Price dispersion and inflation rates: evidence from scanner data

Castellari, E.[1], Moro, D. [1], Platoni, S. [2], Sckokai, P. [1]

[1] Dipartimento di Economia Agroalimentare, Università Cattolica del Sacro Cuore, 29122, Piacenza, Italy
[2] Dipartimento di Scienze Economiche e Sociali, Università Cattolica del Sacro Cuore, 29122, Piacenza, Italy

elena.castellari@unicatt.it

Paper prepared for presentation at the 4th AIEAA Conference
"Innovation, productivity and growth: towards sustainable agri-food production"

11-12 June, 2015
Ancona, Italy

Summary

In this paper we investigate the relationship between price dispersion and inflation; we use weekly retail scanner data from 2009 to 2011 to measure price dispersion and inflation for several dairy products. We implement a linear model to investigate the linkage between price dispersion and consumer price indexes. As in the previous literature, we obtain mixed results with respect to the relationship between price dispersion and inflation, and further investigation and theoretical refinement are needed to identify a common pattern.

Keywords: price dispersion, consumer price indexes, food inflation, scanner data, dairy products

JEL Classification codes: L11, L66

Price dispersion and inflation rates: evidence from scanner data

Castellari, E.[1], Moro, D.[1], Platoni, S.[2], Sckokai, P.[1]

[1] Dipartimento di Economia Agroalimentare, Università Cattolica del Sacro Cuore, 29122, Piacenza, Italy

[2] Dipartimento di Scienze Economiche e Sociali, Università Cattolica del Sacro Cuore, 29122, Piacenza, Italy

INTRODUCTION

Price dispersion, i.e. perfectly homogeneous products sold at different prices by different sellers, has been often observed in grocery markets as a deviation from the law of one price. In the literature, sellers' menu costs and consumers' search costs are often invoked as possible reasons behind such phenomenon.

In a scenario with identical buyers and sellers and perfect information the Bertrand outcome (perfectly competitive prices) becomes the unique Nash equilibrium. However, under imperfect information, firms can act as local monopolist since consumers incur in search cost to find the lower price. Price dispersion can also be explained by product differentiation: an identical product sold by different sellers can be perceived as differentiated due to specificity in locations (Hotelling, 1929). Furthermore, the overtime persistence of price dispersions can be attributed to the fact that sellers may set prices randomly to prevent consumers to identify stores with low-prices (Lach, 2002; Varian, 1980).

Nonetheless, the linkage between price dispersion and inflation is still controversial. Reinsdorf (1994), using micro-data from the U.S. Consumer Price Index (CPI) for nine cities and several products, has found a negative relationship between price dispersion and inflation rates, while other papers have found a positive relationship (see, for example, Caglayan et al, 2008). Most of the available literature is focused on the impact of expected and unexpected general inflation (typically measured by a Consumer Price Index (CPI) or similar) on price dispersion, where the main debate is on the specification of the empirical model that can capture the main economic phenomena analysed by the theory (i.e. consumers' search costs, sellers' menu cost and retail market structure). Very few papers analyse the relationship between Product Specific (PS) inflation and price dispersion (one exception is Caglayan et al, 2008). At this level of analysis, we believe that the causal relationship between the two phenomena may be reversed. In fact, we expect that competition among retail chains and retail store formats may lead to price dispersion, which in turn may impact PS inflation.

For this reason, in this paper we use retail scanner data on several dairy products for analyzing the relationship between price dispersion and PS inflation. Instead of using the standard CPI approach, we measure PS inflation using the superlative Gini-Eltetö-Köves-Szulc (*GEKS*) price index. This type of index is particularly suitable for scanner data, since it allows us to properly consider the changes of the product basket over time, while being free of the so-called "chain drift bias" problem (Ivancic et. al, 2011). Such effect is especially important when analysing price dispersion, since the chain drift bias is linked, for example, to the "price and quantity bouncing" effect due to temporary sales (i.e. the tendency by households of stocking up during sale periods and consuming from inventory at times when the products are not on sale).

More specifically, we analyse price dispersion at the retail chain level, thus focusing specifically on the pricing strategies carried out by large retailers. Thus, our work adds evidence to the recent work by Anania and Nisticò (2014) that have investigated price dispersion between large retailers and traditional shops in Italy.

DATA AND METHODS

We use SymphonyIRI scanner data to compute price indexes and price dispersion within different dairy products segments: Mozzarella cheese, spreadable cheese, butter and three sub-segments within the refrigerated milk (high quality, semi-skim and micro-filtered milk). This scanner database provides brand level weekly prices and sales, with and without promotion, for 400 points of sales which belong to 14 retailing chains along 156 weeks from January 2009 to January 2012.

Since we do not observe the exact location of the 400 points of sales, we aggregate the data at the chain level. This aggregation leads us to capture a price dispersion which is not related to spatial distance but to differences in chain pricing strategies.

We account for price dynamics within each product segment using the *GEKS* index. Let P_{ij} be the Fisher index between entities i and j ($\mathrm{j} = 1, \dots, \mathrm{M}$) and P_{kj} be the Fisher index between entities k and j. The *GEKS* index between i and k will be the geometric mean of the two Fisher indexes (Ivancic et al., 2011). We can, then, write the $\mathrm{GEKS}_{\mathrm{i,k}}$ as follows:

$$GEKS_{i,k} = \prod_{j=1}^{M} \left(\frac{P_{ij}}{P_{kj}}\right)^{\frac{1}{M}} \tag{1}$$

Ivancic et al. (2011) proposed to use the *GEKS* index to make comparison among T different time periods, $t = 1, \dots, T$. Considering the reference time period $\mathrm{t} = 0$, the *GEKS* price index between 0 and t will be:

$$GEKS_{0,t} = \prod_{l=0}^{T} \left(\frac{P_{0l}}{P_{tl}}\right)^{\frac{1}{T+1}} = \prod_{t=1}^{T} GEKS_{t-1,t} \tag{2}$$

As it is clear from equation (2), and differently from other bilateral indexes, multilateral indexes satisfy the Fisher’s circularity test (Fisher, 1922) which allows to directly compare entities among each other or through their relationship with a third one. As a result, we can write the *GEKS* index between time period 0 and t ($GEKS_{0,t}$) as a period-to-period chain index ($\prod_{t=1}^{T} GEKS_{t-1,t}$).

In addition, the *GEKS* index is free of chain drift as it satisfies the multi-period identity test. The multi-period identity test was proposed by Walsh (1901) and Szulc (1983) as a method to test for the presence of a drift chain bias. Given price indexes among different time periods, the price index will not suffer from chain drift bias if the product of indexes among all possible time combinations is equal to one.

We measure the price dispersion as in Reinsdorf (1994) using a simple coefficient of variation (C_{it}). Given $\bar{P}$ as the simple average of prices $\bar{P} = 1/n_{it} \sum_{j=1}^{nit} P_{ijt}$ where n_{it} are the number of chains and P_{ijt} is the price charged by the j^{th} chain for product i in week t. Then the coefficient of variation C_{it} measure the distance in prices among different chains for product i in week t:

$$C_{it} = \frac{1}{\bar{P}} \sqrt{\sum_j \left(P_{jit} - \bar{P}\right)^2 / (n_{it} - 1)} \qquad (3)$$

We compute the coefficient of variation using as P_{jit} a simple average of all the prices for product i recorded in a given chain (CV) or, alternatively, the price of a unique item (CV-item), chosen as being always present along the 156 weeks in all the 14 store chains.

The final dataset for estimation consists of a time series of coefficient of variations and *GEKS* indexes over the 156 observed weeks. In table 1 summary statistics of the variables used are reported. The unemployment rate and the weekly European Brent spot price[1], a proxy for the price of energy, are used as control variables in the econometric model.

INSERT TABLE 1

Figure 1-6 compare the variation CV and CV-item variables with the price dynamics along the 156 weeks for each of products analyzed. As we can notice, the two measures of price dispersion (CV and CV-item) differ consistently among segments. The refrigerated milk is the segment exhibiting less price dispersion (figure1-3). Specifically, the high quality milk is the one with lower price dispersion among the three refrigerated milk segments. Within the high quality refrigerated milk market the CV looks stable, while the variation of the CV-item present a negative spike

[1] The unemployment rate is provided by the Italian Institute of Statistic (ISTAT), www.istat.it. The weekly Europe Brent spot price is collected from the US Energy Information Administration www.eia.gov

between the 60th and 76th weeks. In the micro filtered sub-segment, while the CV shows a repeated fluctuation the CV-item presents few big spikes along the time. Similar behavior of the CV and CV-item can be seen in figure 3, however the CV show smaller fluctuations. The other three market segments (mozzarella cheese, spreadable cheese and butter) show a stronger dispersion in prices, especially if we consider the dispersion of a single item within the segment (figure 3-6).

The patterns of the GEKS present some similarities between the high quality and the semi-skim sub-segments, completely different appears the price variation along the 156 weeks within the micro-filtered milk sub-segment. Observing the patterns of the GEKS in figure 5 and 6 we can observe the two segments are affected by an increase of prices during the last 50 weeks of our observations while the variation of prices within the mozzarella cheese segment does not show a specific upward trend.

INSERT FIGURE 1-6

To determine the relationship between price dynamics and price dispersion we estimate the following econometric model:

$$GEKS_t = \beta_0 + \beta_1 X_t + \beta_2 Y_t + \varepsilon_t \qquad (4)$$

where $t = 1, \dots, 156$ indexes the weeks; X_t is the measure of price dispersion (either CV or CV-item), and Y_t indicate the set of control variables which may affect price dynamics.

We use Durbin-Watson Statistic to test for serial correlation and apply the Cochrane–Orcutt procedure to correct for serial correlated error taking a quasi-difference:

$$GEKS_t - \rho\, GEKS_{t-1} = \beta_0(1 - \rho) + \beta_1(X_t - X_{t-1}) + e_t \qquad (5)$$

where ρ is the coefficient obtained regressing the residuals with their lags: $\varepsilon_t = \rho\, \varepsilon_{t-1} + e_t$.

RESULTS AND DISCUSSION

Table 2 and 3 shows the results with both the CV and CV-item. Overall the coefficient of variation does not seem to have any effect on price indexes, with the exception of micro-filtered milk and mozzarella cheese, where the coefficients are negative and significant only at the 10% level (table3). Differently the price dispersion calculated only considering an item always present in all market chains shows that there is an effect of price dispersion on price indexes. The findings are mixed. In the high quality milk, the one with less variability on the CV-item, we observe a positive relationship between the dispersion on prices and the price index. Similarly the analysis shows a positive relationship between CV-item and price index in the butter milk segment.

On the contrary, in the other two sub-segments of the refrigerated milk market, as well as in the mozzarella cheese and spreadable cheese segments, this relationship is negative and significant, meaning the dispersion on prices of a unique item is related to a decrease of the inflation rate.

Thus, as in the previous literature, we obtain mixed results and further investigation is needed to identify a common pattern between price dispersion and PS inflation. This can be considered as a first step in this direction, which needs several refinements in terms of model specification and estimation strategy. Morever, as noted by Caglayan et al (2008), we are still missing a unified theoretical framework allowing researchers to derive a proper empirical specification for this type of analysis.

INSERT TABLE 2-3

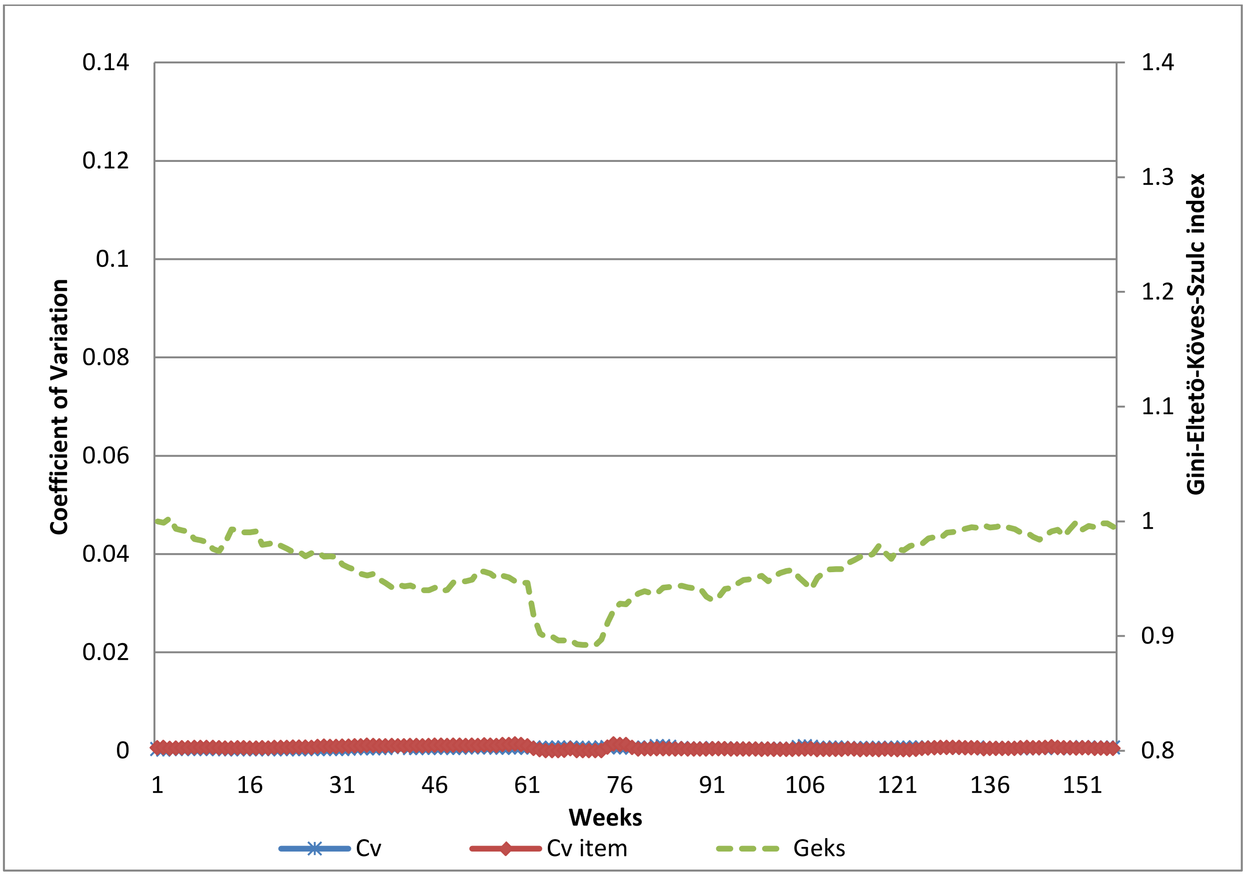
Tables and figures:

Table 1: Summary statistics for each segments analyzed

Segment	Variable	Obs	Mean	Std. Dev.	Min	Max
High Quality	Cv-item	156	0.00	0.00	0.00	0.00
	Cv	156	0.00	0.00	0.00	0.00
	Geks	156	0.96	0.03	0.89	1.00
	Unemployment rate	156	8.19	0.51	7.15	9.58
	Brent spot price	156	84.27	22.36	40.91	125.36
Microfiltered	Cv-item	156	0.00	0.00	0.00	0.00
	Cv	156	0.00	0.00	0.00	0.01
	Geks	156	0.98	0.04	0.93	1.07
	Unemployment rate	156	8.19	0.51	7.15	9.58
	Brent spot price	156	84.27	22.36	40.91	125.36
Semi-Skimmed	Cv-item	156	0.00	0.00	0.00	0.01
	Cv	156	0.00	0.00	0.00	0.00
	Geks	156	0.98	0.02	0.92	1.02
	Unemployment rate	156	8.19	0.51	7.15	9.58
	Brent spot price	156	84.27	22.36	40.91	125.36
Mozzarella	Cv-item	156	0.04	0.01	0.01	0.07
	Cv	156	0.01	0.00	0.01	0.03
	Geks	156	0.95	0.02	0.90	1.00
	Unemployment rate	156	8.19	0.51	7.15	9.58
	Brent spot price	156	84.27	22.36	40.91	125.36
Spredable cheese	Cv-item	156	0.05	0.02	0.01	0.13
	Cv	156	0.01	0.00	0.00	0.03
	Geks	156	0.99	0.05	0.91	1.16
	Unemployment rate	156	8.19	0.51	7.15	9.58
	Brent spot price	156	84.27	22.36	40.91	125.36
Butter	Cv-item	156	0.05	0.02	0.02	0.11
	Cv	156	0.01	0.00	0.00	0.01
	Geks	156	1.04	0.07	0.96	1.18
	Unemployment rate	156	8.19	0.51	7.15	9.58
	Brent spot price	156	84.27	22.36	40.91	125.36

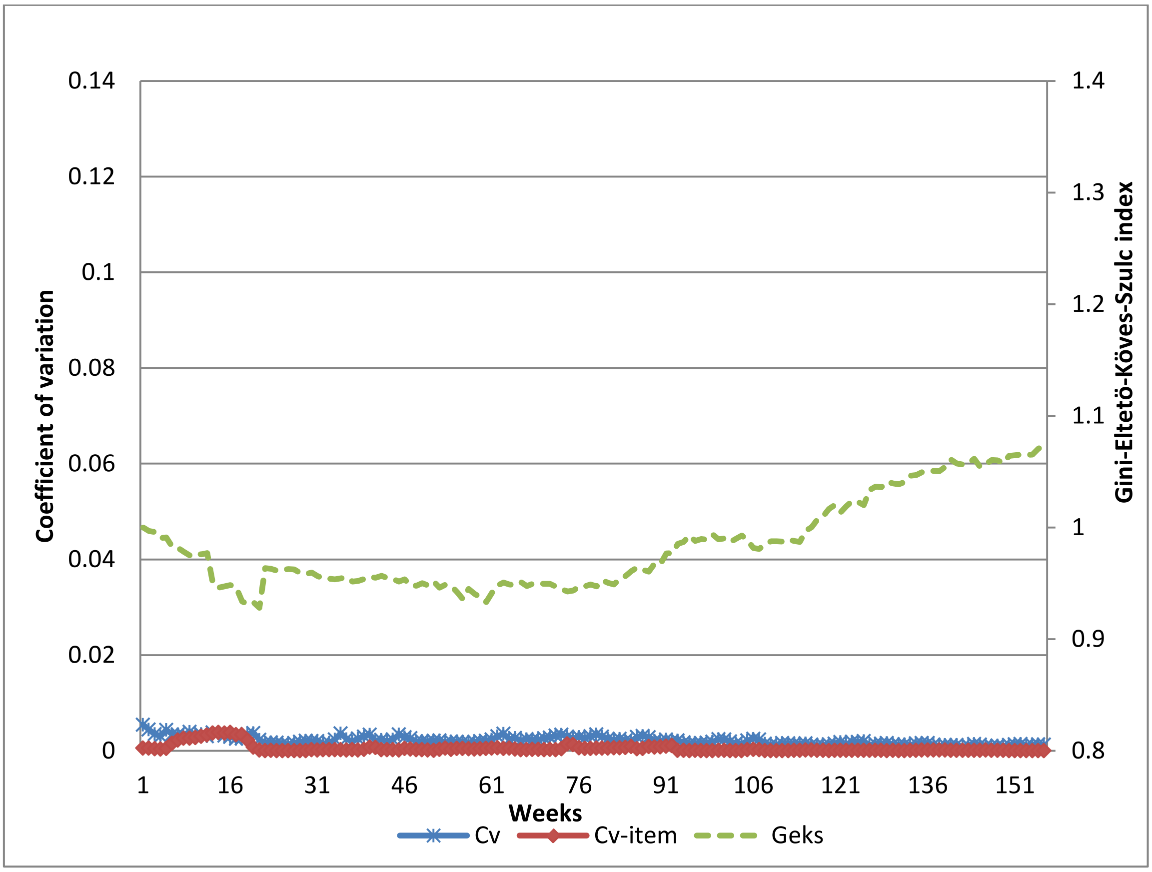
Source: Own Elaboration

Figure1: CV and GEKS index in the high quality refrigerated milk segment



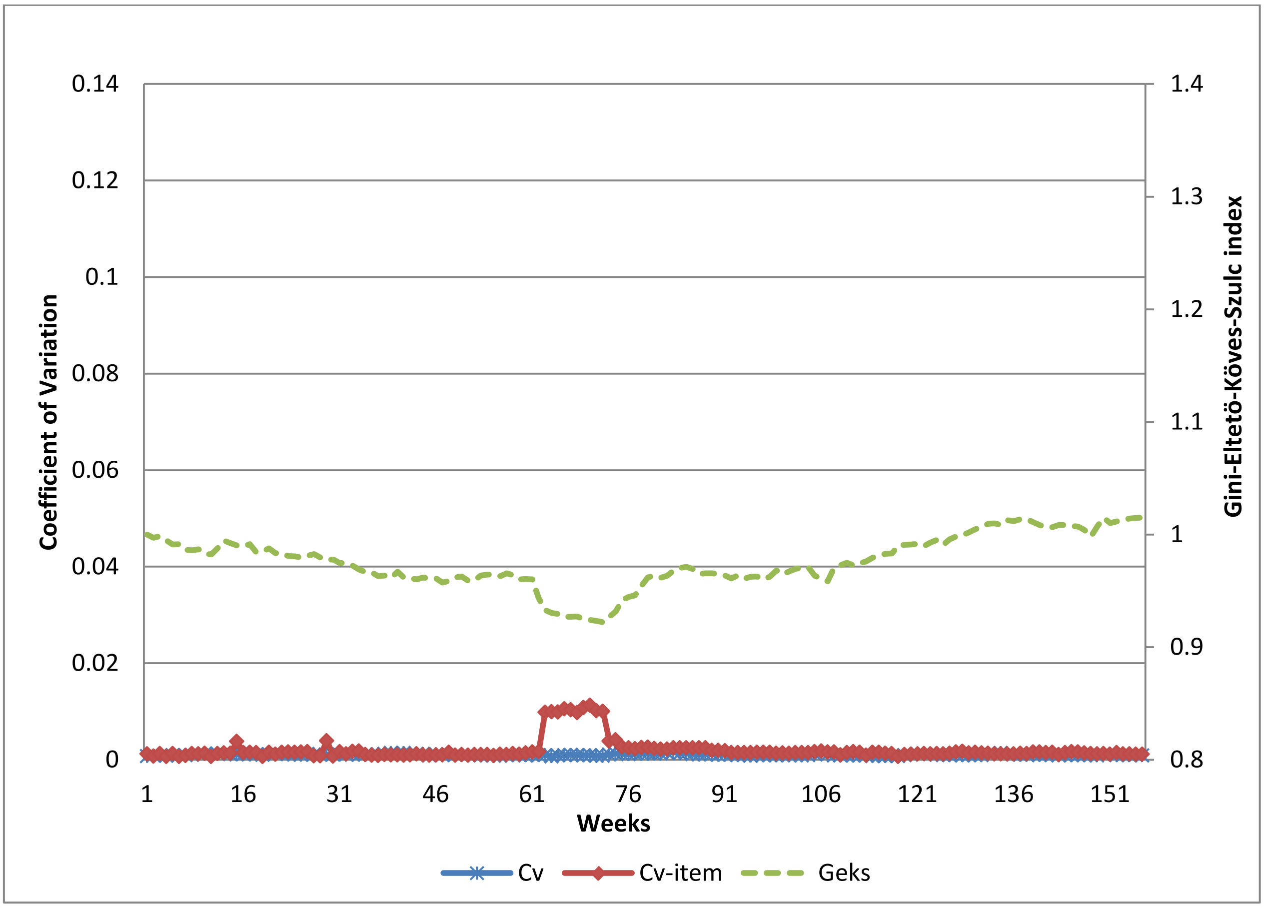
Source: Own Elaboration

Figure 2: CV and GEKS index in the micro-filtered refrigerated milk segment



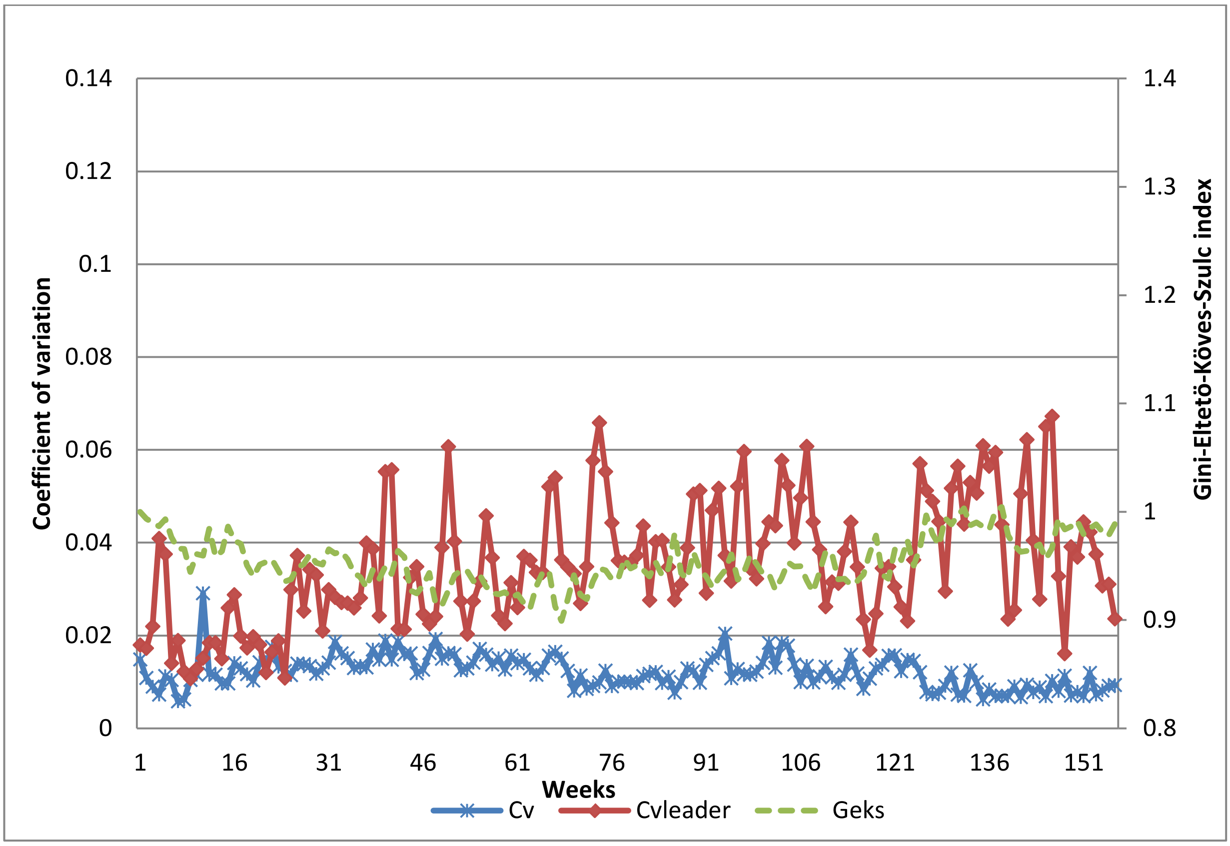
Source: Own Elaboration

Figure 3: CV and GEKS index in the semi skim refrigerated milk segment



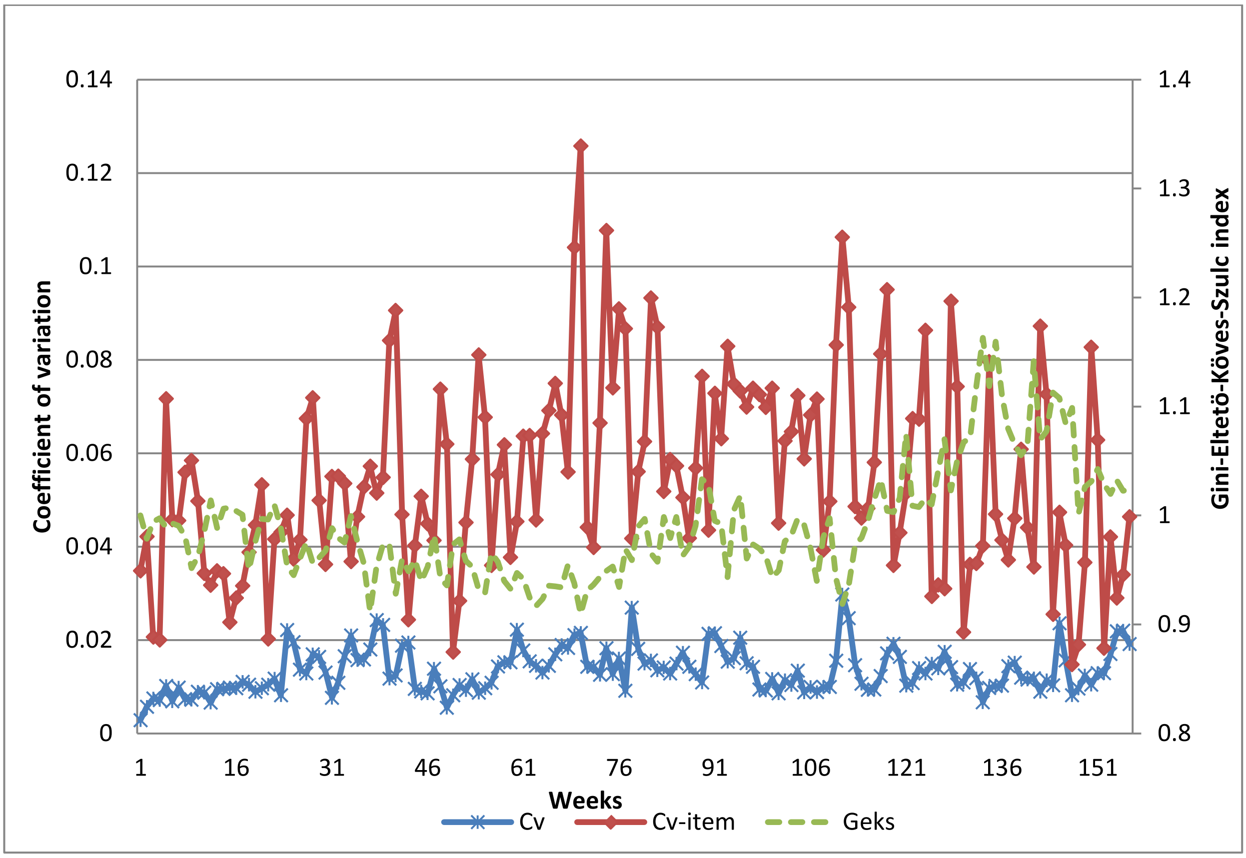
Source: Own Elaboration

Figure 4: CV and GEKS index in the Mozzarella cheese segment



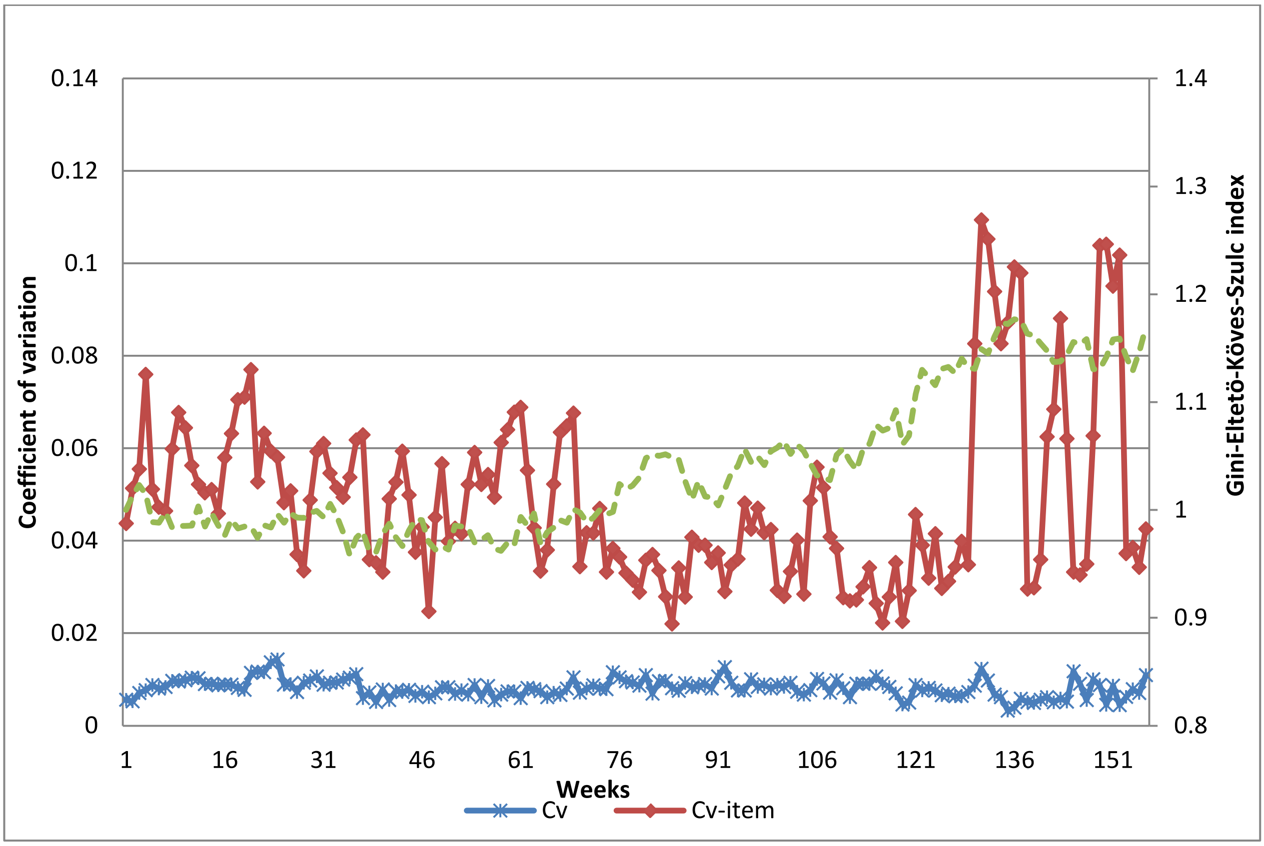
Source: Own Elaboration

Figure 5: CV and GEKS index in the spreadable cheese segment



Source: Own Elaboration

Figure 6: CV and GEKS index in the butter segment



Source: Own Elaboration

Table 2: Results for all different segments: GLS estimates using Cochrane–Orcutt procedure without (1) or with (2) control variables.

	High Quality Milk		Microfiltered Milk		Semi-skim Milk		Mozzarella		Spreadable cheese		Butter	
VARIABLES	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Cv-item	10.20***	10.02***	-3.19**	-3.07**	-0.90***	-0.89***	-0.16*	-0.16*	-0.44***	-0.44***	0.13*	0.16**
	(2.707)	(2.710)	(1.572)	(1.547)	(0.276)	(0.277)	(0.097)	(0.099)	(0.094)	(0.094)	(0.072)	(0.073)
Time trend	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00**	0.00*	0.00***	0.00*	0.00***	0.00***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Brent spot price		0.00		-0.00		-0.00		-0.00		0.00		0.00*
		(0.000)		(0.000)		(0.000)		(0.000)		(0.001)		(0.000)
Unemployment rate		-0.00		0.01***		-0.00		-0.01		-0.02		-0.02*
		(0.003)		(0.004)		(0.003)		(0.008)		(0.015)		(0.009)
Constant	0.84***	0.86***	0.85***	0.78***	0.87***	0.89***	0.94***	1.01***	0.96***	1.10***	0.91***	1.01***
	(0.045)	(0.055)	(0.028)	(0.040)	(0.037)	(0.046)	(0.010)	(0.067)	(0.020)	(0.127)	(0.026)	(0.077)
Durbin-Watson statistic (original)	0.0304	0.1113	0.0499	0.0558	0.146	0.137	0.4209	0.4940	0.5462	0.5971	0.2097	0.2902
Durbin-Watson statistic (transformed)	1.6906	1.7019	2.1836	2.1716	1.848	1.888	2.1224	2.1255	2.1304	2.1268	2.0032	1.9835
Observations	155	155	155	155	155	155	155	155	155	155	155	155
R-squared	0.12	0.13	0.20	0.24	0.12	0.13	0.05	0.06	0.18	0.20	0.21	0.22
Adj. R-squared	0.11	0.11	0.19	0.22	0.11	0.11	0.03	0.03	0.17	0.18	0.19	0.20

Standard errors in parentheses

*** p<0.0105, ** p<0.0505, * p<0.105

Table 3: Results for all different segments: GLS estimates using Cochrane–Orcutt procedure without (1) or with (2) control variables.

	High Quality Milk		Microfiltered Milk		Semi-skim Milk		Mozzarella		Spreadable cheese		Butter	
VARIABLES	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Cv	-4.90	-5.90	-1.56*	-1.67*	-0.54	-0.57	-0.48	-0.46	-0.89*	-0.91*	0.11	0.01
	(4.587)	(4.615)	(0.921)	(0.916)	(2.474)	(2.478)	(0.342)	(0.346)	(0.521)	(0.521)	(0.523)	(0.522)
Time trend	0.00**	0.00***	0.00***	0.00***	0.00***	0.00***	0.00**	0.00*	0.00***	0.00*	0.00***	0.00***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Brent spot price		0.00		-0.00		-0.00		-0.00		0.00		0.00*
		(0.000)		(0.000)		(0.000)		(0.000)		(0.001)		(0.000)
Unemployment rate		-0.01		0.01***		-0.00		-0.01		-0.02		-0.01
		(0.004)		(0.004)		(0.003)		(0.008)		(0.016)		(0.009)
Constant	0.84***	0.87***	0.86***	0.79***	0.86***	0.88***	0.94***	1.00***	0.95***	1.10***	0.91***	0.97***
	(0.049)	(0.060)	(0.023)	(0.037)	(0.042)	(0.051)	(0.011)	(0.067)	(0.021)	(0.135)	(0.027)	(0.077)
Durbin-Watson statistic (original)	0.149	0.124	0.042579	0.058105	0.0301	0.0416	0.60974	0.617999	0.526888	0.517498	0.151679	0.196132
Durbin-Watson statistic (transformed)	1.458	1.479	2.220118	2.209849	1.6395	1.682	2.078384	2.079578	2.200467	2.197564	2.051076	2.024583
Observations	155	155	155	155	155	155	155	155	155	155	155	155
R-squared	0.05	0.06	0.22	0.26	0.06	0.07	0.04	0.05	0.08	0.10	0.18	0.19
Adj. R-squared	0.04	0.04	0.21	0.24	0.05	0.05	0.03	0.03	0.07	0.08	0.17	0.17

Standard errors in parentheses

*** p<0.0105, ** p<0.0505, * p<0.105

References

Anania, G. and R. Nisticò. (2014). Price Dispersion and Seller Heterogeneity in Retail Food Markets. *Food Policy* 44:190-201.

Caglayan, M., A. Filiztekin and M.T. Rauh. (2008). Inflation, Price Dispersion, and Market Structure. *European Economic Review* 52: 1187–1208

Fisher, I. (1922). *The Making of Index Numbers*. Houghton Mifflin, Boston.

Hotelling, H.(1929). Stability in Competition. *Economic Journal* 39:41-57.

Ivanic, L., W.E. Diewert and K.J Fox. (2011). Scanner Data, Time Aggregation and the Construction of Price Indexes. *Journal of Econometrics* 161:24-35.

Lach, S. 2002. Existence and Persistence of Price Dispersion: An Empirical Analysis. *The review of Economics and Statistics* 84: 433-444.

Reinsdorf, M. 1994. New Evidence on the Relation Between Inflation and Price Dispersion. *The American Economic Review* 84: 720-731.

Szulc, B.J. (Schultz) (1983). Linking Price Index Numbers. In: Diewert, W.E.,Montmarquette, C. (Eds.), *Price Level Measurement*. Statistics Canada, Ottawa, 537–566.

Varian, H. 1980. A Model of Sales. *American Economic Review* 70:651-659.

Walsh, C.M. (1901). *The Measurement of General Exchange Value*. Macmillan, NewYork.