Inflation and Price Dispersion: Evidence from the U.K. CPI Micro Data

Ba M. Chu, David T. Jacho-Chavez, Kim P. Huynh, Oleksiy Kryvtsov and Shane Wood

ba_chu@carleton.ca, djachocha@emory.edu, khuynh@bankofcanada.ca, okryvtsov@bankofcanada.ca, shanewood@bankofcanada.ca.

Abstract

We explore the relation between inflation and price dispersion using micro-level consumer prices from the United Kingdom's Office for National Statistics. We find that on aggregate there is a positive relationship. However, the correlation between inflation and price dispersion can be positive or negative depending on the level stratification i.e. the type of good or the region where the good is sold.
Keywords— Inflation; Price Dispersion.

1 Introduction

Understanding the dynamics of inflation is one of the most important challenges in economics. It is also the central goal for some central banks who aim to keep inflation low, stable and predictable. Detailed micro data that recently has become available for researchers provide rich information that could help them glean movements of inflation and its co-movements with other macroeconomic indicators. In this note, we focus on the relationship between inflation and price dispersion, which, according to a large literature can help sorting out competing views on the way consumer prices adjust over the business cycle, see [2]. We contribute to this literature by documenting the inflation-price dispersion relationship for the United Kingdom (UK) using the publicly available consumer priced index (CPI) survey data from the UK Office for National Statistics (ONS) and providing a decomposition of the aggregate relationship into more detailed sections.

The layout of this note is as follows: Section 2 provides information on the UK CPI microdata; Section 3 describes the methodology used to analyse the inflation and dispersion relationship; Section 4 discusses the results; and finally Section 5 concludes.

2 UK CPI Microdata

The UK ONS makes the CPI data publicly available on their online website. The data can be accessed using the following URL: `http://preview.tinyurl.com/heedmk7`. The data comes from a stratified national sample over 13 geographical regions and is classified between two types of stores: stores with more than 10 locations and those with less than 10. Each recorded price includes a descriptor on: type of item i.e. red wine-European 75 cl; 71 classification of Individual Consumption by Purpose (coicop) i.e. "meat." For the purpose of this note, we classify the goods into 12 broader categories in concordance with two-digit coicop divisions i.e. "Food and non-alcoholic beverages."

Most prices are collected monthly, except for some seasonal items. The sample period includes 235 months from February 1996 to September 2015. The ONS provides weights for each shop, stratum, coicop, and item. The sampling weights are based on the relative importance of the good to the representative household. We drop the observations deemed invalid by the ONS which results in a dataset of 22 million observations for over 1,100 individual goods and services each month. [1] provides detailed information on the dataset.

3 Methodology

To compute a measure of monthly inflation and price dispersion, we reduce our sample to only include consecutive observations. For example, if we had two consecutive observations for red wine-European 75cl in January 1999 and February 1999 the information would enter into February but not January. Focusing on consecutive price observations reduces the effect of product churning on inflation and price dispersion. Next, we calculate the weights of each item as a percentage of the total weight in the month using the weights

provided by the ONS for each item strata combination

We calculate inflation and price dispersion as defined as the interquartile range (IQR) of the log prices for all item-strata combinations:

$$\Pi_{S,t} = \sum_{i \in S} w_{i,t}(\ln(P_{i,t}) - \ln(P_{i,t-1})), \quad (1)$$

$$D_{S,t} = \sum_{i \in S} w_{i,t}[IQR((\ln(P_{i,t})) - IQR(\ln(P_{i,t-1}))], \quad (2)$$

where index i denotes individual price quote, t denotes time (month), S is the grouping of interest (e.g., aggregate, by goods, by regions); $\Pi_{S,t}$ is inflation and $D_{S,t}$ is the average percent of the interquartile range of items in the basket i.e. red wine-European 75cl is stratified by location and shop type, we calculate the IQR of red wine-European 75cl for each 24 location-shop type combinations.

These statistics are computed for the aggregate economy and then for different item groups in the aggregate economy. One, results for items that are consistently stratified by shop type that combines the results of goods stratified by shop type, and region and shop type; Two, by region which combines the results of goods stratified by region and region-shop type; Three, produce results for item groups, but only for items consistently stratified by region which is goods stratified by region, and region-shop type. Four, we produce results item groups for the regions of South East and York & Humber, using items consistently stratified by region, region-shop type, and in the regions of South East or York & Humber. We then deseasonalise all the results with monthly dummy variables and then regress each $D_{S,t}$ with the corresponding $\Pi_{S,t}$.

It is important to note that not all disaggregations of the data are created equally because the ONS stratifies its items in one of three ways: by region, by shop type, and by region-shop type, see Table 1. For example, only a weighted 30% of data can be consistently sorted into a store type, with the remaining weighted 70% of items being stratified only by region, data does exist on what type of store the other 70% is collected from, but any results generated by including the extra data may be driven around by random changes in sample composition of each store type through time due to survey methodology, not any underlying trend. 76% of data can be consistently sorted into region. This result is important for when we stratify into item groups per region. Only 4 of the 12 item groups are completely stratified at the region level (Alcoholic Beverages, Tobacco and Narcotics; Communication; Education; Food and Non-Alcoholic Beverages), and another 4 out of 12 seem to be reasonable subsamples (Housing, Water, Electricity, Gas, Fuels; Recreation & Culture; Transport; Hotels, Cafes and Restaurants). The last 4 out of 12 are not representative of their item group at all (Clothing and Footwear; Furnishings, Household Equipment and Maintenance; Health; Miscellaneous goods and services).

Table 1: Weight of Each Stratification as % of Total

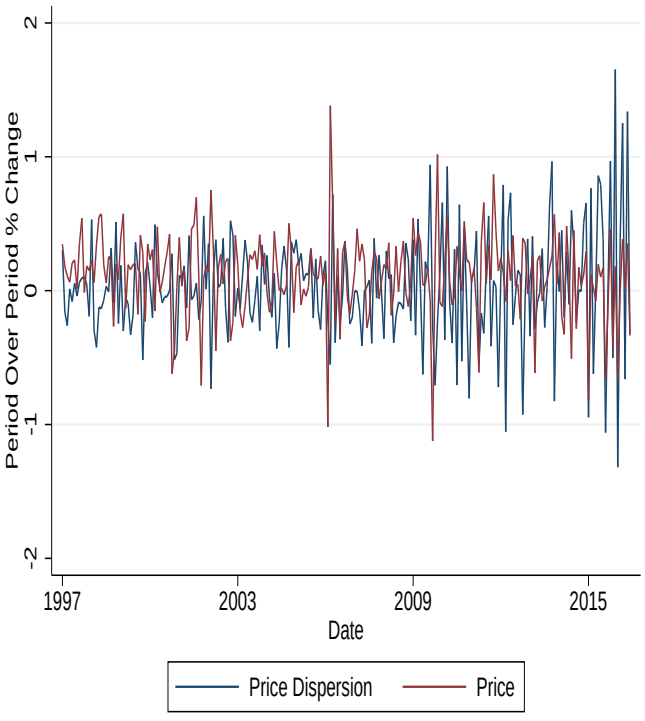
Stratification	Weight
Region	0.70
Region & Shop Type	0.06
Shop Type	0.24

Source: Author's calculations.

4 Results

Figure 1 illustrates the relationship between inflation and price dispersion at the aggregate level.

Figure 1: Monthly Chained Inflation and Price Dispersion, Total



To analyse this multivariate relationship, we estimate the following regression:

$$D_{S,t} = \alpha + \beta \Pi_{S,t} + \varepsilon_{S,t}, \quad (3)$$

where $D_{S,t}$ is the deseasonalised change in Price dispersion, α is the constant, β is the coefficient, $\Pi_{S,t}$ is the deseasonalised change in price, $\varepsilon_{S,t}$ is the error term. The regression

is done using OLS, each regression is done independently of the others.

Table 2 indicates that the overall effect of one percentage-point (ppt) increase in inflation is positively associated with a 0.15 ppt increase in dispersion. Further, we illustrate the effect of the two-digit coicop good category for all the data: Clothing and Footwear, education, Food and Non_Alcoholic Beverages; Housing, Water and Electricity and gas and fuels; Miscellaneous goods and Services, and Transport. There are significant negative relationship between inflation and change in price dispersion, with Education being the stand-out with a 1ppt increase in inflation leading to a 1.2ppt decrease in price dispersion, the rest have on average a 0.25ppt decrease. Communication, and Furnishing Household Equipment and Maintenance have a positive relationship between inflation and price increases, with an effect of 0.64ppt and 0.76ppt respectively.

Table 2: Regression of Dispersion on Inflation, By Item Category (Full Sample)

	Coef.	S.E.	Weight
Total	0.145*	0.087	1
	Coef.	S.E.	Weight
Alcoholic Beverages, Tobacco and Narcotics	-0.018	0.054	0.067
Clothing and Footwear	-0.299***	0.063	0.087
Communication	0.640***	0.140	0.032
Education	-1.178***	0.366	0.018
Food and Non-Alcoholic Beverages	-0.277***	0.052	0.157
Furnishings, Household Equipment and Maintenance	0.762***	0.087	0.086
Health	-0.131	0.145	0.026
Hotels, Cafes and Restaurants	0.164	0.127	0.161
Housing, water, Electricity, Gas, Fuels	-0.164**	0.068	0.078
Miscellaneous goods and services	-0.254**	0.140	0.0742
Recreation & Culture	0.088	0.115	0.116
Transport	-0.254**	0.053	0.110

Notes: Coef. and S.E. denote the coefficients and standard errors, respectively. Asterisks represent the marginal significance level of: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Weights are normalized to add up to 1 within grouping.

Table 3 considers the same results but disregarding items that are only stratified by shop type, using the same data 76% as the regions, there are a couple of changes. Clothing and footwear switches from significantly negative to positive. Health becomes significantly positive. Furnishing Household Equipment and Maintenance, and Miscellaneous goods and Services become insignificant. This result is not surprising since these categories were overly representative in the 24% of observations that were dropped for this run, making up 72% of those observations.

Table 3: Regression of Dispersion on Inflation, By Item Category, Excluding Items Only Stratified By Shop

	Coef.	S.E.	Weight
Alcoholic Beverages, Tobacco and Narcotics	-0.018	0.054	0.087
Clothing and Footwear	0.275*	0.152	0.004
Communication	0.640***	0.140	0.042
Education	-1.178***	0.366	0.024
Food and Non-Alcoholic Beverages	-0.277***	0.052	0.206
Furnishings, Household Equipment and Maintenance	0.012	0.102	0.041
Health	1.003***	0.095	0.023
Hotels, Cafes and Restaurants	0.127	0.161	
0.164			
Housing, water, Electricity, Gas, Fuels	-0.219***	0.060	0.085
Miscellaneous goods and services	0.058	0.081	0.052
Recreation & Culture	0.095	0.115	
Transport	–0.240***	0.052	0.148

Notes: Coef. and S.E. denote the coefficients and standard errors, respectively. Asterisks represent the marginal significance level of: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Weights are normalized to add up to 1 within grouping.

Table 4 illustrates the effect of shop type. Large stores (more than 10 outlets) appear to have a negative effect on price dispersion of a 1 ppt increase in inflation leading to a 0.17 ppt decrease in dispersion. The negative relationship is due to items that are well stratified by shop type; in this case Health, Clothing and footwear, and Miscellaneous Goods and Services have a strong negative relationship between inflation and price dispersion, together with this sample only being representative of 30% of items, so these items have a disproportionately high weight compared to the aggregate.

Table 4 illustrates the effects of region South East, York & Humber, and North West all have a significant positive coefficient of around 0.32, whereas East Anglia has a negative relationship of -0.22. The data used accounts for 0.76% of

the weight.

Table 4: Regression of Dispersion on Inflation, Using Strata Provided

	Shop Type		
	Coef.	S.E.	Weight
Chain (10 or more outlets)	-0.172*	0.097	0.650
Independent(less than 10)	-0.090	.0080	0.350
	Region		
	Coef.	S.E.	Weight
London	-0.075	0.119	0.121
South East	0.342**	0.150	0.158
South West	0.015	0.155	0.087
East Anglia	-0.216**	0.107	0.083
East Midlands	0.050	0.107	0.071
West Midlands	0.009	0.134	0.085
York & Humber	0.300*	0.156	0.084
North West	0.338***	0.130	0.102
North	0.134	0.118	0.049
Wales	0.165	0.131	0.050
Scotland	-0.128	0.131	0.082
Northern Ireland	-0.017	0.112	0.028

Notes: Coef. and S.E. denote the coefficients and standard errors, respectively. Asterisks represent the marginal significance level of: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Weights are normalized to add up to 1 within grouping.

Finally, Table 5 considers the underlying factors in the positive response in York and Humber compared to the negative in East Anglia, we will only discuss good categories whose relationship is consistent with the full sample. In East Anglia, Communication is the only significant positive relationship, with a coefficient of 0.77. In contrast, Education, Food, and transport all have significant negative coefficients of on average -0.46. In addition while communication makes up 5% of the weight in East Anglia, the negative coefficients make up 38% of the weight. In York and Humber, only food has a significant negative coefficients of -0.36. Hotels, Transport, and, surprisingly, Education all have significant positive coefficients, and make up more of the basket as well (33% to 20%).

Table 5: Regression of Dispersion on Inflation, By Item Category, Excluding Items Only Stratified By Shop

East Anglia	Coef.	S.E.	Weight
Alcoholic Beverages, Tobacco and Narcotics	-0.035	0.093	0.070
Clothing and Footwear	0.742***	0.182	0.004
Communication	0.772***	0.193	0.047
Education	-0.619*	0.359	0.013
Food and Non-Alcoholic Beverages	-0.436***	0.089	0.201
Furnishings, Household Equipment and Maintenance	0.103	0.134	0.044
Health	2.802***	0.211	0.030
Hotels, Cafes and Restaurants	-0.069	0.153	0.200
Housing, water, Electricity, Gas, Fuels	0.036	0.108	0.078
Miscellaneous goods and services	0.748***	0.092	0.061
Recreation & Culture	-0.088	0.116	0.085
Transport	-0.344***	0.102	0.166
York & Humber	Coef.	S.E.	Weight
Alcoholic Beverages, Tobacco and Narcotics	-0.081	0.085	0.093
Clothing and Footwear	-0.350**	0.153	0.004
Communication	0.029	0.206	0.045
Education	0.615***	0.217	0.023
Food and Non-Alcoholic Beverages	-0.361***	0.087	0.200
Furnishings, Household Equipment and Maintenance	-0.098	0.117	0.041
Health	1.227***	0.187	0.021
Hotels, Cafes and Restaurants	0.217*	0.129	0.232
Housing, water, Electricity, Gas, Fuels	-0.247	0.151	0.074
Miscellaneous goods and services	-1.114***	0.085	0.049
Recreation & Culture	0.408***	0.145	0.076
Transport	0.181	0.137	0.143

Notes: Coef. and S.E. denote the coefficients and standard errors, respectively. Asterisks represent the marginal significance level of: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Weights are normalized to add up to 1 within grouping.

5 Conclusion

We find that the correlation between price dispersion and inflation depends on stratification at the region-level. Food is the only good category with a consistent negative relationship between dispersion and inflation. Communication always seems to have a positive although sometimes insignifi-

cant relationship. For other goods there seems to be regional factors at play, with different goods behaving differently in different regions. Even if aggregate price dispersion is correlated with inflation then it does not mean it holds true for all goods. Some will be positively correlated, some negative, just more will be positive than negative. So, treating a price increase as harmoniously increasing dispersion in prices is potentially incorrect.

A possible explanation why a negative relationship may exist is that it maybe associated with unanticipated inflationary events which tend to be negatively associated with Dispersion. Whereas the goods with a positive relationship are more likely to be associated with anticipated inflation, in other words more people are more aware of what is going to happen to the price of food than phones, residents of different regions may have different knowledge of goods so what is a surprise in one region may be a known in another, etc. However, testing these hypotheses is outside the scope of this paper. A fruitful avenue for further investigation is to consider the evolution of price distributions using functional principal components analysis that allows for survey weighting, see [3] for more details.

Acknowledgements

We are grateful to the UK Office for National Statistics, especially, Derek Bird and Ainslie Restieaux for their assistance with this project; Rosie Foster and Rowan Kelsoe for their detailed responses to our queries regarding the data. We would like to acknowledge Boyan Bejanov for his expert technical assistance for this project. All computations were completed using EDITH or the Bank of Canda's high performance cluster. Tony Chernis, Geoff Halmo, and Anderson Nzabandora provided excellent research assistance. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada. All remaining errors are the responsibility of the authors.

References

[1] ONS. "Consumer Price Indices - Technical Manual," The Office of National Statistics, Prices Division, Newport, United Kingdom, 2014.

[2] David Fielding, Christopher Hajzler, and James MacGee. "Price Leve Dispersion versus Inlation Rate Dispersion: Evidence from Three Countries," Bank of Canada Staff Working Paper, 2016.

[3] Ba M. Chu, Kim P. Huynh, and David Jacho-Chávez, and Oleksiy Kryvtsov. "On the Evolution of the United Kingdom Price Distributions," mimeo, February 2016.