

French grocery store prices: What do we know?

(Very preliminary)

E. Chamayou*
CREST-LEI

October 30, 2015

Abstract:

This note provides an overview of supermarket price data made available by the comparison website quiestlemoinscher.com.

Keywords:

JEL Classification Numbers: XXX

*e-mail: *etienne.chamayou@ensae.fr*

1 Introduction

The price comparison website quiestlemoinscher.com was launched in XXX by the retail grocery store group Leclerc with a view to prove to consumers that Leclerc was cheaper than all its competitors.

2 Overview of price records

Data mainly consist of 13 pdf files containing prices observed at various grocery stores across France. The original purpose of these data was to create indexes reflecting the price level of the main competitors of Leclerc.

Table 1: Overview of price records

P	File name	Ordered by	Price range
0	200705_relevés_QLMC.pdf	Rayon Famille Produit Magasin	[0.28; 45.99]
1	200708_relevés_QLMC.pdf	Rayon Famille Produit Magasin	[0.25; 26.95]
2	200801_relevés_QLMC.pdf	Rayon Famille Produit Magasin	[0.15; 31.52]
3	200804_relevés_QLMC.pdf	Rayon Famille Produit Magasin	[0.16; 28.38]
4	200903_relevés_QLMC.pdf	No order	[0.36; 22.23]
5	200909_relevés_QLMC.pdf	Magasin Produit	[0.30; 31.46]
6	201003_relevés_QLMC.pdf	Produit Magasin (Chaîne)	[0.30; 09.99]
7	201010_relevés_QLMC.pdf	Rayon Famille Produit Magasin	[1.00; 99.00]
8	201101_relevés_QLMC.pdf	Rayon	[0.37; 09.99]
9	201104_relevés_QLMC.pdf	Produit	[0.31; 09.99]
10	201110_relevés_QLMC.pdf	Rayon Famille Produit	[0.16; 35.02]
11	201201_relevés_QLMC.pdf	Rayon Famille Produit	[0.15; 34.99]
12	201206_relevés_QLMC.pdf	Rayon Famille Produit	[0.16; 29.14]

Problems spotted at period level regarding prices:

- Period 6: it can be observed that all prices above 10.00 euros have been divided by 10 (hence min price for "Ricard pastis 45 degrés 50cl" is 1.00 euros. Since only two digits are provided after decimal point, this induces a loss in information. Furthermore,

there is no general trivial fix based on spread as there are products for which all prices are above 10 euros (e.g. "Ricard pastis 45 degrés 70cl"). One way to correct the data is to compare average price across periods whenever the product is part of another record. However, after multiplication by 10, prices can not be compared directly with other periods' prices due to the loss of one digit. One may thus want to set affected prices to missing values.

- Period 7: it can be observed that all prices below 1.00 euro have been multiplied by 100. It appears that the real highest price product within the period is "Mumm Cordon rouge champagne brut 75cl", the price of which does not exceed 35 euros so a simple fix consists in dividing all prices above that level by 100. The products thereby affected have been checked to ensure this strategy was valid (in the last resort, one could have proceeded to inter period comparisons).
- Period 8: same issue as in period 6
- Period 9: same issue as in period 6

3 Overview of stores and products by period

Table 2: Overview of stores and products

P	Date start	Date end	Nb rows	Nb stores	Nb products	Avg nb products by store
0	09/05/2007	25/05/2007	554,691	344	2,325	1,612
1	10/08/2007	31/08/2007	570,193	335	2,384	1,702
2	21/01/2008	12/02/2008	544,366	318	2,374	1,712
3	04/04/2008	30/04/2008	417,272	246	2,443	1,696
4	01/04/2009	30/04/2009	414,911	701	1,471	592
5	01/09/2009	28/09/2009	432,510	726	1,463	596
6	05/03/2010	03/04/2010	446,309	739	1,466	604
7	18/10/2010	16/11/2010	385,253	624	1,479	617
8	28/01/2011	22/02/2011	357,188	634	1,383	563
9	28/04/2011	20/05/2011	240,710	637	954	378
10	17/10/2011	09/11/2011	430,968	640	1,674	673
11	30/01/2011	22/02/2011	464,604	617	1,657	753
12	12/05/2012	01/06/2012	607,185	605	1,805	1,004

Within each period, instances of duplicates ("Produit", "Magasin") are looked for i.e. cases in which two prices are reported when only one is expected. It appears that it may be due to

- two products being reported under the same name (pack of n bottles of milk of bottle bought alone, duplicates are found for many stores but not all?)
- two stores being reported under the same name (duplicates are found for virtually all products and only one store)
- ???

Table 3: Overview of product duplicates

P	Nb of prod w/ duplicates	Nb prod w/ abnormal price spread
0	2	1
1	10+	1
2	7	1
3	8	1
4	4	1
5	3	1
6	2	need to fix prices first
7	4	need to fix prices first
8	4	need to fix prices first
9	0	0 (one due to price pbms)
10	10+	1+2 suspects
11	10+	0+1 suspect
12	10+	0+1 suspect

4 Overview of product categories

Within each period, a product label is uniquely matched with a "Rayon" and "Famille". Each "Famille" is a subcategory of a "Rayon" (really?). Categories nevertheless vary across periods.