

Advisory Panel on Consumer Prices

Technical Panel

18 July 2025



Agenda

Time	Item	Presenter & Paper	Description
10:00	Introductions, apologies, and actions	Richard Heys	Make introductions if necessary. Inform Panel members of any apologies. Update on outstanding actions.
10:05 – 10:20	Update on variance estimate	Paul Smith	Update on CPI variance estimate work since April 2024
10:20 – 11:05	Retailer type stratification	Mario Spina	Further update on retailer type stratification transformation work last present at April 2025
11:05 – 11:45	New method proposal for Northern Ireland CPI	David Hearne (external) & James McGregor	Alternate methodology proposed for calculating Northern Ireland CPI by David Hearne .
11:45 – 11:55	Publication status of papers & AOB	Stephen Burgess	Publication status of papers PPI update from Stephen Burgess Next meeting date
12:00	Meeting close		

Actions from previous meeting

No.	Action	Person responsible	Progress
APCP-T (April 2025)			
19	Send a follow-up document relating to HPI imputation to the panel	Ms North and Mr Khalid	Complete
20	Send more information on implicit weights to confirm decision to stop with the panel	Mr Payne and Dr Spina	Complete
21	Send worked example of PPI correction to the panel	Mr Carey and Mr Sova	
22	Send seasonal adjustment paper to panel	Secretariat	Complete
23	Invite suggestions of future panel members	Secretariat	Complete

Update on variance estimate

Prof Paul Smith

Retailer type stratification

Mario Spina

Aim of the meeting and state of art

- Follow-up analysis after April panel
- Show big/small retailer stratification analysis based on local collection data only
- Discuss results for the market share retailer stratification
- Some results indicate this method is unsuitable for introduction in March 2026
- Question: advice on the short- and long-term improvement strategies

Retailer types

- Current:
 - Multiple/independent, stratified on the number of outlets
 - Retailers with ≥ 10 outlets are "multiple"
- Exploring
 - Big/small, stratified on the market share
 - Retailers with a market share $\geq 2\%$ are "big"

Analysis overview

Current approach

- Analysis of traditional collection data
- Analysis spanning from January 2019 to June 2024

Three main lines of analysis:

1. Weights change
2. Impact on indices and annual growth rate, and their volatility
3. Change in the sample sizes of the strata

Rethink our strategy due the annual growth rate volatility and the size of the strata results

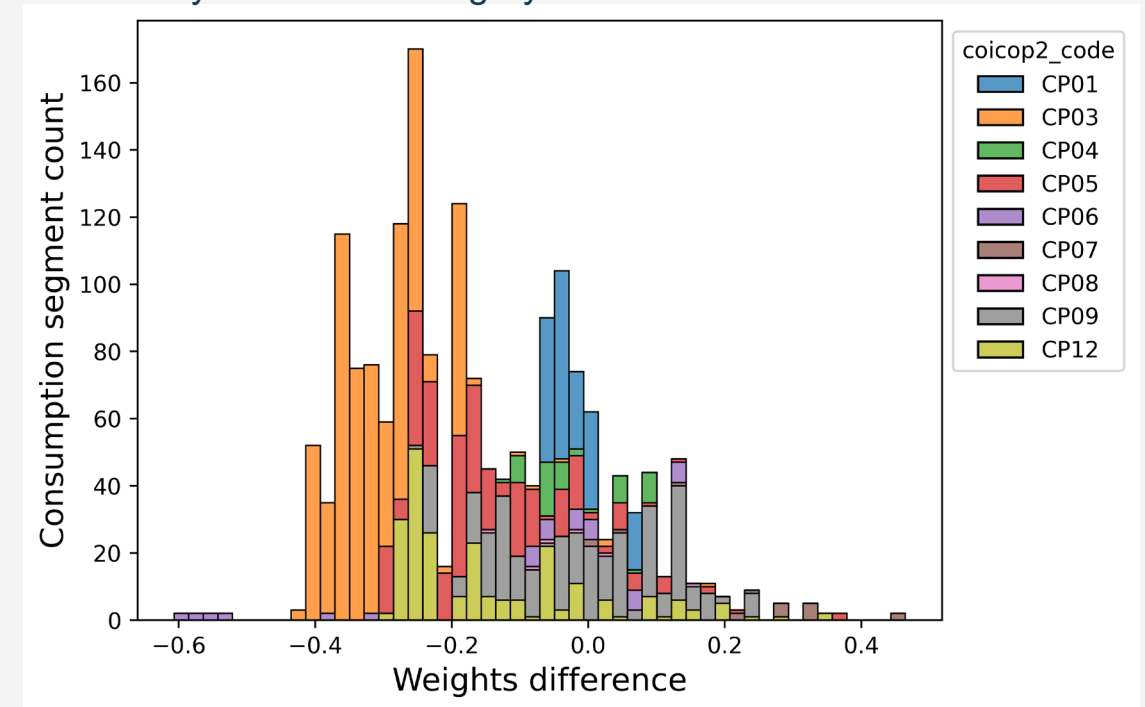
Weights change

Figure 1- Distribution of ABS retailers on a number-of-shops vs market share plane (Amended)

REDACTED

- Different classification relevant for divisions not dominated by large retailers
- Shift towards lower weights moving from "multiple" to "big"

Figure 4- Difference of 'multiple' and 'big' weights, by COICOP2 category



Impact on index and growth rate

Figure 6 - Comparison of headline CPI annual growth for both stratification methods

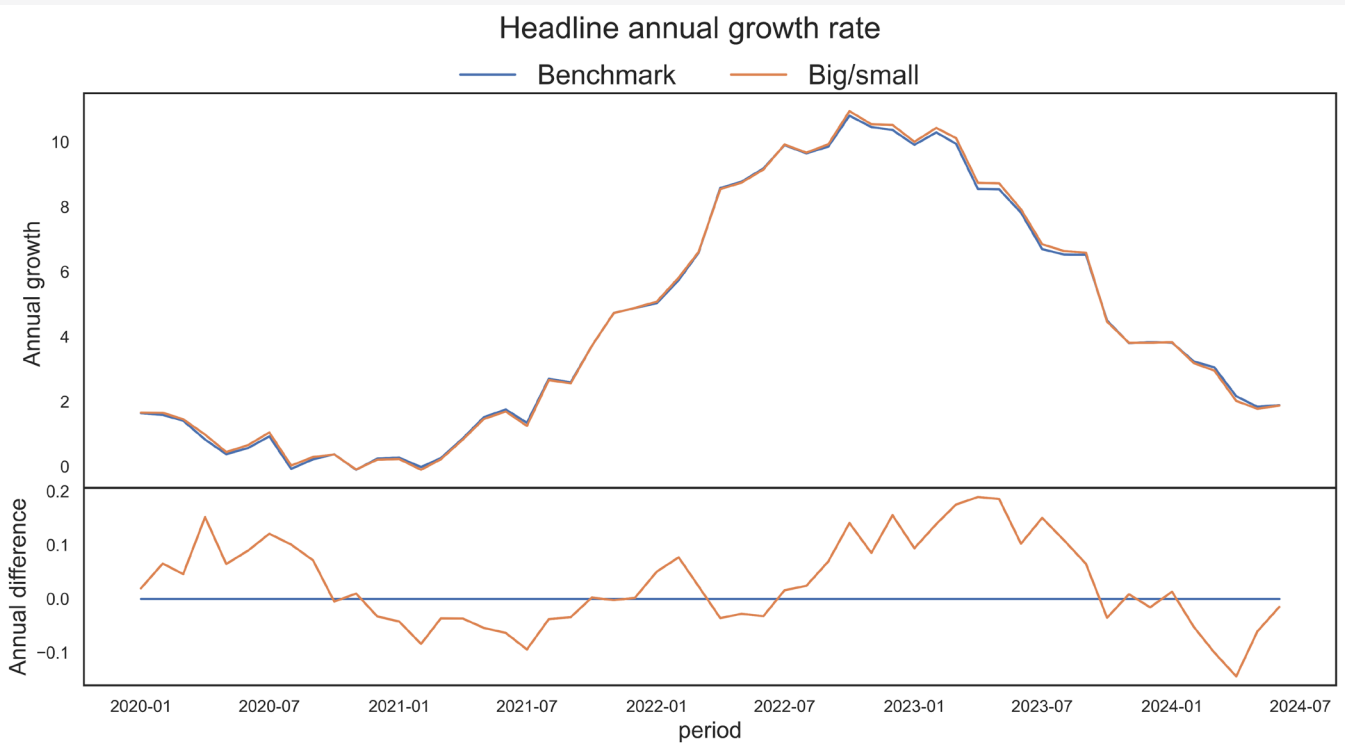


Table 1- RS3 for headline and COICOP2 CPI index

coicop2_code	coicop2_name	RS3
CP00	Headline	0.0003806
CP01	Food and non-alcoholic beverages	0.0003847
CP02	Alcoholic beverages and tobacco	0
CP03	Clothing and footwear	0.0047632
CP04	Housing, water, electricity, gas and other fuels	0.0001280
CP05	Furnishings, household equipment and maintenance	0.0024188
CP06	Health	0.0026658
CP07	Transport	0.0003896
CP08	Communications	0.0006852
CP09	Recreation and culture	0.0011721
CP10	Education	0
CP11	Restaurants and hotels	0
CP12	Miscellaneous goods and services	0.0022767

- Annual growth rate show similar trends, with maximum difference smaller than 0.2
- Change of retailer stratification doesn't seem to introduce bias to the growth rate
- RS3 scores small at headline level, and increases at lower levels of aggregation

Impact on index and growth rate

Figure 10 - Comparison of CPI annual growth for a footwear consumption segment for multiple and big retailers

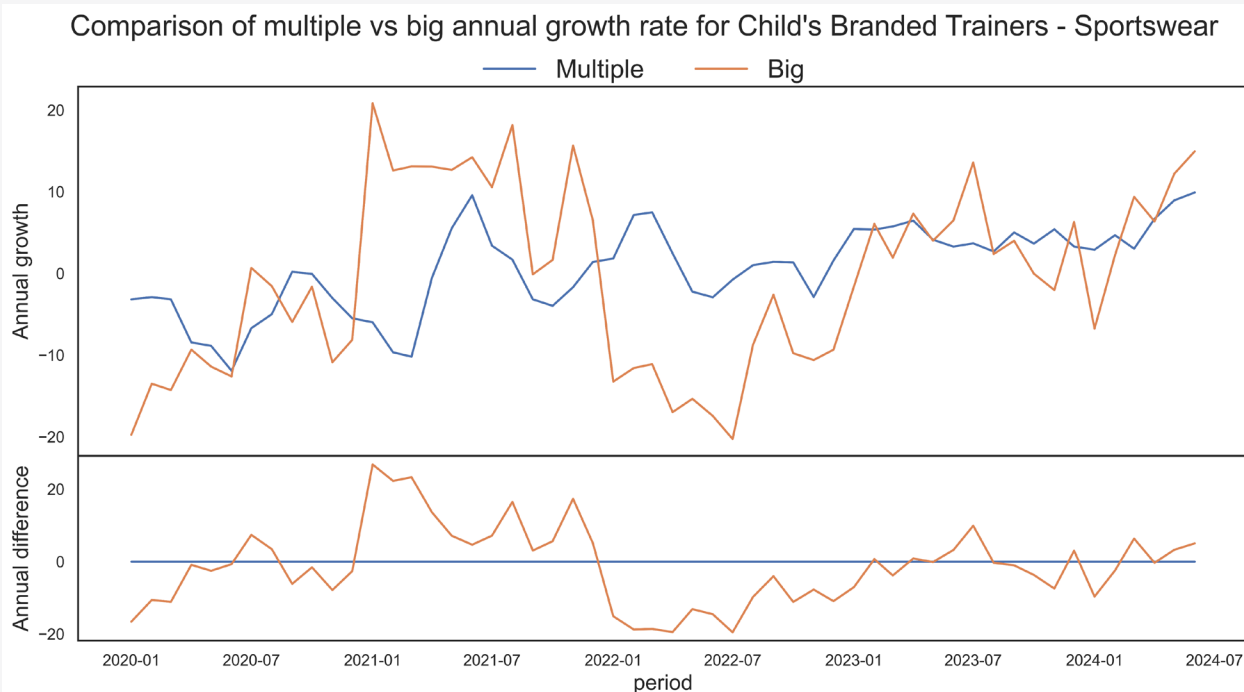
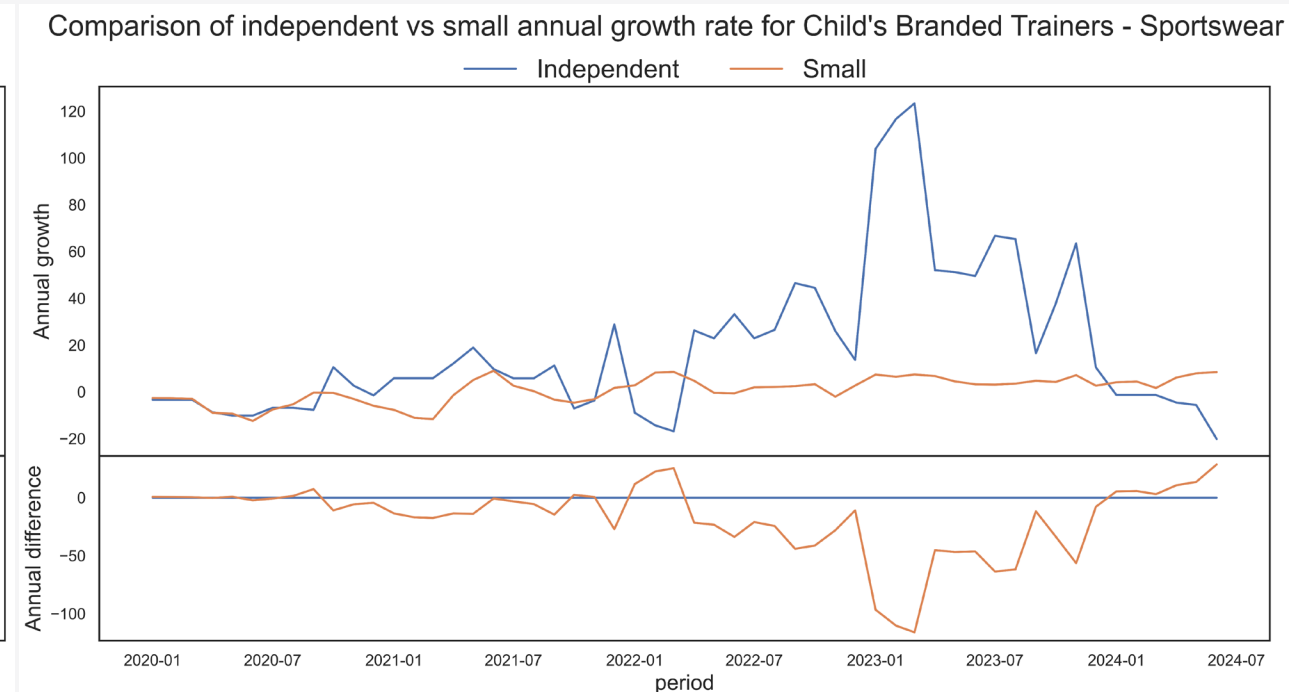


Figure 11 - Comparison of CPI annual growth for a footwear consumption segment for independent and small retailers



- Annual growth rate differences increase at lower levels of aggregation, and largest at the retailer type level
- Largest annual growth volatility by retailer type changed with the retailer stratification

Volatility and strata sizes

Figure 8- Box plot of the differences of annual growth.

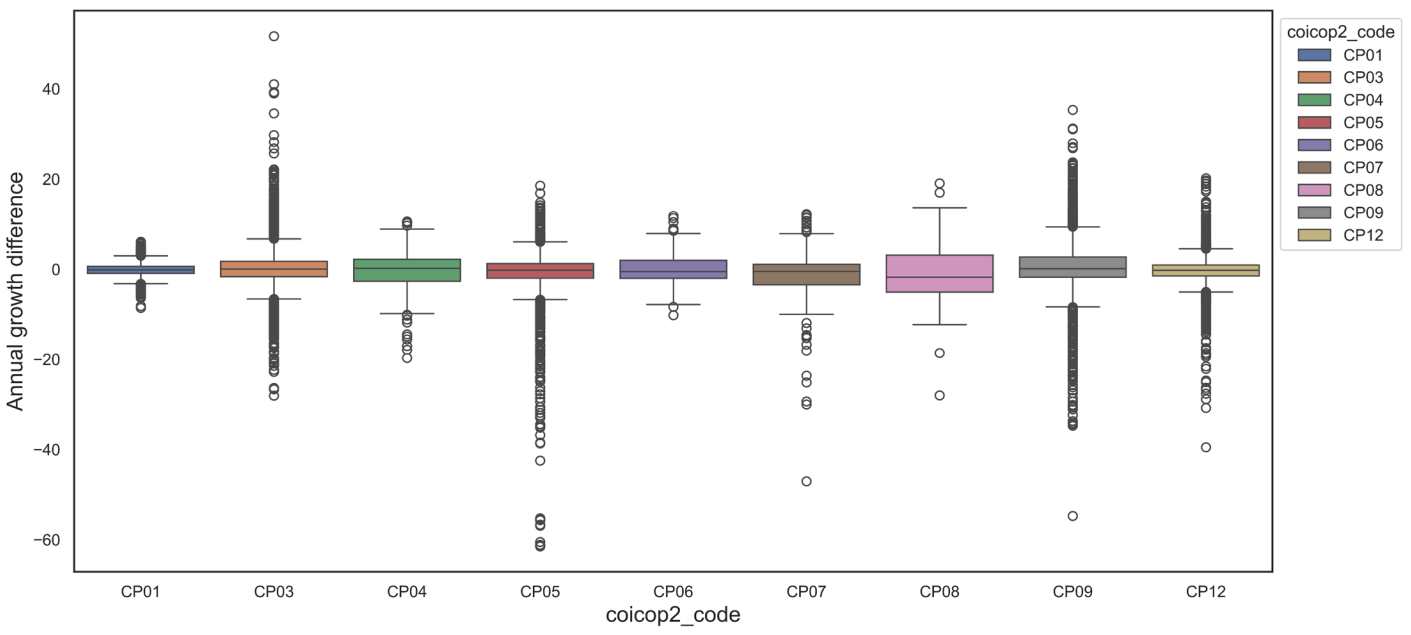


Table 3 - Comparison of the percentage of quotes stratified by retailer type for the two stratification methods

coicop2_code	Multiple retailers	Big retailers
	Sample size (%)	Sample size (%)
CP01	79.63	51.90
CP03	88.31	26.78
CP04	85.72	24.30
CP05	82.70	28.56
CP06	96.28	35.88
CP07	69.12	2.87
CP08	91.68	25.15
CP09	91.74	28.02
CP12	91.56	47.32

- Large differences and high level of volatility in annual growth rate observed at the consumption segment level
- The large change in the percentage of quotes might cause the differences observed, but it seems unreasonable.
- The extreme change in the percentage of quotes prompted us not to recommend the market share based retailer stratification for March 2026 and to rethink our retailer stratification strategy.

Proposed improvements

Improvements

- We still want to improve the multiple/independent stratification strategy.

Some avenues we want to explore include

- Short term improvements:
 - Re-assign "independent" retailers with large turnover or market share
 - Splitting retailers based on total turnover
- Long term improvements:
 - More holistic approach to the retailer stratification
 - Exploring multi-threshold market share based stratification
 - To be prioritised against other internal work

Feedback from panel

- Are our short- and long-term improvement strategies reasonable?
- Should we prioritise our work differently?
- Can you advice on any other investigation/strategy we can pursue?

Timelines

- **July** – further investigate strata size change, and explore short-term improvements
- **August** – report analysis findings and feedback to our internal decision forums
- **September** – Update on this work at the next APCP-T meeting in September

Summary

- Presented the retailer type stratification we undertook
- Performed an analysis based on local collection data only
- Discussed three main areas of investigation:
 - Change of weights
 - Impact on the index and annual growth rate
 - Change of price quote strata sample size
- Impact on volatility at lower levels and sample size made us rethink our strategy
- Presented some options for short- and long-term improvements and the next steps

Backup slides

Weights change

Figure 2 - 'Multiple' weights distribution by COICOP2 category

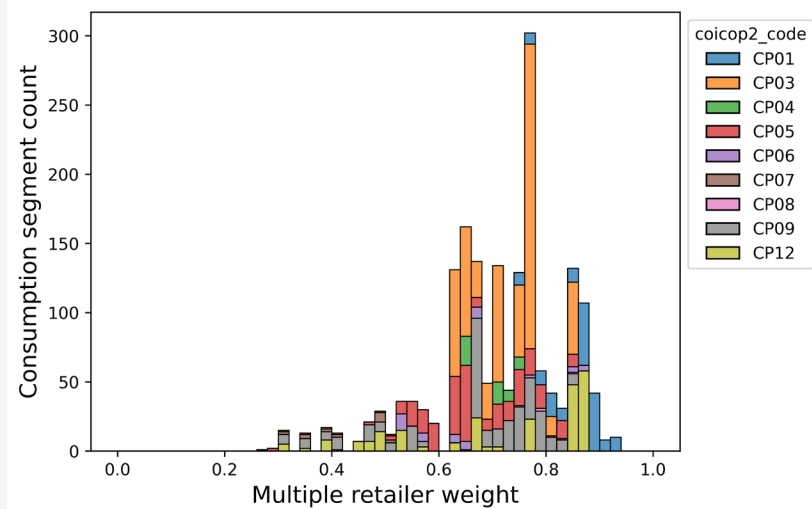


Figure 3 - 'Big' weights distribution by COICOP2 category

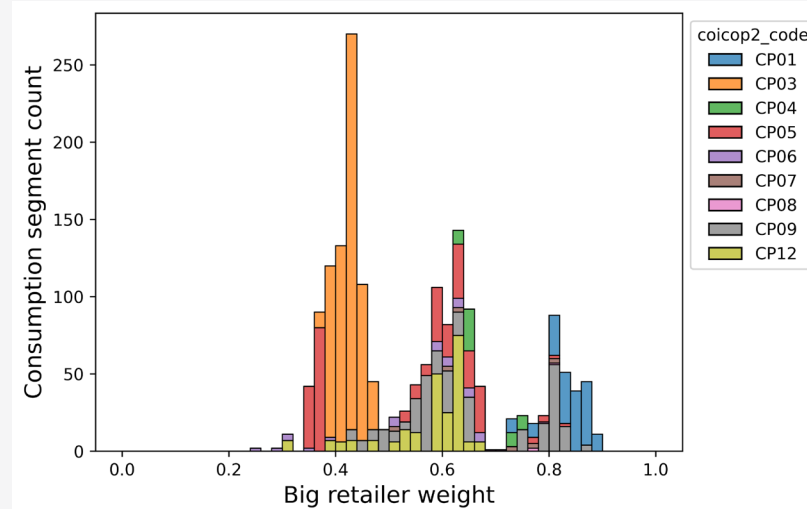
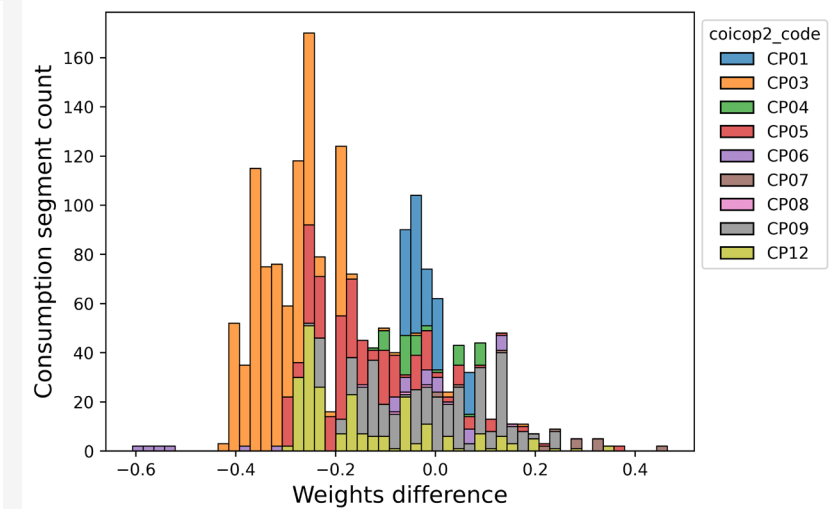


Figure 4- Difference of 'multiple' and 'big' weights, by COICOP2 category



Headline index and annual growth rate

Figure 5 - Comparison of headline CPI index for both stratification methods

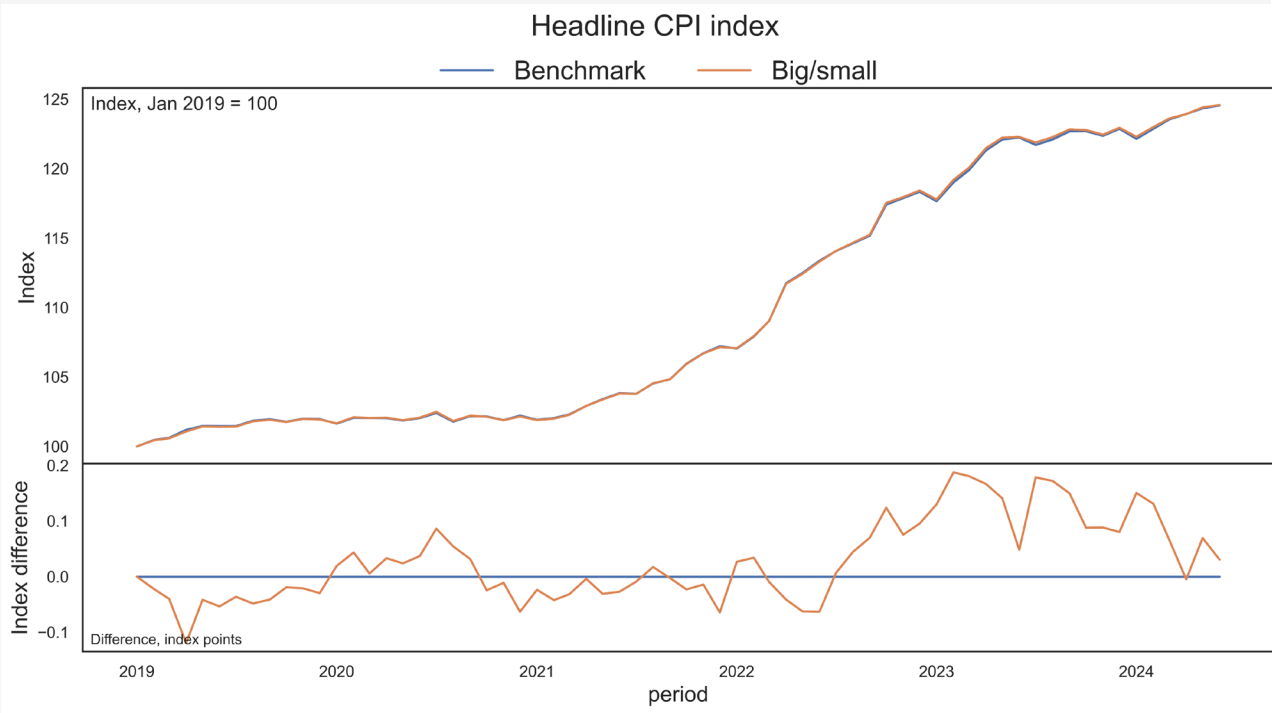
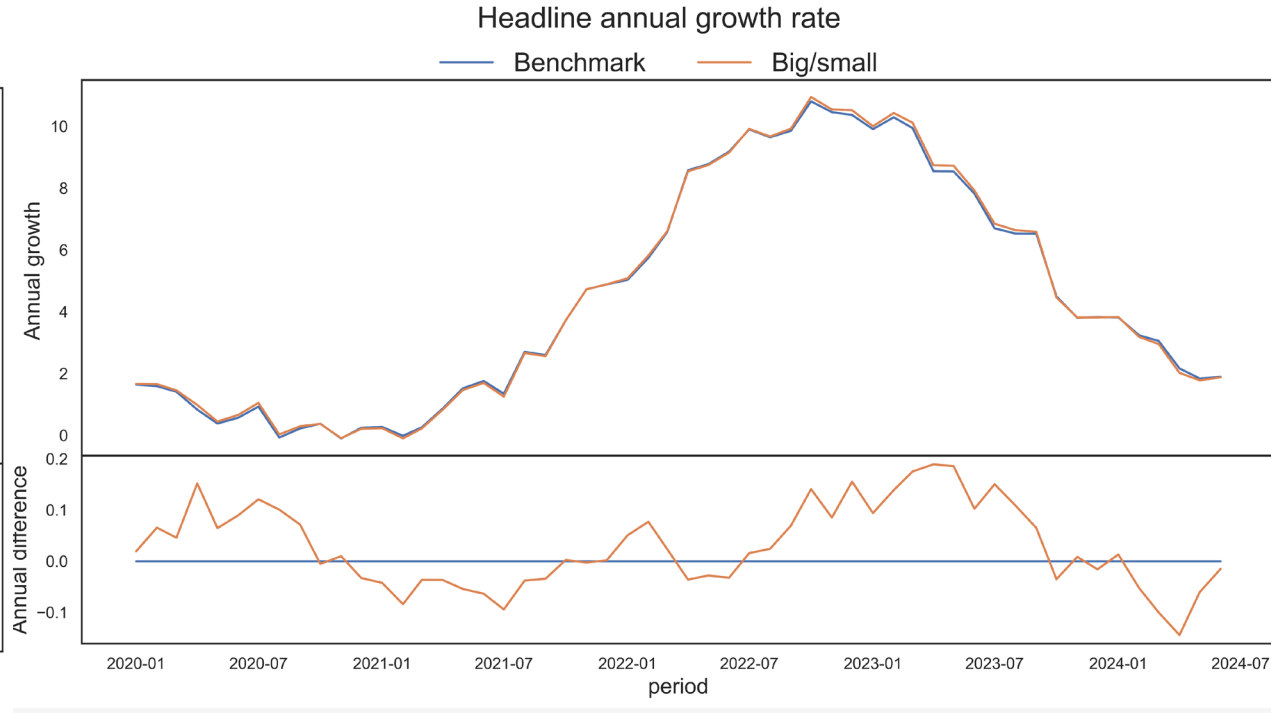


Figure 6 - Comparison of headline CPI annual growth for both stratification methods



Clothing and footwear annual growth rate and RS3

Figure 7- Comparison of Clothing and footwear CPI annual growth for both stratification methods

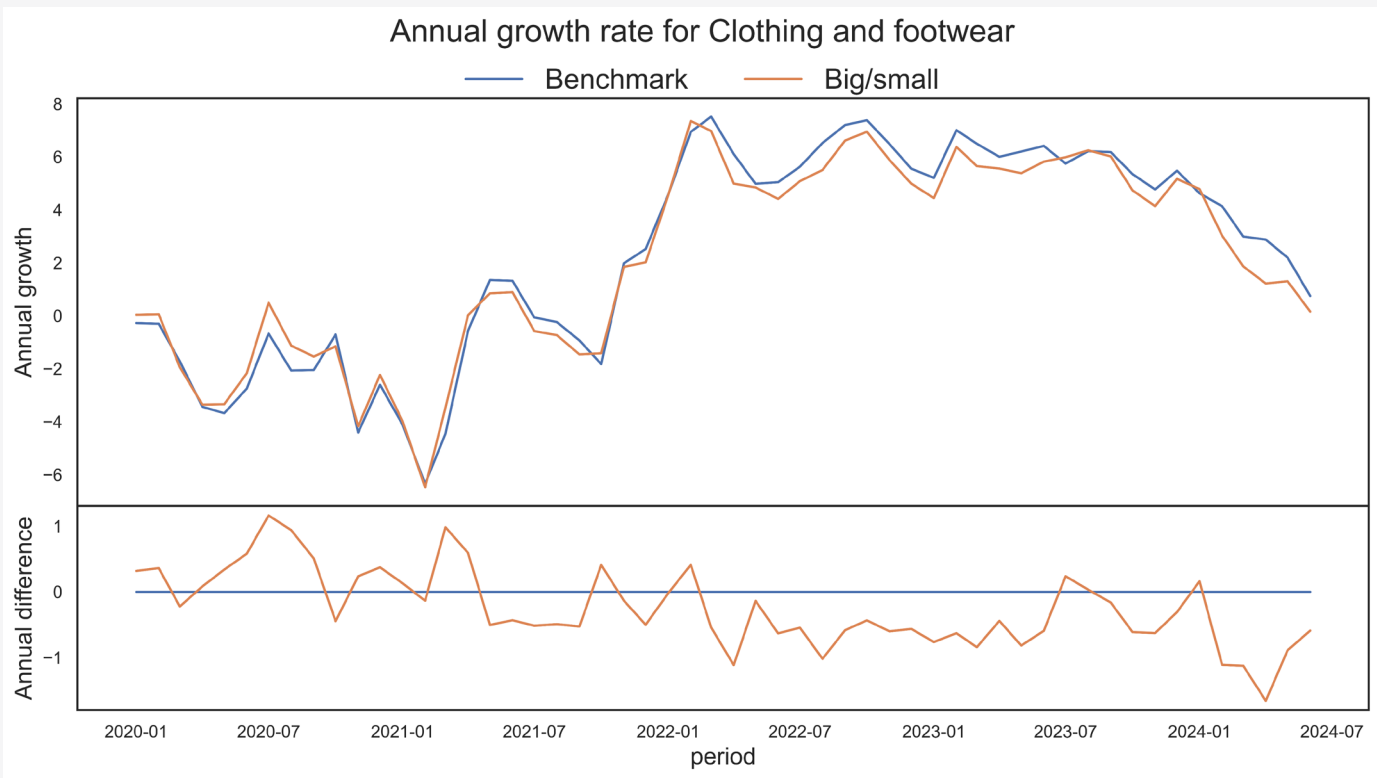


Table 2- RS3 for clothing and footwear CPI index by COICOP3 category

coicop2_code	coicop3_name	coicop4_name	RS3
CP03			0.0047632
CP03	Clothing		0.0058650
CP03	Clothing	Garments	0.0063808
CP03	Clothing	Other clothing and clothing accessories	0.0104355
CP03	Clothing	Cleaning, repair and hire of clothing	0
CP03	Footwear including repairs		0.0083703
CP03	Footwear including repairs	Shoes and other footwear	0.0083703

Annual growth rate volatility

Figure 9 - Comparison of CPI annual growth for both stratification methods for a footwear consumption segment

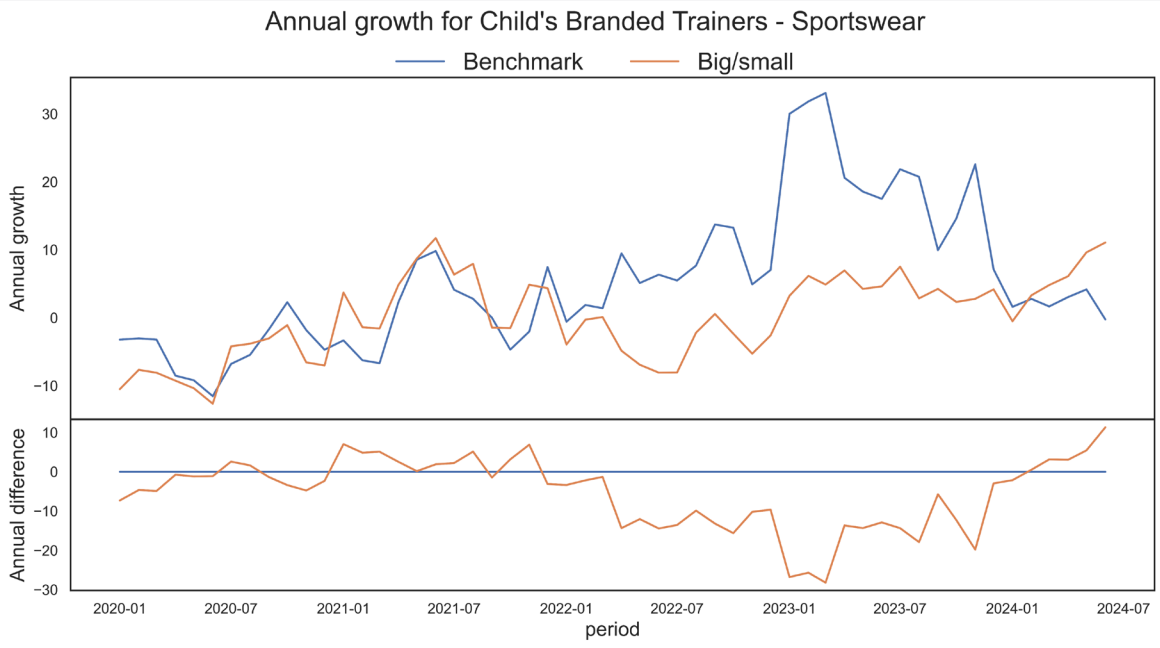


Figure 10 - Comparison of CPI annual growth for a footwear consumption segment for multiple and big retailers

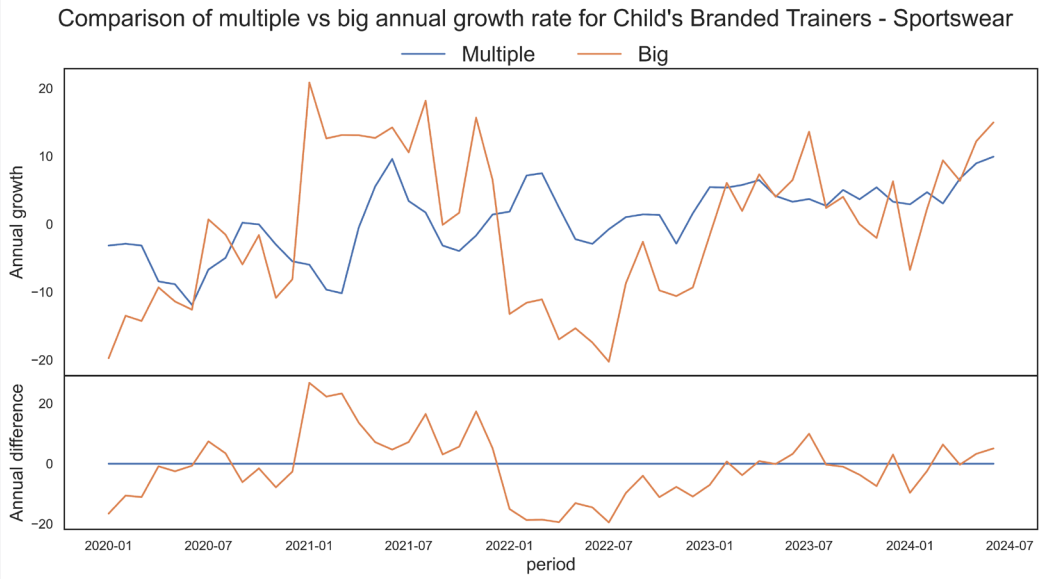
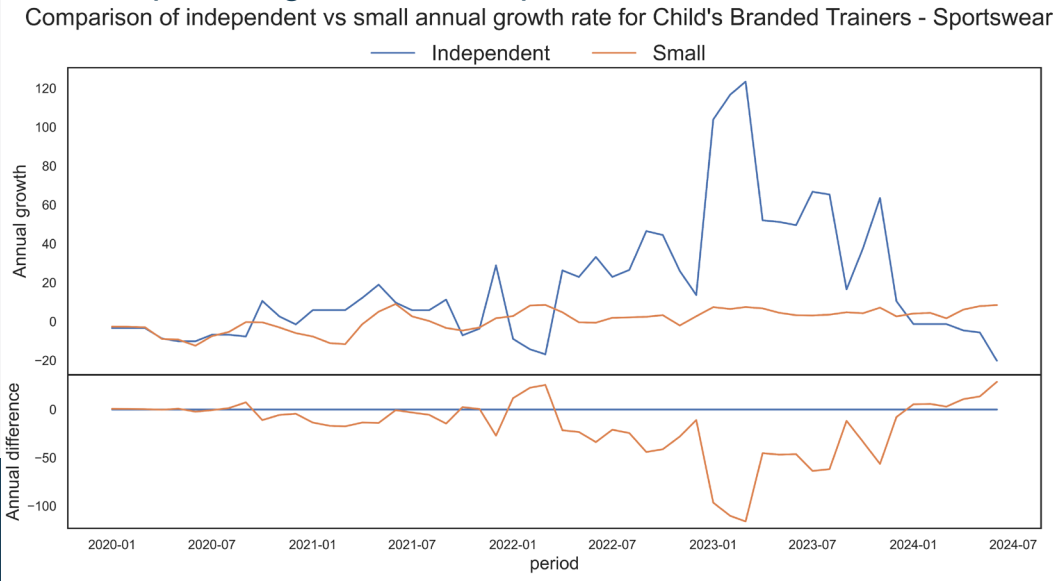


Figure 11 - Comparison of CPI annual growth for a footwear consumption segment for independent and small retailers



New method proposal for Northern Ireland CPI

David Hearne & James McGregor

Towards a regional HICP

Developing COICOP weights in a UK context

David Hearne

Birmingham Business School

18th July 2025

Subnational inflation measures

By the inflationary episode of 2022, it was clear that not all places were being impacted equally. Three recent pieces of work relate to the UK:

- ▶ Dawber, J. & Smith, P. (2017). *Feasibility study into producing CPIH consistent inflation rates for UK regions*. Office for National Statistics.
<https://www.ons.gov.uk/economy/inflationandpriceindices/methodologies/feasibilitystudyintoproducingcpihconsistentinflationratesforukregions>
- ▶ Connolly, K. & Spowage, M. (2021). *Improving the quality of regional economic indicators: Regional consumer prices*. Economic Statistics Centre of Excellence.
<https://www.escoe.ac.uk/publications/improving-the-quality-of-regional-economic-indicators-regional-consumer-prices/>
- ▶ Dawber, J., Würz, N. Smith, P., Flower, T., Thomas, H., Schmid, T. & Tzavidis, N. (2022). Experimental UK Regional Consumer Price Inflation with Model-Based Expenditure Weights. *Journal of Official Statistics*, 38(1), 213–237. <https://doi.org/doi:10.2478/jos-2022-0010>

Common problem: inflation ‘jumps’ each Jan/Feb

- ▶ Chain-linking or expenditure weights (the latter)

My background: estimating subnational PPPs in the UK (Hearne, 2021; Hearne Bailey, 2025).

Known expenditure survey challenges

The Living Costs & Food Survey

Representativeness - a national issue

- ▶ Britons overstate their consumption of fruit...
- ▶ ...and understate their consumption of confectionary!

Variability - predominantly a regional one

Large jumps in importance from one year to the next, especially in infrequently purchased expenditure classes.

We wish to estimate

$$w_{ijt} = \frac{Y_{ijt}}{\sum_I Y_{ijt}} \quad (1)$$

Where Y_{ijt} is the unknown total spending on COICOP class i in region j at time t . Eq (2) is a direct estimate of the unknown quantities from the LCFS. Households in the survey are indexed by h . In the chosen notation, H_j represents the set of households in the survey in region j .

$$\hat{w}_{ijt} = \frac{\sum_{H_j} \omega_{ht} y_{hit}}{\sum_I \left(\sum_{H_j} \omega_{ht} y_{hit} \right)} \quad (2)$$

ω is used for the household weights that exist in the survey. Hence, ω_{ht} represents the weight received by household h in year t .

Indirect expenditure weights

A compositional approach

Note we can define

$$w_{ijt} = \frac{\theta_{ijt} \sum_J Y_{ijt}}{\sum_I \left(\theta_{ijt} \sum_J Y_{ijt} \right)} = \frac{\theta_{ijt} Y_{it}}{\sum_I \theta_{ijt} Y_{it}} \quad (3)$$

Where $\sum_J Y_{ijt}$ equals Y_{it} (which we treat as known from the national total) and:

$$\theta_{ijt} = \frac{Y_{ijt}}{\sum_J Y_{ijt}} \quad (4)$$

Hence, we can define \tilde{w}_{ijt} as an *indirect* estimate of w_{ijt} , where:

$$\tilde{w}_{ijt} = \frac{\tilde{\theta}_{ijt} Y_{it}}{\sum_I \tilde{\theta}_{ijt} Y_{it}} \quad (5)$$

The problem is thus reframed as one in which we estimate $\tilde{\theta}_{ijt}$.

The estimation problem

We wish to estimate θ_{ijt} but observe $\hat{\theta}_{ijt}$, where:

$$\hat{\theta}_{ijt} = \frac{\sum_{H_j} \omega_{ht} y_{hit}}{\sum_J \sum_{H_j} \omega_{ht} y_{hit}} \quad (6)$$

Treat $\hat{\theta}_{ijt}$ as a noisy observation of underlying process. Can use Dirichlet but ALR transformation (Aitchison, 1982; 1986) more flexible.

Hence, model ALRs as:

$$\hat{\phi}_{it} = \boldsymbol{\mu}_i + \boldsymbol{\beta}' \mathbf{t} + \tilde{\phi}_{it} + \epsilon_{it} \quad (7)$$

Where the vector $\hat{\phi}_{it}$ is the ALR transform of $\hat{\theta}_{it}$. The linear predictor gives an estimate of the underlying process.

Since $P > N$, the model is unidentified without constraints. Bayesian approach works well in practice, because:

- ▶ Natural language in which to express constraints (priors).
- ▶ Gives an estimate of whole posterior (crucial for transformations!)
- ▶ Strong track-record with multilevel models in practice
- ▶ Rich software ecosystem

Each COICOP class has an overall intercept μ_i , which is left unconstrained* (via improper flat priors). Each region also has a linear time trend (to account for gradual changes in relative size, which is common across all COICOP classes and this β' is also left unconstrained (via a flat prior).

*I think empirical performance would be improved by using weakly regularising priors for μ_i .

Constraining ϵ_{it}

Conversely, the model imposes structure on ϵ_{it} and $\tilde{\phi}_{it}$, through the covariance matrices.

$$\epsilon_{it} \sim t_{\nu}(\mathbf{0}, \Sigma_i) \quad (8)$$

$$\Sigma_i = \begin{pmatrix} \sigma_{1i} & 0 & \cdots & 0 \\ 0 & \sigma_{2i} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{ji} \end{pmatrix} \quad (9)$$

Constraining ϵ_{it} cont.

Hence, the noise term is heteroscedastic across COICOP classes, *but* this is of known form, related to the *average* (across regions & years) number of non-zero observations in the LCFS (N_i).

$$\sigma_{ji} = \alpha_j \left(\frac{1}{\sqrt{N_i}} \right)^{\gamma_j} \quad (10)$$

In practice, this defines a linear relationship for $\log \sigma_{ji}$.

Constraining $\tilde{\phi}_{it}$

This term is not noise: it represents idiosyncratic variation in actual expenditure. It can be modelled as following a Gaussian distribution.

$$\tilde{\phi}_{it} \sim N(\mathbf{0}, \Psi) \quad (11)$$

Decompose Ψ into a diagonal matrix of standard deviations \mathbf{S} and a correlation matrix \mathbf{R} with the following regularising hyper-priors:

$$\varsigma_j \sim \text{half-}t_3(0, 2.5) \quad \forall j \quad (12)$$

This is (very) weakly regularising, but is sufficient for identification. An LKJ(1) prior over the correlation matrix is likewise weakly regularising.

$$\mathbf{R} \sim LKJ(1) \quad (13)$$

Theoretical performance

Performance with simulated data

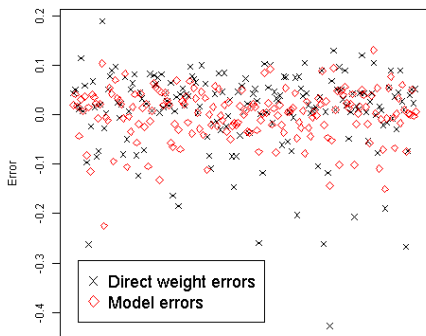
A simplified model: 12 COICOP types, 5 years 3 regions. National spending random across COICOPs.

Each COICOP is distributed across regions as follows:

$$\mu_i \sim \text{Dir} \left(\begin{array}{c} 4 \\ 10 \\ 6 \end{array} \right) \quad (14)$$

Annual variations around μ_i are tightly constrained (also via a Dirichlet). In pseudo-code:

```
theta <- apply(mu, 1, function(x)
  as.data.frame(rdirichlet(5,
    alpha = 1000*x)))
Y <- array(dim = c(5,3,12))
for(i in 1:12) {
  Y[, , i] <- as.matrix(
    theta[[i]] * national[i])
}
```



Theoretical performance

A closer look

#Z is matrix of "true" expenditure weights

```
Z <- apply(Y, 1:2, function(x) x/sum(x))
```

#Create sample of 500 households derived from Z

```
a <- array(dim = dim(Z))
```

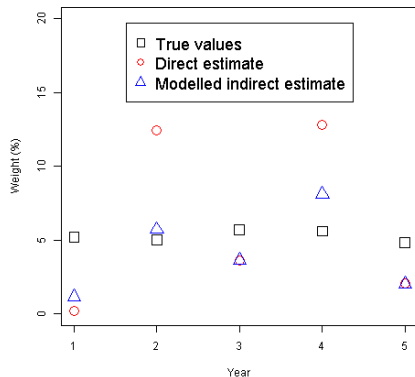
```
for(i in 1:5)
```

```
for(j in 1:3)
```

```
  a[,i,j] <- colSums(colMeans(mu)[j])* #Correct for region size
```

```
  rexp(500, 1)* #Random hh financial 'size' (rich vs. poor)
```

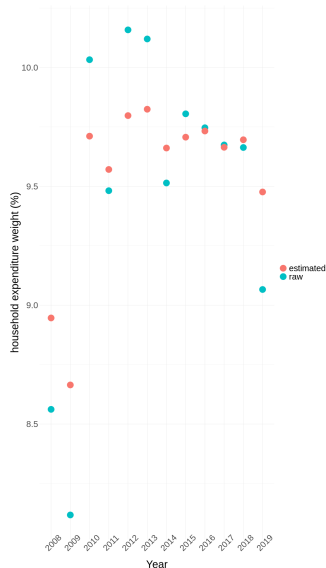
```
  t(rmultinom(500, 1, rdirichlet(n = 1, alpha = 10*Z[,i,j]))) #Random spending
```



Empirical performance

Weights in practice: classes well covered by the LCFS

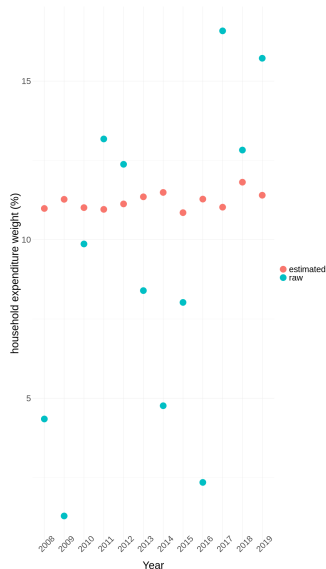
East of England's Percentage of UK
expenditure on COICOP 01.1.1



Empirical performance

Weights in practice: classes poorly covered by the LCFS

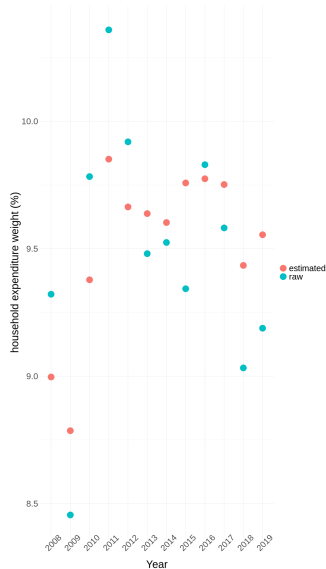
East of England's Percentage of UK
expenditure on COICOP 09.2



Empirical performance

Weights in practice: classes with intermediate coverage by the LCFS

East of England's Percentage of UK
expenditure on COICOP 08.2/3



Empirical performance

Modelled expenditure weights and inflation in East England



Green = Direct weights

Red = Modelled weights

Blue = National CPI

Further thoughts

Some unanswered questions

- ▶ There remain some 'problem' expenditure classes (e.g., airfares).
- ▶ How far should priors go in constraining μ_i ?
- ▶ What measure of central tendency should we use?
- ▶ Should cross-correlations be modelled in ϵ or is θ more sensible?
- ▶ Should we constrain γ ?
- ▶ Can these methods be integrated into existing NSI workflows in practice?

Northern Ireland CPI Weights

- Previous methodology was based around 2015 5-year average uplifted with UK changes
- Panel Previously suggested pooling multiple years of LCF data
- Now focusing on proportion of spend in each region given by the LCF
- Comparing with other data sources – Regional Accounts
- How we incorporate David Hearne's research into our approach

AOB

- Publication status of papers presented today
- PPI update
- September 5th Joint and individual panel date