Price Rigidity in the Italian Grocery Retailing Sector: Evidence from Scanner Data

Santioni, Raffaele (1st author)
Bank of Italy, Economic Research Unit, Branch of Rome
Via XX Settembre 97/e
Rome (00100), Italy
E-mail: raffaele.santioni@bancaditalia.it

Carbonaro, Isabella (2nd author)
Tor Vergata University of Rome, Department
Street Address
Rome (00100), Italy
E-mail: isabella.carbonaro@uniroma2.it

The views expressed in this paper are those of the authors and do not necessarily reflect those of their organizations.

This paper provides an empirical evidence on the relevance of price rigidity and on the role of some major determinants of food price structure in Italy. Recent empirical research has made extensive use of monthly data to analyze price stickiness and to control for the impact of temporary sales. The analysis is based on a comprehensive weekly scanner data set collected at regional level from large grocery-retailers for different product groups (pasta, oil olive and other mixed seeds oil, tomato pulp, bread, cheese, *etc.*). These products represent an important share of italian households food consumptions sold by large and small-medium size manufacturers.

1 Introduction

This paper makes an empirical contribution to the literature by estimating price stickiness in Italy using high frequency scanner data collected by the AC Nielsen in the Grocery Retailing Sector. Indeed, besides CPI data which are gathered by government officials, there are scanner data[1], another important source of micro data that offers a huge quantity of pricing quotes over time and space and for a wide range of products. Scanner data contain all price quotes of items that are actually being bought by consumers. Although CPI data are in general much broader in terms of products coverage than scanner data from any given outlet, the latter have an important asset in that they are collected on weekly basis, whereas CPI data are gathered monthly, at best.

Following the approach proposed by Bils and Klenow (2004) and Klenow and Kryvtsov (2008), our aim is to use higher frequency data to revisit the results on price stickiness and to document the main characteristics of consumer price changes, focusing on the frequency of price changes for different products. We also treat the issue of the relevance of downward and upward price rigidities. At the same time we compare the frequency of price changes that we obtain at base frequency with the ones from data set at monthly frequency, artificially derived from the original data. In this way, we are able to gange the effects of short-term pricing tactics on the price stickiness.

The rest of the paper is structured as follows. In section 2 we briefly describe the scanner data set. Section 3 defines the main methodological steps and statistics on which our empirical analysis is based. Section 4 presents the empirical results at weekly and monthly frequency. Section 5 concludes.

[1] They are obtained through the scanning process of product-specific bar codes that is common use in supermarkets and hypermarkets.

2 The scanner data set

In Italy, AC Nielsen has developed a weekly data set, aggregated at regional level, on retail prices, quantities and product characteristics for over 130.000 products sold by major supermarket (400 to 2,499 m^2 area) and hypermarket (at least 2,500 m^2 area) chains.[2] It captures scanner data from the Italian Large-Scale Retailing Sector over 17 regions - all Italian regions with the exception of Basilicata, Valle d'Aosta and Molise - with a time span of 217 weeks, i.e. the period from June 2002 to July 2006.

Product groups included are those typically sold in supermarkets (food, beverages and a large number of other commodities for personal care, home cleaning, *etc.*). By "product" we denote each of the items included in our sample. Take, as example, "Milk", several price quotes are available, each referring to a specific brand, variety and package sold. Instead, by "individual product", we mean a specific product, of a specific brand and quality. Each individual product has an elementary price quote at particular time. In terms of our example, by elementary price quote we mean the price of "1 liter package of milk, of brand X, sold at time t in region j". Individual product are identified by a unique bar code (European Article Number[3]) which allows us to follow the elementary price quote through the time and space. The sequence of records corresponding to one individual product is referred as a "price trajectory".

In this paper we concentrate on the sub-groups pasta, oil olive and other mixed seeds oil, tomato pulp, bread, cheese, and so on. For each item at temporal and regional level we have weekly revenues, quantity and number of packages sold. Prices are not included: weekly and monthly average revenue per unit sold has been used as a proxy. Information on product characteristics (brand name, weight and variety of product) is also provided.

3 Some methological issues

Price rigidity can be measured by the duration of a price spell in months following two alternative methodologies: the duration approach measures the average durations of price spells directly from the data, whereas the frequency approach aims at calculating the frequency of price change which can then be inverted to obtain the implied duration of a price spell[4].

In this work, we employ the latter approach because it offers several practical advantages. First of all, gaps in price trajectories are a serious concern in the duration approach, as it is impossible to derive the spell length if price quotes are missing[5] whereas in the frequency approach, a long and uninterrupted span of time series is not a necessary condition to calculate the frequency of price change[6], as long as homogeneity and stationarity are valid assumptions. Secondly, the duration approach suffers from serious censoring issues that complicate the measurement of price spell durations, whereas frequency approach does not require an explicit treatment of censoring

In order to figure out the duration of a price spell using the frequency approach, we need to invert the estimates of the frequency of price change. In this respect, it is important to note that we use data in their base frequency. As a consequence, we can calculate the duration of a price spell

[2] Although the data are collected from a sample of supermarket and hypermarket chains, the coverage within the overall sector is very high. Indeed, products sold by large retailers represent an important share of personal consumptions (in Italy, 77% of food and 88% of beverage and grocery expenditures according to AC Nielsen).

[3] EAN code contains 8 or 13 numbers which are printed on the product and it allows the identification of the manufacturer and the details of the product.

[4] See Baudry *et al.* (2004) and Veronese *et al.* (2005), among others, for an analysis of both approaches.

[5] This problem can be tackle by carry forward the last available price to fill in gaps, but this can be generate an upward bias for the duration of price spells (Baudry *et al.*, 2004; Klenow and Kryvtsov, 2008).

[6] This is a very important asset in our setting, as scanner data generally contain a reasonable amount of gaps in price series, due for example to stockouts.

simply by inverting the frequency parameter, call it F_{ij}. The implied duration of a price spell therefore equals to $T_{ij} = \frac{1}{F_{ij}}$, expressed in time units corresponding to the intervals at which price quotes are available. Working with data that are not in base frequency complicates the picture, because if prices can change at any point in between data points, then the instantaneous probability of a price change is $-ln(1 - F_{ij})$ and the implied mean time between price changes is equal to $T_{ij} = \frac{-1}{ln(1-F_{ij})}$ (Bils and Klenow, 2004). This is the way frequencies and durations are linked when using CPI data, and it is the formula that we will use when we work with our derived monthly scanner data.

We have to take different relevant aspects into account in the implementation of frequency approach. Firstly, for the homogeneity and stationarity assumptions to be fulfilled, is important to perform the analysis both at a disaggregated level and for homogenous sub-periods. Secondly, aggregation method raises some issues. In case of uncensored data, one indicator of the overall duration can be measured as the inverse of the weighted average frequency of price changes. However, acknowledging that homogeneity is likely to be satisfied at the product level only, a more relevant approach is to compute the average duration[7] for each product level i and then averaging over all products using specific weights. Thirdly there are a number of statistics that can be used to capture the implied duration of a price spell. Among others, we use also the weighted median of inverse frequencies (Bils and Klenow, 2004). This measure is particularly remarkable given that for some sector the frequency is close to zero, leading to a very large value of implied duration which strongly influences the overall mean. However, this statistic cannot be interpreted as an estimator for the average duration of prices.

4 Results

The results in table 1 show implied mean duration of price spell, breakdown by Italian regions, measured both in weekly and monthly frequencies. Compared to previous researchs on the extent of price stickiness in Italy, our estimates of duration are very low at base frequency, if we do not admit a threshold for price changes (i.e. if we set s equal to zero)[8]. The most important reason to explain the high frequency of price change is that we study retail prices (some very sticky classes of consumer goods and services, as apartment rents or restaurant meals, are not included in our sample). Another factor is the type of retailer, as large supermarkets and hypermarkets like the ones we analyze change their prices more swiftly compared to traditional retailers. Furthermore, this is likely due to that elementary data available to us are aggregated across outlets[9]. Indeed, if we introduce small price variation threshold in the price stickiness measurement, to overcome temporal price variability related to this issue, we obtain different pictures. With a threshold of 1 percent the price stickiness growths on average to one month. Allowing a more restrictive price variation boundary (2 percent), alike the European inflation target, the prices are stickier: on average for one month an half, with a mean implied duration of about 2 months for some regions such as Lazio and Liguria.

In table 1, we also present the results of a similar analysis based on monthly price series that we directly derived from high frequency scanner data by withholding only the first or the last price observation of each month. As such, we mimic the structure of typical monthly data sets like those used in empirical work to analyze price rigidity. On top of that, prices in general appear to be more sticky when monthly data are considered. This it true also if we do not introduce a price variation threshold over time. In this case the regional mean implied duration ranges from 1.1 to 1.3 months. This is a direct consequence of the fact that temporary sales are very short-lived and hence not picked

[7]Appendix shows how we compute the mean frequency of price change for a specific product i across t time periods.

[8]Indeed, the price rigidity measure is about one third a month in each region.

[9]It is important to remark that for each individual product trajectory, we do not observe price quote at store level, but unit value at regional level.

up in monthly data. It is therefore an inherent problem of price data that are not available in their base frequency, as they do not provide direct evidence about the critical issue of how many temporary price changes happen in between data points (Kehoe and Midrigan, 2010). Also in this context, allowing for a price variability boundary, we observe an increasing price stickiness: with s fixed at 1 or 2 percent the mean duration varies respectively from 2.2 and 2.7 months for Sicily to 3.5 and 5.0 months for Liguria.

Table 1: Mean implied duration of a price spell in months (weighted statistics)

	Product level								
	weekly base			monthly base (1)			monthly base (2)		
	(3)	(4)	(5)	(3)	(4)	(5)	(3)	(4)	(5)
Piedmont	0.3	0.9	1.5	1.2	2.7	3.6	1.2	2.6	3.5
Lombardy	0.3	0.9	1.3	1.2	2.5	3.4	1.2	2.5	3.4
Liguria	0.3	1.3	2.3	1.2	3.7	5.4	1.2	3.5	5.0
Trentino	0.3	1.2	1.7	1.2	3.4	4.2	1.2	3.2	4.0
Veneto	0.3	0.9	1.3	1.2	2.6	3.4	1.2	2.6	3.5
Friuli-Venezia	0.3	1.0	1.5	1.3	3.0	3.9	1.3	2.7	3.6
Emilia Romagna	0.3	1.0	1.5	1.2	2.7	3.6	1.2	2.6	3.5
Tuscany	0.3	1.0	1.5	1.2	2.7	3.5	1.2	2.7	3.5
Umbria	0.3	0.9	1.2	1.2	2.8	3.4	1.2	2.7	3.3
Marche	0.3	0.9	1.3	1.3	2.7	3.6	1.3	2.6	3.6
Lazio	0.3	1.1	1.9	1.2	3.0	4.5	1.2	3.0	4.1
Abruzzo	0.3	0.9	1.2	1.1	2.6	3.3	1.1	2.6	3.2
Campania	0.3	0.9	1.4	1.2	2.6	3.4	1.2	2.6	3.4
Apulia	0.3	1.0	1.6	1.2	2.7	3.5	1.2	2.8	3.8
Calabria	0.3	1.0	1.4	1.3	3.0	3.6	1.3	2.9	3.6
Sardinia	0.3	0.9	1.2	1.2	2.7	3.3	1.2	2.6	3.2
Sicily	0.3	0.7	0.9	1.2	2.2	2.7	1.2	2.2	2.7

Source: AC Nielsen ScanTrack.
Note: (1) First price observation of each month. (2) Last price observation of each month.
Threshold s fixed at (3) 0 percent. (4) 1 percent. (5) 2 percent.

Looking at the mean implied duration of a price spell does not give any insight into the composition of the price changes. Therefore, we break up all price movements into increases and decreases. In table 2 we report the mean frequency and the size[10] of both price variations obtained with a boundary s fixed at 1 percent at weekly basis. The estimates show that price increases are slightly more frequent than decreases as long as the order of price increases are in general greater than decreases. Furthermore, both price movements are sizable if we compare them with the low inflation rates that are prevalent in Euro area during this period. Focusing the analysis on the median size of a price increase, we find some regional differences: it spans from about 8 percent for many smaller regions in the Central and Northern area to about 7 percent in the South. The same happens for the median size of a price decrease. Secondly, we notice that the median size of a price change is closer to the mean, implying that the distribution of price changes is not particularly skewed towards large or small price changes. However, this is not the case for some regions in the Central and Northern area, as for Liguria, Emilia and Lazio where the distribution of price variations has a great number of big price movements in the tails.

In table 3 we show the same statistics for the analysis derived on monthly scanner data. In this setting occurs that monthly price variations are bigger than inflation rate with relevant differences for the smallest regions, as Trentino, Friuli, Marche, Calabria and Sicily, where price changes are about 9 per cent (one point per cent above the others). Looking at the median the price changes are much pronounced (about 10 percent). Also in this case, the distribution of price changes is not specially skewed, as can be observed comparing mean with median statistics, with the exception as before for same regions.

[10] It is important to note that we compute the size of price increases and decreases as the difference of logarithm, so that the two successive price changes recorded during a temporary sale are equal in absolute terms (Dhyne et al., 2005).

Table 2: Mean frequency and size of price changes, weekly base (weighted statistics)

	Frequency (%)		Size (%)			
	Up	Down	Mean		Median	
			Up	Down	Up	Down
Piedmont	16.6	15.3	7.6	-7.1	7.4	-7.0
Lombardy	17.5	16.1	6.7	-6.4	7.2	-6.9
Liguria	13.2	11.9	6.9	-6.0	7.6	-6.9
Trentino	13.7	11.9	8.0	-8.0	8.4	-8.5
Veneto	17.9	16.7	7.2	-6.9	7.5	-7.2
Friuli-Venezia	15.7	14.2	8.1	-7.3	8.0	-7.4
Emilia Romagna	15.8	14.7	7.5	-6.8	8.2	-7.2
Tuscany	15.3	14.0	7.0	-6.6	7.1	-6.9
Umbria	15.4	15.0	8.0	-7.5	7.7	-7.6
Marche	17.2	16.0	7.2	-6.5	7.9	-7.2
Lazio	16.9	15.3	6.3	-6.0	7.0	-7.1
Abruzzo	16.9	16.0	7.6	-7.3	7.7	-7.5
Campania	17.3	15.7	6.2	-6.3	6.5	-6.7
Apulia	16.8	15.6	6.7	-6.3	6.8	-6.5
Calabria	13.3	12.8	7.9	-7.8	7.9	-7.9
Sardinia	16.6	15.5	7.6	-7.5	7.8	-7.7
Sicily	19.2	17.8	7.0	-6.6	7.0	-6.8

Source: AC Nielsen ScanTrack.

Table 3: Mean frequency and size of price changes, monthly base (weighted statistics)

	Frequency (%)		Size (%)			
	Up	Down	Mean		Median	
			Up	Down	Up	Down
Piedmont	21.4	19.7	8.6	-8.4	9.0	-8.2
Lombardy	22.1	20.6	7.9	-7.8	8.0	-8.5
Liguria	17.4	15.1	7.5	-6.9	8.9	-7.8
Trentino	18.0	16.3	9.5	-9.0	10.1	-9.4
Veneto	22.1	20.5	8.2	-8.2	8.9	-8.8
Friuli-Venezia	20.5	18.3	9.0	-8.6	9.9	-8.5
Emilia Romagna	20.7	19.2	8.1	-8.0	8.9	-8.6
Tuscany	20.2	18.6	8.2	-8.0	8.9	-8.6
Umbria	19.7	18.7	9.1	-9.0	9.3	-9.1
Marche	21.7	19.7	8.6	-8.1	9.4	-8.9
Lazio	21.1	18.9	7.7	-7.4	8.8	-8.7
Abruzzo	21.4	19.5	9.0	-8.6	9.0	-8.8
Campania	22.1	19.7	7.9	-7.6	8.8	-8.1
Apulia	21.3	19.2	7.9	-7.4	7.9	-8.1
Calabria	18.1	17.0	9.1	-9.3	8.8	-9.4
Sardinia	20.9	19.0	9.1	-8.7	9.4	-8.5
Sicily	23.9	22.4	8.4	-8.2	8.3	-7.9

Source: AC Nielsen ScanTrack.

5 Conclusion

Looking at high frequency data, price rigidity seems to be lower than founded by empirical studies who make use of CPI data. The explanation of this differences could be related to our data availability. It is to be note that in Italy it is very difficult to have scanner data at outlet level[11]. So, by the researcher viewpoint there are two alternatives: analyzing the CPI micro data available from National Bureau of Statistics or, if possible the aggregated scanner data bought by some Institutions. Nonetheless our data set have many disadvantages in that they are not elementary price quotes at store level but unit values averaged across store, actually they represent better chance to analyze price rigidity and other issues. Empirical analysis shows that price stickiness measured at regional level are linked to threshold admitted for price changes. We have introduced a minimum boundary to take into account the aggregation error has been previously explained. Comparing with others paper that make use of scanner data at store level, the price rigidity differences are not very marked.

[11]Indeed, these data are provided not free of charge by private institution.

REFERENCES

Baudry, L., Le Bihan, H., Sevestre, P. and Tarrieu, S. (2004). Price Rigidity: Evidence from the French CPI Micro-Data. ECB Working Paper, No. 384

Bils, M. and Klenow, P. (2004). Some Evidence on the Importance of Sticky Prices. Journal of Political Economy, 112 (5)

Klenow, P. and Kryvtsov, O. (2008). State-Dependent or Time-Dependent Pricing: Does It Matter for Recent US Inflation? Quarterly Journal of Economics, 123 (3)

Dhyne, E., Alvarez, L., Le Bihan, H., Veronese, G., Dias, D., Hoffman, J., Jonker, N., Lnneman, P., Rumler, F. and Vilmunen, J. (2005). Price Setting in the Euro Area: Some Stylized Facts from Individual Consumer Price Data. ECB Working Paper, No. 524

Kehoe, P. and Midrigan, V. (2010). Prices Are Sticky After All. Federal Reserve Bank of Minneapolis - Research Department Staff Report, No. 413

Fabiani, S., Gattulli, A., Sabbatini, R. and Veronese, G. (2005). Consumer Price Behaviour in Italy: Evidence from Micro CPI Data. ECB Working Paper, No. 449

DEFINITION OF THE STATISTICAL MEASURES

In order to calculate our statistics with respect to the stickiness of prices in Italy, we create four binary variables based on the price trajectories in our scanner data sets, in a similar way as in Dhyne *et al.* (2005). Let p_{ijt} the price of a product i, sold in region j, at time t. Because we only include adjacent price observations in our analysis, we first create an indicator variable I_{ijt}^{obs}, attached to p_{ijt}, that takes a value of 1 if a price quote for product i in region j from the previous period is observed and 0 otherwise. The second binary variable that we create is a price change indicator for product i in region j at time t:

$$I_{ijt} = \begin{cases} 1 & \text{if } abs(p_{ijt}/p_{ijt-1} - 1) > s \qquad \text{where s=0, 0.01, 0.02} \\ 0 & \text{otherwise} \end{cases}$$

In the same way, we also create a price increase and price decrease indicator for product i in region j at time t:

$$I_{ijt}^{+} = \begin{cases} 1 & \text{if } p_{ijt}/p_{ijt-1} - 1 > s \qquad \text{where s=0, 0.01, 0.02} \\ 0 & \text{otherwise} \end{cases}$$

$$I_{ijt}^{-} = \begin{cases} 1 & \text{if } p_{ijt}/p_{ijt-1} - 1 < s \qquad \text{where s=0, -0.01, -0.02} \\ 0 & \text{otherwise} \end{cases}$$

Adding these indicator variables to our data set, calculation of the frequency of price change, increase and decrease for product i in region j across all t time periods are straightforward:

$$F_{ij} = \frac{\sum_{t=2}^{T} I_{ijt}}{\sum_{t=2}^{T} I_{ijt}^{obs}}$$

$$F_{ij}^{+} = \frac{\sum_{t=2}^{T} I_{ijt}^{+}}{\sum_{t=2}^{T} I_{ijt}^{obs}}$$

$$F_{ij}^{-} = \frac{\sum_{t=2}^{T} I_{ijt}^{-}}{\sum_{t=2}^{T} I_{ijt}^{obs}}$$

The average size of price change, price increases and decreases are calculated as follows:

$$\overline{\Delta} p_{ij} = \frac{\sum_{i=2}^{t} I_{ijt}(lnp_{ijt} - lnp_{ijt-1})}{\sum_{i=2}^{t} I_{ijt}}$$

$$\overline{\Delta} p_{ij}^{+} = \frac{\sum_{i=2}^{t} I_{jt}^{+}(lnp_{ijt} - lnp_{ijt-1})}{\sum_{i=2}^{t} I_{ijt}^{+}}$$

$$\overline{\Delta} p_{ij}^{-} = \frac{\sum_{i=2}^{t} I_{jt}^{-}(lnp_{ijt} - lnp_{ijt-1})}{\sum_{i=2}^{t} I_{ijt}^{-}}$$