of the first differences (SDFD) of the regional CPI series, that is $\frac{1}{T-2}\sum_{t=2}^{T}\left(y_t - y_{t-1}\right)^2$.

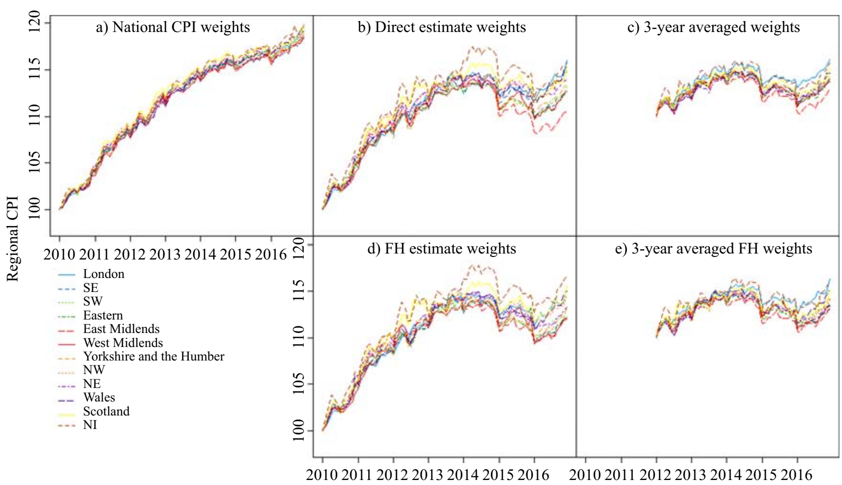

Fig. 4. Regional CPI series for different expenditure weights. Note that three-year averaged weights are rooted at 110 rather than 100.

The higher this SDFD measurement, the more variable the monthly changes in the index. As a comparison, we find that the national CPI between 2010 and 2016 has a SDFD of 0.29. This provides an approximate guide for an appropriate level of temporal variability. As shown in Figure 5 we find that the SDFD for the series with regional prices and national weights in Figure 4a ranges from 0.36 for Wales, to 0.46 for Northern Ireland. This corresponds to 1.24 and 1.60 times that of the national level of variability, from using regional prices alone. The SDFD for the regional prices and weights from the LCF data ranges from 0.48 for North East, to 0.87 for Northern Ireland, corresponding to 1.65 and 2.99 times that of the national SDFD respectively. In all but two regions, the majority of the additional variability derives from the expenditure weights rather than the prices. These two regions were the North East and South West. Averaged over regions, approximately 60% of the additional variability according to the SDFD is due to the expenditure weights. In the next section we explore smoothing and SAE methods to reduce this volatility.

4. Improving the Expenditure Weights

4.1. Smoothing and Small Area Estimation

Since direct estimation of the expenditure weights leads to substantial increases in temporal variability, we assess whether smoothing methods and small area estimation can make improvements. We test whether the three-year moving average of the expenditure weights substantially reduces the temporal instability. This serves to increase the sample sizes, as well as strengthen the temporal correlation. Regional CPIs using these smoothed weights were calculated and the series are shown in Figure 4c. Two years are removed due to the smoothing, and the resulting SDFDs range from 0.41 for the North East to 0.72 for Northern Ireland. This amounts to an 8–20% reduction in the variability. Hence there is evidence of improved stability compared to the one-year direct estimates, although only moderate improvement. This suggests that the three-year moving average approach does not eliminate all problems of volatility in the expenditure weights.

Small sample sizes are demonstrably a substantial limitation to developing reliable regional CPIs in the UK. Like smoothing, SAE can potentially improve the reliability,

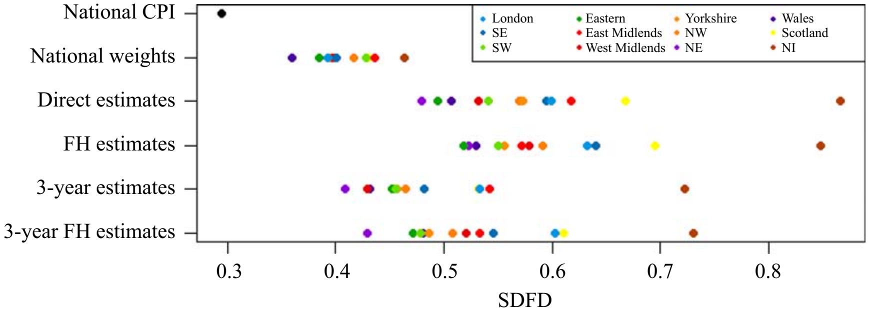

Fig. 5. Standard deviation of first differences of each regional CPI series and the national CPI.

as it utilises model-based methods to borrow strength from a wider or population-level data source to improve estimates for small domains. Although SAE methods can improve reliability by reducing the variance, this comes at the cost of introducing bias. Due to the high variance this seems a worthwhile trade-off in this case. For a general overview of SAE methodology, Pfeffermann (2013), Rao and Molina (2015) and Tzavidis et al. (2018) are highly recommended. SAE is most effective when a wider data source with high quality is available, containing strong predictors for the variables of interest within each region. In the case of the regional CPI, having region-level predictors of expenditure within COICOP classes will improve the precision of the estimates.

For the price quotes, it is difficult to utilise the beneficial aspects of SAE. This is because the sampling units are shops or prices. To effectively use SAE, comprehensive and informative data about the shop or price population would be required. This could include population numbers of shop types (independent and multiple) by region, number of supermarkets, and shop type by COICOP level. Alternative data sources from web-scraping or scanner data might also provide suitable predictors, although such sources typically have only partial coverage. However, such data sources were not available. Furthermore, the expenditure weights cause more of the variability than the price quotes, so SAE should be more effective at improving estimates of the weights.

SAE is much more feasible for the expenditure estimates, because the sampling units for expenditure data are households, about which plenty of national and regional data are available. The detailed national-level data for potential predictors (such as household types, total salary, total expenditure, head of household's age and the number of children) can be used in a model to get improved estimates of the expenditure for each class. We aim to use these small area models to give regional expenditure estimates with lower variances, and in particular look for them to be relatively smooth across years.

4.1.1. Fay-Herriot Models

The SAE method used for expenditure estimation was the Fay-Herriot (FH) model (Fay and Herriot 1979). The FH model is a commonly used region-level model for SAE (Rao and Molina 2015, sec. 4.2, chap. 6), and comprises of two stages. The first stage simply models the sampling variation of the direct estimate which was defined in Equation (1):

$$\hat{\theta}_{ict}^{direct} = \theta_{ict} + \varepsilon_{ict}$$

where the sampling errors ε_{ict} are assumed to be independent and normally distributed with $\varepsilon_{ict} \sim N(0, \sigma^2_{\varepsilon_{ict}})$. From the total household expenditure and the sampling weights we get regional direct estimators for each COICOP class and year. However, the true variance of this estimator cannot be determined from only sample data. For this reason, the variance $\sigma^2_{\varepsilon_{ict}}$ must be estimated. Possible options are the Poisson approximation or a bootstrap procedure. We used the bootstrap method of the laeken package (Alfons and Templ 2013) in R (R Core Team 2019). For each COICOP class within every year and each region i a bootstrap sample was drawn from the sample data with replacement. Using the household weights and expenditures within each bootstrap sample, the corresponding direct estimator was obtained. This was repeated many times and the variance $\sigma^2_{\varepsilon_{ict}}$ was then estimated by taking the sample variance of all bootstrap direct estimates.

The second stage of the FH model is to fit a linear model which can be used to predict θ_{ict}:

$$\theta_{ict} = x_{ict}^T \beta + u_{ict}$$

where x_{ict}^T denotes the region-level covariates for year t, β denotes the regression parameter vector, and u_{ict} represents the random effects which are assumed to be $u_{ict} \sim N(0, \sigma^2_{u_{ct}})$. The combination of the two stages of modelling leads to the combined FH model:

$$\hat{\theta}_{ict}^{direct} = x_{ict}^T \beta + u_{ict} + \varepsilon_{ict}.$$

The estimates $(\hat{\beta}, \hat{u}_{ict}, \hat{\sigma}^2_{u_{ct}})$ of these unknown parameters can be estimated using a standard linear random-effects model. From this the FH estimates can be derived as:

$$\begin{aligned}\hat{\theta}_{ict}^{FH} &= x_{ict}^T \hat{\beta} + \hat{u}_{ict}\\ &= \gamma_{ict}\hat{\theta}_{ict}^{direct} + (1-\gamma_{ict})x_{ict}^T\hat{\beta}\end{aligned}$$

where $\gamma_{ict} = \hat{\sigma}^2_{u_{ct}}(\hat{\sigma}^2_{u_{ct}} + \sigma^2_{\varepsilon_{ict}})^{-1}$. In cases when an area has zero observations the estimator simply becomes: $\hat{\theta}_{ict}^{FH} = x_{ict}^T\hat{\beta}$. Estimates of the precision of the FH estimates can be made using the mean squared error (MSE) which is estimated using restricted maximum likelihood (REML). Further details can be found in Rao and Molina (2015, chap. 6). Once the estimates are calculated, the weights can be created using the same adjustment as in Equation (2).

One of the improvements of SAE comes from borrowing strength from other areas, using the association between the region-level covariates and the regional expenditure. Region-level associations strengthen each region's estimate by using the region-level covariates, which are assumed to be informative. In the next section, we describe how the covariates were selected for use in the FH models.

A challenge for the FH model in this situation is due to having only twelve regions. With so few regions, model assumptions are more difficult to assess, covariate associations are more likely to occur by chance and the number of covariates that can be included in the model is restricted because there are few degrees of freedom. Another notable limitation in the application of the FH model to the expenditure weights is that some classes have no expenditure in the region. This becomes problematic, as zero expenditure for more than a few regions will lead to violations of the normality assumptions. For some classes, a zero-inflated model (Pfeffermann et al. 2008; Chandra and Sud 2012) may be beneficial, which we leave for future research.

The same FH model was also applied to the three-year averaged data to assess the collective impact of both smoothing and SAE on the regional CPI series.

4.1.2. Covariate Variable Selection

The LCF survey provides a large number of variables at the regional level which can be used to estimate expenditure. These variables relate to socioeconomic status, household composition and household features, for example, tenure type, number of adults, weekly income. These variables were aggregated to the region level. For a FH model to be effective at estimating expenditure, the region-level covariates should be predictive of the expenditure of the COICOP classes. The challenge is to select the best combination of variables that ensures the relationship is predictive but not over-fitted to the sampled data. This over-fitting is especially a concern since there are only twelve regions, and hence twelve points from which to fit a model. Furthermore, the covariates should not have high multicollinearity, as this can greatly exaggerate over-fitting. Over-fitting will lead to small area estimates with under-estimated precision, as well as overly biased point estimates. Hence the explanatory variables must be selected carefully.

The variables were chosen based on associations with the class expenditure for a pooled data set across all years from 2010 to 2016. This ensures that the covariate is predictive across all years rather than a certain year. This will also ensure consistency across time. A forward selection approach using AIC was used to select the variables for each COICOP class, with at most five selected variables. We made five the maximum since any more would be superfluous when estimating only twelve regions. At each step, the multicollinearity was assessed using the variance inflation factor (VIF). If the VIF was greater than ten then no more variables were added. This ensured that a minimal number of variables were selected and that none of the variables were highly collinear.

4.1.3. Model Assessment

A FH model relies on a strong level of prediction with explanatory variables. Figure 6 shows the R^2 values of the fitted linear models averaged over the years for each COICOP class. It shows that in Divisions 01 and 02, which include food and alcoholic beverages, the R^2 is generally high, but some classes have low R^2 values. For example, COICOP class 12.6.2 ‘Other financial services’ with a mean R^2 of just 0.03, which will be unlikely to improve the expenditure estimate in the FH model. Classes with higher R^2 may be a better reflection of the general economic differences between regions.

With the FH models fitted, it remains to be seen what the effect on the stability of the expenditure estimates is when we use the model predictions in place of the direct estimates.

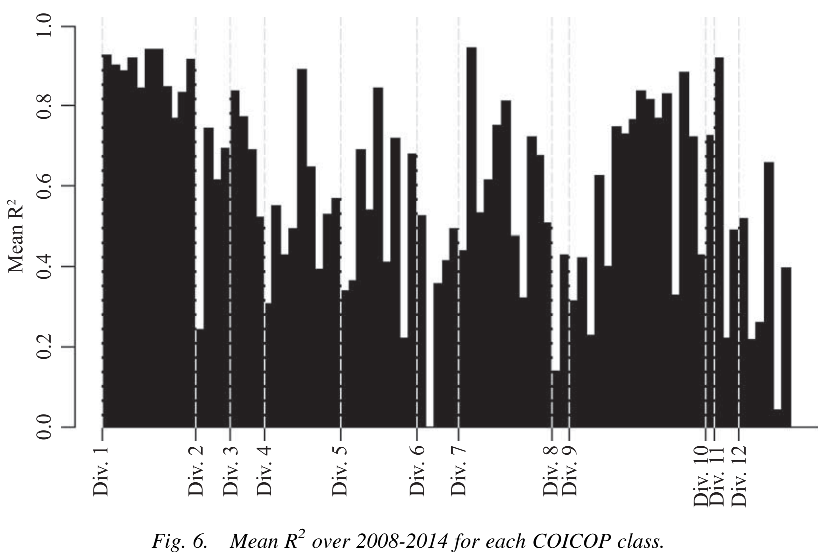

Fig. 6. Mean R^2 over 2008-2014 for each COICOP class.

As part of the model assessment, the assumptions of the models were checked, particularly the normality assumption of $\hat{u}_{ict} = x_{ic}^{T}\hat{\beta} - \hat{\theta}_{ict}^{FH}$. A Shapiro-Wilk test was used, and based on this test, there was evidence to reject the normal distribution for many classes for each year. Across years, between 26 and 43 (out of 80) classes were not rejected, with a median of 36. However, as the focus is on improving the temporal stability of the estimates we include all the estimates in calculating the experimental index, even if there is strong evidence to reject the normality assumptions. Although the FH model is somewhat robust to these violations (Lahiri and Rao 1995), further development of the models may provide a better basis for such estimation.

4.2. *Assessment of Fay-Herriot Estimates*

To assess the effect that FH estimation has on the expenditure we first measure how different the FH estimates are from the direct estimates. Figure 7 again uses the North East region as an example, and shows the percent difference between the FH and direct estimate, averaged over the seven years. This reveals up to a 30% difference in the estimates with many COICOP classes showing non-trivial relative differences. Clearly, FH estimation has some effect, but it is unclear what effect this is. Some classes have no percent difference because FH estimates could not be calculated, where too many regions had zero reported expenditure. In total, FH estimates could not be calculated for 16 of the 85 COICOP classes (18.8%). In these cases the direct estimates are used for the weights.

We expect that expenditure patterns are in reality rather stable between adjacent years, and change slowly as consumer spending is influenced by changes in products and their availability, particularly at higher levels of aggregation in COICOP. Therefore we judge that the more stable the estimates are over 2008–2014, the better the estimates are and the more stable the regional CPI will be. This is because the instability is likely caused by small sample sizes. Note that some expenditure patterns may vary substantially over time, so an expert opinion on what level of variability is realistic would need to be considered too.

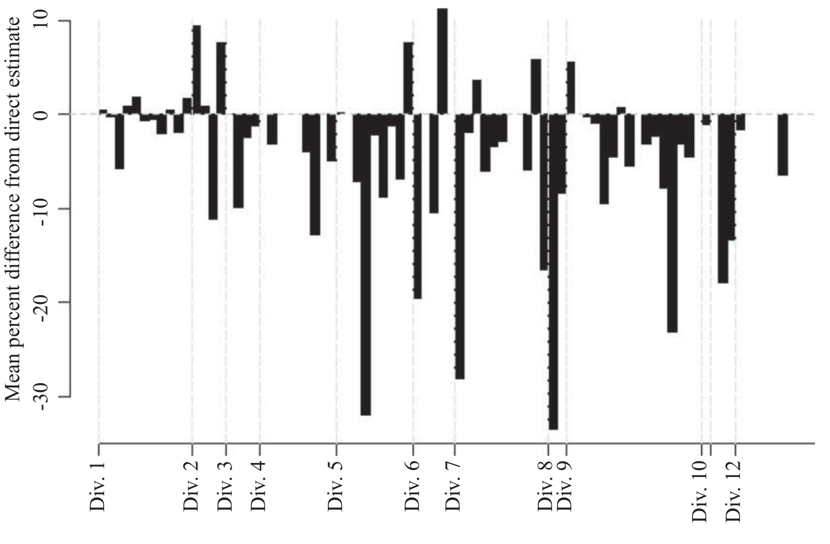

Fig. 7. North East region – mean percent difference between FH estimate and direct estimate.

To measure this stability over time we measure the variability of the yearly estimates of expenditure; this will include a small element of real change in expenditure patterns, but we expect that this is much smaller than the random variation we are trying to smooth using SAE. Typically the standard deviation or variance is used to measure variability, however this will not be appropriate in this case, because the variance is greater for COICOP classes with higher expenditure. To accommodate this, we use the coefficient of variation (CV) which is the standard deviation divided by the mean. This ensures the measure is standardised by the amount of expenditure, hence making the metric comparable across all COICOP classes. This 'temporal' CV is calculated using:

$$CV_{ic} = \frac{SD_t\left(\hat{\theta}_{ict}\right)}{Mean_t\left(\hat{\theta}_{ict}\right)}$$

where SD_t and $Mean_t$ are the standard deviation and mean of the estimates across years t respectively. We use CV_{ic} to compare the stability of the FH estimates compared to the direct estimates. The lower this temporal CV, the more stable the estimates over time.

Figure 8 compares this temporal CV for the direct and FH estimates for all twelve regions. A positive difference in temporal CV indicates that the direct estimate is less

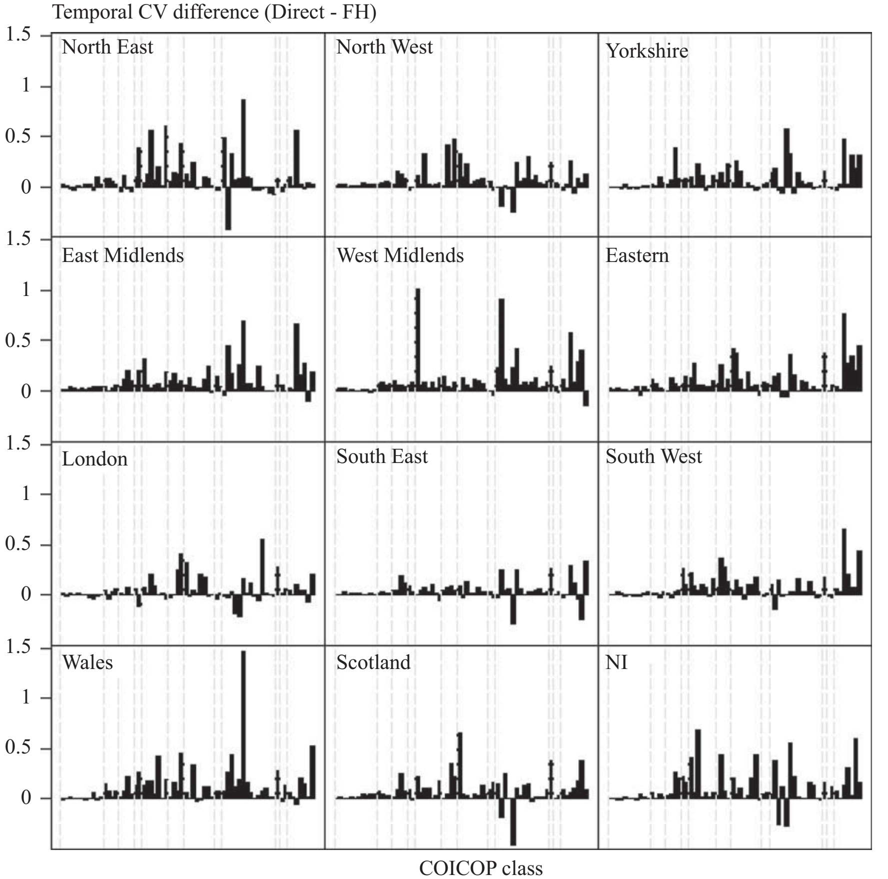

Fig. 8. All regions – difference in temporal CV between direct and FH estimates.

stable, which is the case for the vast majority of classes and regions. This suggests that FH estimation is generally improving the stability compared to the direct estimates. Notably, in well-represented classes like Division 01 the differences are very small.

Figure 9 displays the North East region estimates of γ_{ict} for each class averaged over all seven years. The higher the value of γ_{ict} the more the FH estimate utilises the data directly as opposed to the model-based component. There is a lot of variation between COICOP classes, ranging from 0 to 0.7, so there is no clear trend about what types of classes have higher values of γ_{ict}. Again, other regions showed similar patterns.

Table 3 reports the ten COICOP classes that have the greatest improvements in temporal CV due to FH estimation. These ten classes have generally few observations, ranging from 6 to 61. Interestingly, the R^2 values are not particularly high, which suggests that FH estimation does not require strongly predictive explanatory variables to provide additional stability to the estimates. Classes 6.1.2/3 'Other medical and therapeutic equipment' and 9.2.1/2 'Major durables for in/outdoor recreation' each have relatively large ppt values, which shows that it is not just the trivially small classes (such as 12.7 'Other services') that show improvements. The last column in Table 3 shows the mean values of γ_{ict} which range between 0.10 and 0.38 which is quite typical of all classes.

Figure 10 shows more broadly the effect of sample size on improved stability due to FH estimation. A smoothing spline has also been added to give an idea of the average effect for varying numbers of observations. The results show that for COICOP classes with relatively few observations, the benefit of the FH estimation is generally better although highly variable for all regions. We also see how the COICOP classes with many observations show negligible benefit from FH estimation. The improvement of FH estimation becomes reasonably small after approximately 100 household observations.

In combining all these results, we consider four attributes of the COICOP classes which relate to their suitability for SAE. These four attributes are:

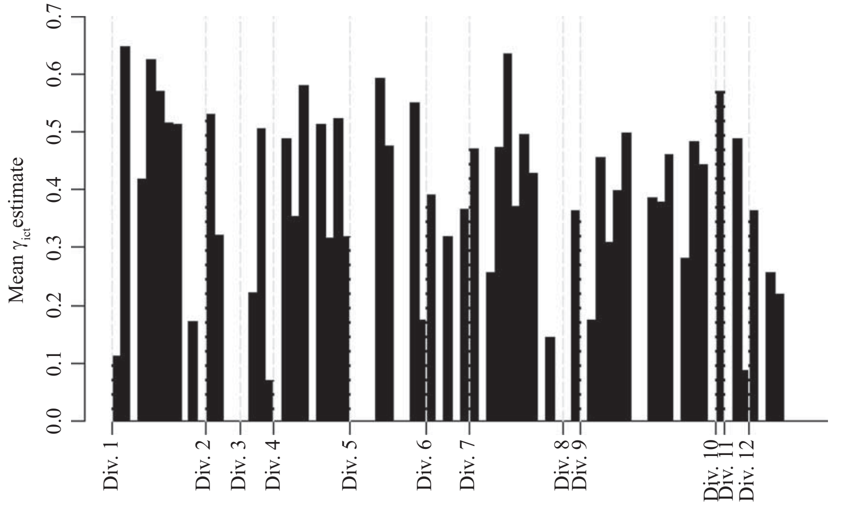

Fig. 9. North East region – mean estimates of γ_{ict} across years for each COICOP class.

Table 3. Top ten COICOP classes with the most improvement in stability due to FH estimation.

COICOP class	Temporal CV difference	Mean number of observations	Mean ppt	Mean R^2	Mean γ_{ict}
9.2.1/2	0.46	6.5	9.11	0.23	0.15
12.3.1	0.37	58.2	5.64	0.22	0.10
10	0.23	11.2	5.23	0.33	0.16
5.1.1	0.23	61.0	3.01	0.39	0.26
6.2.2	0.23	20.6	6.23	0.53	0.32
12.7	0.23	72.1	0.46	0.40	0.13
5.3.1	0.22	39.4	4.40	0.34	0.33
7.1.1A	0.20	16.6	2.91	0.36	0.38
6.1.2/3	0.19	59.7	15.03	0.22	0.21
3.1.4	0.15	15.2	1.20	0.69	0.32

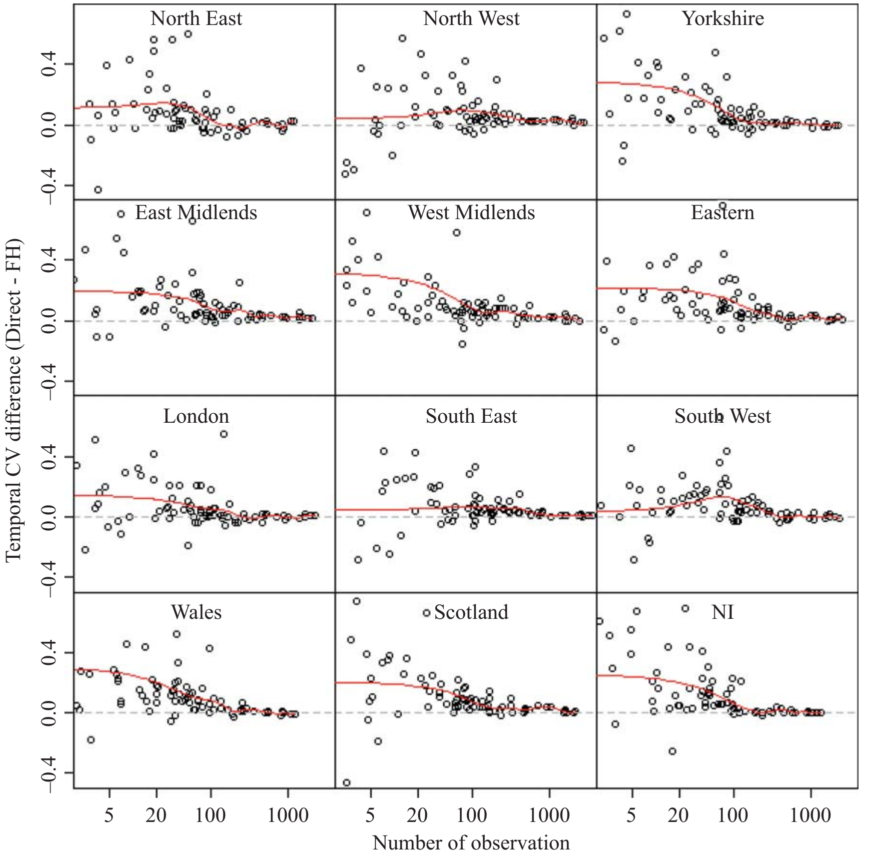

Fig. 10. Temporal CV difference by sample size within each region, with a trend line added. Each point is a COICOP class.

1. Observations recorded in all regions for all years – 63 out of 85 classes (74.1%) meet this criterion. This means at least one reported price within a class for each region and year,

2. The number of observations not being so large that SAE remains useful, chosen to be all COICOP classes where all regions have at most 100 household observations – 63 out of 85 classes (74.1%),
3. A non-negligible expenditure share in ppt, chosen to be the COICOP classes which have at least 0.5 ppt share in all regions and years (in line with the conceptual framework in section 2.1) – 58 out of 85 classes (68.2%), and
4. A non-negligible ($>$ 0.03) decrease in temporal CV for at least one region when using FH estimation – 60 out of 85 classes (70.6%).

These four attributes are not mutually exclusive, so the total number of COICOP classes which possess all four attributes is 36 out of 85 (42.4%). Hence, based on these criteria, 42.4% of COICOP classes have distinguishable improvements in the stability of the expenditure estimates through the use of SAE.

The regional CPI series with the FH estimates for each year and also FH estimates with weights based on the three-year average are shown in Figure 4d and Figure 4e. In comparison to the series with the direct weights, all regions but Northern Ireland appear much closer together. This is expected given that SAE adds national-level information, making the estimates closer to the national value. Generally, the observable differences between series with direct and FH expenditure estimates are not large. To assess in more detail we again calculate the SDFD for the two FH-based series, these are displayed for each region in Figure 5, as well as for the other regional CPI series. The SDFD metric shows that FH estimation does not generally decrease the variability over time compared to the direct estimates. There is a slight increase in the SDFD for most regions. So it appears that while SAE appears to make inter-regional differences smaller in the series, it does so without improving the temporal stability, and appears to mildly increase variation over time.

The results show that smoothing and SAE using FH models for individual classes can improve the stability of the expenditure weights in some ways, but the smoothing appears to have a greater effect. The classes that benefit the most tended to have fewer than 100 observations, but also with enough observations for a good model to be fitted. Although FH models do ensure that the regional indices stay closer together, it does not provide reduced temporal variability for the regional CPIs.

5. Discussion

5.1. Regional Indices and Data Sources

We have shown that it is possible to construct a regional CPI series from the available data sources. Although these experimental regional CPIs are somewhat useful, the reliability of specific components of the data and procedures is generally low. Small sample sizes and absence of data sources create increased irregularities and variability over time. We show that the source of this variability is generally from the expenditure weights more than the price quotes.

Some further exploration of data sources is possible. ONS has produced experimental regional HFCE estimates, which are balanced through the national accounts (ONS 2018). These are a less detailed level of the COICOP classification than is needed for the CPI, but could be integrated into a framework of weight calculation. Alternative data sources such as web-scraped prices and scanner data may also provide useful regional data sources or predictors for either or both of the prices and weights. These could improve the expenditure weights and make the methods more similar to the national CPI. Incorporating regional estimates of owner occupiers' housing costs (OOH) and council tax to produce a regional CPIH would also be useful, and is likely to show greater regional differences, because there are large disparities in housing costs between regions.

It would also be worthwhile exploring whether it is possible to calculate a version of the regional item weights which were not available for our analysis. Marchetti and Secondi (2017) go one step further by adjusting the regional expenditure estimates for differences in PPPs, which is potentially an important extension to the conceptual framework of Subsection 2.1.

5.2. Smoothing, SAE and Bias-Variance Trade-Off

Smoothing methods like the three-year moving average were shown to reduce the variability moderately. Although FH estimation improved temporal stability of the expenditure weights there was no evidence that this reduced the temporal variability of the regional CPI series. We conclude that smoothing and SAE do generally make improvements to the series and the stability of the expenditure weight estimates, and are plausible options for improving the reliability of the regional CPI. Fengki et al. (2020) have applied FH models to estimate regional CPIs in Indonesia (using the CPI as a target, because there are city CPIs available for modelling), with some improvement over direct estimation, but in their case too, further research is needed to deal with data deficiencies.

Smoothing and SAE both involve a bias-variance trade-off, so using them in a regional CPI results in biased estimates of the regional price indices. However, since the main purpose of the index is to provide regional estimates of *inflation*, a bias may be acceptable as long as it evolves slowly, and as long as it accompanies a substantial reduction in the variability of the index. However, for long-run comparisons the bias may be important, particularly where it has different effects in regions of different size (as suggested by the estimates of γ_{ict} (Table 3, Figure 9)). A periodic benchmarking may therefore be needed to correct the path of the index if measures of long-term change are important.

It would be useful to have quality targets that a regional CPI or CPIH should aim to achieve, particularly an estimate of the variance to identify an acceptable level of volatility. How to measure the variance of a CPI is an open question, (see Smith 2021; O'Donoghue 2017; Zimmermann et al. 2020, sec. 2.a.6) for some initial assessments at national level for the UK. This work should be extended to regional measures.

5.3. Model Extensions

The FH model used does not account for the zero-inflated, longitudinal or compositional properties of the data, and these could be addressed using more advanced methods. Methods have been proposed for small area estimation with zero-inflated data (Pfeffermann et al. 2008; Chandra and Sud 2012), and it would be interesting to explore these. In particular, accounting properly for the observed zeroes in expenditure might allow modelling at a more detailed level in the COICOP classification.

Extension of the FH model to account for correlation over time should improve the small area estimates. Esteban et al. (2016) provide a comprehensive review of the literature on temporal extensions to the FH model that – among many variations – includes a model with an autoregressive structure in the sampling errors (Choudry and Rao 1989) and a model with an autoregressive structure in the random effects (Esteban et al. 2011). A further extension of the FH model simultaneously accounts for spatial and temporal effects (Marhuenda et al. 2013).

A further line of investigation is to treat the estimation of weights as the estimate of a composition. Scealy and Welsh (2017) have developed an approach to estimate expenditure proportions as compositions within small domains of a population, different from the FH models, and Esteban et al. (2020) provide a FH model for compositions. Applying similar methods may give regional expenditure weights for the regional CPI with smaller variances and less change from year to year which would lead to fewer irregularities and smoother indices.

5.4. Conclusions

The work reported here is a stepping-stone to the development of a regional CPI in the UK. We have provided a clear framework, and highlighted the deficiencies and issues faced in calculating a regional temporal CPI with available data sources. Smoothing and small area estimation offer reasonable reductions in the volatility of the weights, and could be used in a regional CPI, if the addition of other data sources does not provide the necessary level of reliability. Assessing whether the bias induced by these methods in the index affects their suitability remains an open question.

6. References

Alfons, A., and M. Templ. 2013. “Estimation of Social Exclusion Indicators from Complex Surveys: The R package laeken.” *Journal of Statistical Software* 54, no. 15: 1–25. DOI: www.dx.doi.org/10.18637/jss.v054.i15.

Baran, D., and J. O’Donoghue. 2002. “Price levels in 2000 for London and the regions compared with the national average.” *Economic Trends* 578: 28–38. Available at: www.escoe.ac.uk/research/historical-data/etarticles/ (accessed April 2021).

BLS. 2018. *BLS handbook of methods: Chapter 17 The Consumer Price Index*. Washington DC. Available at: www.bls.gov/opub/hom/pdf/homch17.pdf (accessed April 2021).

Borooah, V.K., P.P.L. McGregor, P.M. McKee, and G.E. Mulholland. 1996. “Cost of living differences between the regions of the United Kingdom.” In *New Inequalities. The Changing Distribution of Income and Wealth in the United Kingdom*, edited by J.Hills: 103–132. Cambridge: Cambridge University Press.

Brown, M., R. de Haas, and V. Sokolov. 2018. “Regional Inflation, Banking Integration, and Dollarization.” *Review of Finance* 22 8(6): 2073–2108. DOI: www.doi.org/10.1093/rof/rfx021.

Chandra, H. and U.C. Sud. 2012. "Small area estimation for zero-inflated data." *Communications in Statistics–Simulation and Computation* 41: 632–643. DOI: www.dx.doi.org/10.1080/03610918.2011.598991.

Choudry, G.H., and J.N.K. Rao. 1989. "Small area estimation using models that combine time series and cross sectional data." In Proceedings of Statistics Canada Symposium on Analysis of Data in Time, edited by A.C. Singh, and P. Whitridge: 67–74. Ottawa: Statistics Canada.

Department for Environment, Food and Rural Affairs and Office for National Statistics. 2019. Living Costs and Food Survey, 2008-2014. [data collection]. 3rd Edition. UK Data Service. SN: 7992, www.doi.org/10.5255/UKDA-SN-7992-4 (DOI for 2014 only) and also SN: 6385, 6655, 6945, 7272, 7472, 7702).

Duran, H.E. 2016. "Inflation differentials across regions in Turkey." *South East European Journal of Economics and Business* 11(1): 7–17. DOI: www.doi.org/10.1515/jeb-2016-0001.

Esteban, M.D., M.J. Lombardía, E. López-Vizcaíno, D. Morales, and A. Pérez, A. 2020. "Small area estimation of proportions under area-level compositional mixed models." *TEST* 29: 793–818. DOI: www.doi.org/10.1007/s11749-019-00688-w.

Esteban, M.D., D. Morales, A. Pérez, and L. Santamaría. 2011. "Two area-level time models for estimating small area poverty indicators." *Journal of the Indian Society of Agricultural Statistics* 66: 75–89.

Esteban, M.D., D. Morales, and A. Pérez. 2016. *Analysis of Poverty Data by Small Area Estimation, First Edition*. Edited by Monica Pratesi. Chichester: John Wiley & Sons.

EU regulation 1998. EU regulation 1687/98 on HICP available at: (www.eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri = CELEX:31998R1687&from = EN) (accessed April 2021).

Fay, R., and R. Herriot. 1979. "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data." *Journal of the American Statistical Association* 74 (366): 269–277. DOI: www.doi.org/10.2307/2286322.

Fengki, A.O., K.A. Notodiputro, and K. Sadik. 2020. "Bisakah memperoleh statistik indeks harga konsumen tingkat provinsi di Indonesia dengan ketelitian yang lebih baik? (Can provincial level consumer price indices statistics in Indonesia be obtained with better accuracy?)." *Seminar Nasional Official Statistics* 1: 297–306. DOI: www.doi.org/10.34123/semnasoffstat.v2019i1.178.

Fenwick, D., and J. O'Donoghue. 2003. "Developing estimates of relative regional consumer price levels." *Economic Trends* 599: 72–83. Available at: www.webarchive.nationalarchives.gov.uk/20160120195958/www.ons.gov.uk/ons/rel/cpi/regional-consumer-price-levels/developing-estimates-of-relative-regional-consumer-price-levels/-developing-estimates-of-relative-regional-consumer-price-levels-.pdf (accessed April 2021).

Gajewski, P. 2017. "Sources of regional inflation in Poland." *Eastern European Economics* 55(3): 261–276. DOI: www.doi.org/10.1080/00128775.2017.1284572.

Hayes, P. 2005. "Estimating UK regional price indices, 1974–96." *Regional Studies* 39: 333–344. DOI: www.doi.org/10.1080/00343400500087216.

Kosfeld, R., H.-F. Eckey, and M. Schüßler. 2009. "Ökonometrische Messung regionaler Preisniveaus auf der Basis örtlich beschränkter Erhebungen". German Council for

Social and Economic Data (RatSWD) Research Notes: 33. Available at: www.hdl.handle.net/10419/43644 (accessed April 2021).

Lahiri, P., and J.N.K. Rao. 1995. "Robust estimation of mean squared error of small area estimators." *Journal of the American Statistical Association* 90, no. 430: 758–766. DOI: www.doi.org/10.1080/01621459.1995.10476570.

Marchetti S., and L. Secondi. 2017. "Estimates of Household Consumption Expenditure at Provincial Level in Italy by Using Small Area Estimation Methods: ''Real'' Comparisons Using Purchasing Power Parities." *Social Indicators Research* 131: 215–234. DOI: www.doi.org/10.1007/s11205-016-1230-8.

Marchetti S., G. Bertarelli, L. Biggeri, G. Giusti, M. Pratesi, and F. Schirripa-Spagnolo. 2019. "Small area poverty indicators adjusted using local price indexes." Italian Conference on Survey Methodology (ITACOSM), 5–7 June, 2019, Florence. Available at: www.centrodagum.it/wp-content/uploads/2019/07/Presentation_Pratesi.pdf (accessed April 2021).

Marhuenda, Y., I. Molina, and D. Morales. 2013. "Small area estimation with spatio-temporal Fay–Herriot models." *Computational Statistics and Data Analysis* 58: 308–325. DOI: www.doi.org/10.1016/j.csda.2012.09.002.

Nagayasu, J. 2011. "Heterogeneity and convergence of regional inflation (prices)." *Journal of Macroeconomics* 33: 711–723. DOI: www.doi.org/10.1016/j.jmacro.2011.07.002.

O'Donoghue, J. 2017. "The effect of variance in the weights on the CPI and RPI." *Survey Methodology Bulletin* 77: 1–27. Available at: www.ons.gov.uk/file?uri = /methodology/methodologicalpublications/generalmethodology/surveymethodologybulletin/surveymethodologybulletinno.77autumn2017a.pdf (accessed April 2021).

ONS. 2011. *UK Relative Regional Consumer Price levels for Goods and Services for* 2010. ONS, Newport. Available at: www.webarchive.nationalarchives.gov.uk/20160108054525/www.ons.gov.uk/ons/rel/cpi/regional-consumer-price-levels/2010/uk-relative-regional-consumer-price-levels-for-goods-and-services-for-2010.pdf (accessed April 2021).

ONS. 2014. *Consumer price indices technical manual.* Available at: www.ons.gov.uk/ons/rel/cpi/consumer-price-indices--technical-manual/2014/index.html (accessed April 2021).

ONS. 2018. *Development of regional household expenditure measures.* Office for National Statistics, Newport. Available at: www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome/articles/developmentofregionalhouseholdexpendituremeasures/latest (accessed April 2021).

ONS. 2019. *Consumer prices indices technical manual.* Available at: www.ons.gov.uk/economy/inflationandpriceindices/methodologies/consumerpricesindicestechnicalmanual2019 (accessed April 2021).

ONS. 2020. *Consumer price inflation item indices and price quotes.* Available at: www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceindicescpiandretailpricesindexrpiitemindicesandpricequotes/ (accessed April 2021).

Pfeffermann, D. 2013. "New important developments in small area estimation." *Statistical Science* 28(1): 40–68. DOI: www.doi.org/10.1214/12-STS395.

Pfeffermann, D., B. Terryn, and F.A. Moura. 2008. "Small area estimation under a two-part random effects model with application to estimation of literacy in developing countries." *Survey Methodology,* 34: 235–249. Available at: www.eprints.soton.ac.uk/153903/1/Pfeffermann_Moura-Terryn.pdf (accessed April 2021).

Purwono, R., M.Z. Yasin, and M.K. Mubin. 2020. "Explaining regional inflation programmes in Indonesia: Does inflation rate converge?" *Economic Change and Restructuring* 53: 571–590. DOI: www.doi.org/10.1007/s10644-020-09264-x.

R Core Team. 2019. "R: A language and environment for statistical computing". *R Foundation for Statistical Computing*, Vienna, Austria. Available at: www.R-project.org (accessed April 2021).

Rao, J.N., and I. Molina. 2015. *Small area estimation (Second Edition).* Hoboken: John Wiley & Sons.

Rienzo, C. 2017. "Real wages, wage inequality and the regional cost-of-living in the UK." *Empirical Economics* 52: 1309–1335. DOI: www.doi.org/10.1007/s00181-016-1122-4.

RPI Advisory Committee. 1971. *Proposals for retail prices indices for regions*. H.M. Stationery Office, London.

Scealy, J. and Welsh, A. 2017. "A Directional Mixed Effects Model for Compositional Expenditure Data" *Journal of the American Statistical Association,* 112, 24–36. DOI: www.doi.org/10.1080/01621459.2016.1189336.

Smith, P.A. 2021. "Estimating sampling errors in Consumer Price Indices." *International Statistical Review*. DOI: www.doi.org/10.1111/insr.12438.

Statistisches Bundesamt. 2018. *Preise Verbraucherpreisindex für Deutschland Qualitätsbericht. Statistisches Bundesamt, Wiesbaden.* Available at: www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Preise/verbraucherpreis.pdf?__blob = publicationFile (accessed April 2021).

Statistisches Bundesamt. 2020. *Regional CPIs for Germany*. Available at: www.www-genesis.destatis.de/genesis/online (search "61111-0011"), (accessed April 2021).

Statistics Bureau of Japan. 2020. *Regional CPIs for Japan*. Available at: www.e-stat.go.jp/en/stat-search/files?page = 1&layout = datalist&toukei = 00200573&tstat = 000001084976&cycle = 1&year = 20200&month = 11010303&tclass1 = 000001085955 (accessed April 2021).

Swanson, D.C., S.K. Hauge, and M.L. Schmidt. 1999. "Evaluation of Composite Estimation Methods for Cost Weights in the CPI." In *Proceedings of the Section on Survey Research Methods, American Statistical Association.* Washington D.C. American Statistical Association. Available at: www.bls.gov/osmr/research-papers/1999/st990050.htm (accessed April 2021).

Tillmann, P. 2013. "Inflation targeting and regional inflation persistence: Evidence from Korea." *Pacific Economic Review* 18(2): 147–161. DOI: www.doi.org/10.1111/1468-0106.12016.

Tzavidis, N., L.C. Zhang, A. Luna, T. Schmid, and N. Rojas-Perilla. 2018. "From start to finish: A framework for the production of small area official statistics." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 181(4): 927–979. DOI: www.doi.org/10.1111/rssa.12364.

UK Statistics Authority. 2013. *Statistics on Consumer Price Inflation – Assessment Report 257.* UK Statistics Authority, London. Available at: www.osr.statisticsauthority.gov.uk/wp-content/uploads/2015/11/images-assessmentreport257statisticsonconsumerpriceinflatio_tcm97-43135.pdf (accessed April 2021).

Weber, A.A., and G.W. Beck. 2005. *Price stability, inflation convergence and diversity in EMU: Does one size fit all?* CFS Working Paper 2005(30). Goethe University, Center for Financial Studies (CFS), Frankfurt. Available at: www.nbn-resolving.de/urn:nbn:de:hebis:30-23427 (accessed April 2021).

Wingfield, D., D. Fenwick, and K. Smith. 2005. "Relative regional consumer price levels in 2004". *Economic Trends* 615: 36–45. Available at: www.ons.gov.uk/ons/rel/elmr/economic-trends–discontinued-/no–615–february-2005/relative-regional-consumer-price-levels-in-2004.pdf (accessed April 2021).

Yesilyurt, F., and J.P. Elhorst. 2014. "A regional analysis of inflation dynamics in Turkey." *The Annals of Regional Science* 52(1): 1–17. DOI: www.doi.org/10.1007/s00168-013-0570-4.

Zimmermann, T., F. Polidoro, F. Di Leo, M. Fedeli, J. Burger, J. van den Brakel, M. Pratesi, C. Giusti, S. Marchetti, L. Biggeri, G. Bertarelli, F. Schirripa Spagnolo, T. Laureti, I. Benedetti, P.A. Smith, J. Dawber, N. Tzavidis, A. Luna, J. O'Donoghue, T. Flower, H. Thomas, N. Würz, T. Schmid, C. Articus, J.P. Burgard, C. Caratiola, H. Dieckmann, F. Ertz, J. Krause, R. Münnich, and A.-L. Wölwer. 2020. Regional poverty measurement as a prototype for modern indicator methodology: Guidelines for best practices implementation for transferring methodology. MAKSWELL. Available at: www.makswell.eu/attached_documents/output_deliverables/deliverable_3.2.pdf (accessed April 2021).

Received May 2020
Revised January 2020
Accepted May 2021