Stylized Facts From Consumer Prices of Multi-Channel Retailers in Mexico

Diego Solórzano (jsolorzano@banxico.org.mx)

Banco de México | General Directorate of Economic Research

This paper characterizes the frequency, size and dispersion of price changes as observed on the website of eight large retailers in Mexico between 2016 and 2020 and compare them with the same price moments using data from brick and mortar stores of the same retailers. That is, this study focuses on the stylized facts from consumer prices at multi-channel retailers in Mexico.

Methodology and Data

Price moments are calculated for fairly homogeneous product categories by sales channel (online and offline) and retailer.[1] These statistics are then compared at retailer level across sales channels. Examples of product categories include milk, butter, soap, dish washer, socks, TVs, etc.

Two main data sources in this study: (i) online prices gathered by Banco de México through web scraping, it encompasses over 14 million price quotes from more than 150 thousand different products across eight retailers; and (ii) prices at brick and mortar stores gathered by INEGI for CPI calculation purposes (observations come from the same retail chains for which online prices are available), it comprehends a little less than one million price quotes from about 22 thousand different products.

Results

Data from 2016-2019 suggests that, for most retailers, prices observed in brick and mortar stores (offline) change more frequently than those observed on websites (online). However, given a price change, online prices tend to change by larger amounts than offline prices. Moreover, the categories' extensive margins of price adjustments across channels exhibit a positive relationship, while that is less clear for the intensive margin.

In 2020, for the product categories and retailers in the study, the less frequent but larger online prices changes relative to offline prices changes holds; and the frequency increased by around 5 p.p., on average, in both online and offline sales channels, while the size of adjustments barely changed relative to previous years.

Other results in the paper include an analysis on the distribution of price changes and standardized price adjustments.

Finally, the paper provides evidence on the compositional differences between sales channels (e.g. average price and share of missing products) stemming from the data collection techniques (i.e. census-like and survey approaches).
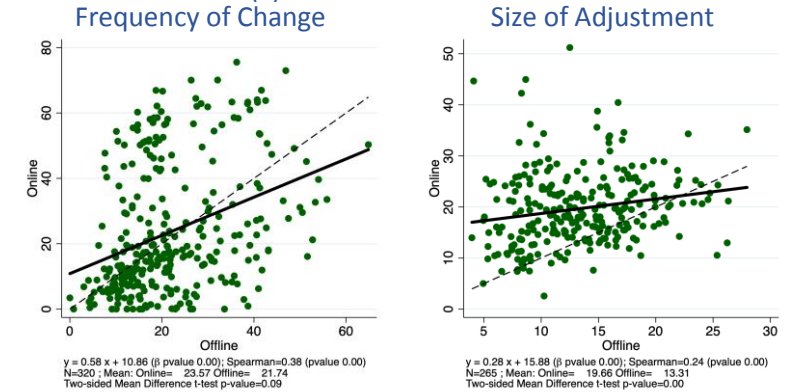
Comparison by Retailer[2]

(a) Frequency of Change

	Data 2016-2019				Data 2020				Difference		
	Average		Ho: Equality	Categories	Average		Ho: Equality	Categories			
	Online	Offline	p-value		Online	Offline	p-value		Online	Offline	Categories
	(a)	(b)	(c)	(d)	(f)	(g)	(h)	(i)	(f) - (a)	(g) - (b)	(i) - (d)
Retailer 1	11.16	16.02	0.00	38	12.10	16.16	0.01	27	0.94	0.14	-11
Retailer 2	14.47	22.95	0.00	38	17.76	24.13	0.00	26	3.29	1.18	-12
Retailer 3	15.27	19.30	0.09	52	14.46	20.98	0.01	42	-0.81	1.68	-10
Retailer 4	10.96	19.60	0.00	23	23.76	25.96	0.68	13	12.80	6.36	-10
Retailer 5	18.10	25.86	0.00	48	23.39	29.61	0.00	48	5.29	3.75	0
Retailer 6	55.33	23.03	0.00	48	53.05	42.74	0.00	33	-2.28	19.71	-15
Retailer 7	22.91	20.24	0.26	39	36.03	25.22	0.00	40	13.12	4.98	1
Retailer 8	32.49	26.09	0.09	34	40.92	34.32	0.02	32	8.43	8.23	-2

(b) Size of Adjustment

	Data 2016-2019				Data 2020				Difference		
	Average		Ho: Equality	Categories	Average		Ho: Equality	Categories			
	Online	Offline	p-value		Online	Offline	p-value		Online	Offline	Categories
	(a)	(b)	(c)	(d)	(f)	(g)	(h)	(i)	(f) - (a)	(g) - (b)	(i) - (d)
Retailer 1	12.60	10.95	0.01	31	12.69	11.78	0.35	15	0.09	0.83	-16
Retailer 2	24.29	9.72	0.00	33	9.79	8.95	0.40	22	-14.50	-0.77	-11
Retailer 3	22.65	13.47	0.00	41	26.71	13.73	0.00	32	4.06	0.26	-9
Retailer 4	22.22	14.27	0.00	13							
Retailer 5	15.79	9.44	0.00	41	17.94	10.97	0.00	37	2.15	1.53	-4
Retailer 6	19.16	16.85	0.00	48	18.00	18.05	0.93	32	-1.16	1.20	-16
Retailer 7	21.45	14.13	0.00	33	22.69	18.33	0.00	35	1.24	4.20	2
Retailer 8	20.93	18.64	0.10	25	23.84	23.06	0.44	27	2.91	4.42	2

Stylized Facts From Consumer Prices[3]

(a) Data from 2016-2019

Frequency of Change

y = 0.58 x + 10.86 (β pvalue 0.00); Spearman=0.38 (pvalue 0.00)
N=320 ; Mean: Online= 23.57 Offline= 21.74
Two-sided Mean Difference t-test p-value=0.09

Size of Adjustment

y = 0.28 x + 15.88 (β pvalue 0.00); Spearman=0.24 (pvalue 0.00)
N=265 ; Mean: Online= 19.66 Offline= 13.31
Two-sided Mean Difference t-test p-value=0.00

(b) Data from 2020

Frequency of Change

y = 0.84 x + 4.80 (β pvalue 0.00); Spearman=0.61 (pvalue 0.00)
N=261 ; Mean: Online= 28.08 Offline= 27.67
Two-sided Mean Difference t-test p-value=0.65

Size of Adjustment

y = 0.65 x + 9.88 (β pvalue 0.00); Spearman=0.54 (pvalue 0.00)
N=201 ; Mean: Online= 19.77 Offline= 15.32
Two-sided Mean Difference t-test p-value=0.00

Conclusions

Rationalizing sample differences across sales channels is fundamental for drawing conclusions from price statistics originated from online and offline data sources. This is important as nominal rigidities are a key ingredient in macro models.

1/ These product categories are the lowest level of aggregation of the Mexican CPI, known as *genéricos* by Mexico's National Statistical Institute (INEGI). 2/ Statistics based on all non-food product categories in a given retailer. Food-related categories are reported in the paper. In both panels, columns (a), (b), (f) and (g) are unweighted averages of the categories observed in the retailer; columns (c) and (h) are p-values from two-sided mean difference tests; columns (d) and (i) report the number of (non-food) product categories in the retailer. Few non-zero price changes in 2020 in both online and offline channels prevent computing price statistics for Retailer 4 in Panel (b). 3/ Each scatter represents a product category in a give retailer. Subfigures depict all non-food product categories from all retailers in the sample. Food-related categories are reported in the paper. The vertical axis comes from the online (web scraped) data, while the horizontal axis comes from the offline (survey) data.