VIETNAM E-COMMERCE PLATFORMS: PRICING BEHAVIOR AND INFLATION NOWCASTING

Nguyen Duc Hieu, Bui Thi Thuy Linh

VIETNAM E-COMMERCE PLATFORMS: PRICING BEHAVIOR AND INFLATION NOWCASTING

I. Introduction

Inflation is one of the key interests of economists when it comes to evaluating the well-being status of an economy. However, estimating inflation is not an easy task. In partially dollarized and goldarized economies like Vietnam, individuals tend to base their expectation on inflation on the gold price and dollar price when making financial decisions. However, some study (Adrangi 2003, Blose 2010) reported that gold price and dollar price do not necessarily reflect the changes in the Consumer Price Index. This further emphasizes the need for a reliable inflation measure in Vietnam.

In Vietnam, for many years, the survey for the Consumer Price Index has been conducted by the national General Statistics Office (GSO). Although the GSO provides the CPI regularly on a monthly basis, the procedures they use to collect inflation data is expensive, complex and requires a lot of modifications. This results in (1) a heavy burden on official resources, and (2) potentially bias statistics that mismatch with the demand for faster and more accurate data from policy-makers and other users. Thus, there is a need for driving towards innovation and exploiting alternative sources of data for inflation statistics that are less costly and more accurate.

Data about the goods prices sold by online sellers have several advantages over offline data. One of the main advantages is the daily frequency of data allowing us to measure real-time inflation without any delays. Further, online data include the full history of goods prices that are useful to control for new-good biases arising from traditional data collection methods. In Vietnam, with the increasing popularity of e-commerce websites and a young population, data from the online market is becoming a promising alternative to traditional price data, collected by surveys from brick-and-mortar stores.

This research collaboration between Swansea University, UK and Vietnam Institute for Economic and Policy Research (VEPR) aims to meet this demand by making use of technological development in inflation measurement. Our study draws on and contributes to the growing research on the use of "web-scraping" technique in economic

studies (e.g., Cavallo 2013, Gorodnichenko 2017) and use the data collected to form the inflation expectations. This method provides several advantages over the traditional method. First, it offers savings on the costs of conducting data and reduces the work of the General Statistics Office (GSO). Second, real-time data could be useful for a wide range of users like policy-makers, households, investors and firms. Third, this method can be expanded to collect online prices of identical goods sold outside Vietnam and data that can be useful in international economics.

The e-commerce market introduces new behaviors from both sellers and buyers. In this study, we take a look at the pricing behavior of online sellers and study some of its mechanics like dynamic pricing, the pattern of sales, and the synchronization (Klenow and Malin 2010) of prices between different types of price changes and categories of related products. Secondly, we prototype a simple price index that mimics the formula of the GSO's Consumer Price Index, using the data we collected online. By comparing our Online Price Index and the CPI, we can assess the application of using online data as a compliment or alternative to the traditional methods of calculating CPI in Vietnam.

Overall, our study reported that some evidence of dynamic pricing is observable in the data collected. Empirical results from the subset data of food, foodstuffs, and related services showed that synchronization is strong between positive and negative changes in price and weak between price changes within different subcategories. On an aggregated level, a trend of inflation is observable, but comparisons between our computed price indexes and the CPIs of some subcategories showed that online prices may move in different patterns to the traditional price data collected by survey.

II. Methodology

1. Data

In order to collect data from the internet, we use the new technique so-called "web scraping". We utilize Python, a programming language to build an automated procedure that can record online prices from available retailer webpages (e.g., lazada.vn) every day. Next, we screen data to extract and store the relevant pieces of information in a database. To document and study the properties of online prices, we collect the data on a daily basis, starting from April 2018. Each observation includes the item's name, its posted price – before and after reduction, categorization of the item according to the website section naming and the date in which data is collected. Information about the seller's name, selling location and other information are also collected if available. Unified information that can be used to identify the same goods across different websites (e.g. barcodes) are not available on most websites, so we generalize the data into groups of

similar products in our statistics. Our data are classified according to the official Consumer Price Index basket structure, provided by the General Statistics Office of Vietnam, which in turn are built based on the United Nations' Classification of Individual Consumption According to Purpose (COICOP) with 11 main groups and over 650 smaller specifications.

2. Calculating the Online Price Index

We calculated the Online Price Index (OPI) for 4 subcategories in the Consumer Price Index (CPI) basket structure: (a) *Food, foodstuffs and related services;* (b) *Garments, hats and footwear;* (c) *Household, equipment and appliance;* and (d) *Culture, entertainment and tourism*. The indexes are aggregated on a weekly basis, starting from week 14 to week 47 of 2018 (from April 2nd to November 25th), where the Online Price Index for week 14 is set as 100. We compare our OPI with the official CPI, provided by the General Statistic Office of Vietnam, from April to November 2018. The CPI are aggregated on a monthly basis, with the index for March is set as 100.

The Online Price Index (OPI) use a variation of the geometrics mean Laspeyres formula as follow:

$$I^t = \sum_{i=1}^{n} W_i \times dp_i^t$$

And for each category i that contains s types of items:

$$dp_i^t = \left(\prod_{j=1}^{s} \frac{p_j^t}{p_j^{t-1}} \right)^{\frac{1}{s}}$$

III. Pricing Behavior in Vietnam's E-commerce Market

1. Descriptive Statistics

Although the time span of 8 months (from April to December 2018) is slightly shorter than the time span available for researchers studying online prices (typically one year or longer), the data provided is sufficient for us to start analyzing the properties of online prices on Vietnamese platforms. There are also other advantages of using our data. The number of goods is much broader than previous researches of the same type. By the end of November 2018, the collected data covered a broad range of nearly 7 million items, across 50 websites. Table 1 shows the number of goods collected in 10 different main categories. Of the data collected, approximately 33% are categorized as Household

equipment and appliances. The two groups Garments, hats and footwear and Culture, entertainment and tourism also account for a large proportion of the data collected (18.27% and 14.68%, respectively). Only 1.7% of the goods collected are not classified and thus are omitted from our statistics.

Category	**Number of goods**	**Percentage**
Household equipment and appliances	2,292,898	33.01%
Garment, hats and footwear	1,268,839	18.27%
Culture, entertainment and tourism	1,019,761	14.68%
Transportation and communication	535,123	7.70%
Food, foodstuffs and services	160,459	2.31%
Medicine and health-care services	118,882	1.71%
Education	91,211	1.31%
Housing, electricity, water, fuels and building materials	27,602	0.40%
Beverage and tobacco	19,934	0.29%
Other goods and services	1,293,239	18.62%
Not classified	117,774	1.70%
Total	**6,945,722**	**100.00%**

Table 1: Data coverage by categories

Across 50 websites, *Lazada.vn* accounts for the largest number of items (more than 5.5 million items). *Adayroi.com* has the second largest number of items (332,970 items). *Shopee.vn* – a market platform that connects buyers and sellers – has 281,080 items (accounts for approximately 4% of all items collected). Table 2 shows the numbers of goods we collected by each website.

No.	Websites	Number of goods	Percentage
1	Lazada	5,569,319	80.18%
2	Shopee	281,080	4.05%
3	Adayroi	332,970	4.79%
4	Sendo	230,480	3.31%
5	Vatgia	124,673	1.79%
6	Yes24	107,999	1.55%
7	Tiki	96,656	1.39%
8	Meta	17,471	0.25%
9	Vuivui	15,009	0.21%
10	Chopp	13,237	0.19%
	Others	156,828	2.25%
	Total	**6,945,722**	**100.00%**

Table 2: Data coverage by websites

With the data collected, we take a cursory look at the distribution of prices. The results are shown in Table 3. *Household, equipment and appliances* items varies in prices the most (with standard deviation of approximately 0.5 between all items), Which is approximately 0.05 point higher than the second most price deviated group (*Culture, entertainment of tourism*). Although the medians of posted prices do not differ significantly between groups, at the 95% percentile, *Household, equipment and appliances* is the most expensive category. *Transportation and communication* comes second, followed closely by *Culture, entertainment and tourism*.

Category	Log prices		Percentile				
	Mean	SD	5%	25%	50%	75%	95%
Household, equipment and appliances	5.38069	0.53318	4000	101000	200000	439000	2999400
Culture, entertainment and tourism	5.37366	0.48628	8000	103000	208000	450000	1830000
Garment, hats and footwear	5.27533	0.35165	9000	105941	190000	325000	715000
Transportation and communication	5.26161	0.46226	2000	90000	168000	312000	1850000
Beverages and tobacco	5.13319	0.42092	9000	68000	135000	269000	700000

Food, foodstuffs and related services	5.03934	0.42839	3000	54300	110000	225000	550000
Education	4.99009	0.40350	4000	52000	89000	181000	505000
Other goods and services	5.24463	0.36793	4000	98000	174000	545000	786453

Table 3: Distribution of prices

2. Price changes and speed of adjustments – The case of Food, foodstuff and related services

In online stores, changing the price is less costly. The sellers do not have to, for example, manually print new price tags and re-attach them to items. Websites can also use algorithms to change the prices in response to anticipated variation in demand (throughout the week, month, or year) or current market conditions (competitors' prices, number of customers, inventories, etc.). This behavior is called "dynamic pricing". An evident of dynamic pricing in Vietnamese online stores can be seen in Table 4, where we document the frequency of changes in Food, foodstuffs and related services prices over different weekdays. Overall, the change in online prices of food and foodstuffs tends to fluctuate on Monday. Prices decreased on Monday and rise back again on Tuesday.

Day	Positive	Negative	Overall
Sunday	0.007695	0.005461	0.013156
Monday	0.016996	**0.028574**	**0.04557**
Tuesday	**0.022064**	0.005917	0.027981
Wednesday	0.008614	0.008417	0.017031
Thursday	0.009191	0.006702	0.015893
Friday	0.006112	0.014423	0.020535
Saturday	0.021676	0.004727	0.026404

Table 4: Daily Frequency of Price changes within a week

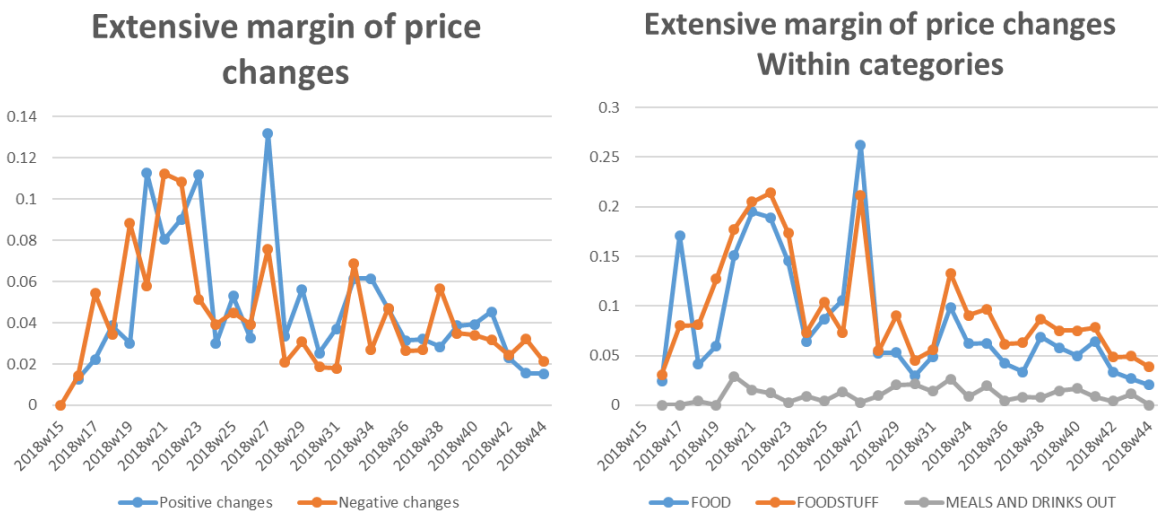

Figure 1: Margin of prices changes between positive and negative changes (left) and within different categories (right)

Figure 1 (left) shows the margins of price changes, in comparison between the positive and negative changes. The sizes of changes between positive and negative changes are quite similar over time. On the other hand, the sign of synchronization between different group of goods is not as evident, as show in *Figure 1* (right). This shows that the sign of synchronizations between different sub-categories of *Food, foodstuffs and related services* is weak. When prices of food and foodstuffs increases, the sellers don't see the need to increase the prices of meals and drinks out.

In order to separate price changes that reflect inflation and temporary price changes (temporary sales), following previous literatures, we defined temporary price changes as v-shaped sales – where the prices if the price returns to its original level within one week. With that in mind, we make a distinction between posted prices (i.e. prices we observe in the data) and regular prices (i.e., prices that exclude temporary sales).

Category	**Size (mean, %) (1)**	**Frequency (mean, %) (2)**	**No. of goods (3)**
Food	12.56	0.1573	110891
Foodstuff	14.62	0.1409	1284309
Meals and drinks out		0	24150

Note: one week V-shaped

Column (1): the absolute size of sales for the mean good measured by the log difference between the sale and regular price (multiplied by 100).
Column (2): reports the average weekly frequency of sales across goods (%).
Column (3) reports the number of goods.

Table 5: Sales: Size and Frequency (V-shaped, weekly)

This weekly frequency is comparable to the frequency of sales reported for prices in regular stores. The mean weekly frequency of sales (columns (2)), without weights, is in the range of 0.14% to 0.16%. There is substantial heterogeneity in the frequency across products: we do not find sales in meals and drinks out (see row (3)). The median size of sales is 12.6% to 14.6% with equal weights. Compare with previous research from Ukraine online market (Y. Gorodnichenko, O. Talavera, and M. Tian), the mean frequency is 1.7%–2.7% and the median size of sales is 10.5%–11.9% with equal weights. Thus we can conclude that in Vietnam e-commerce market, sales are less frequent, but are generally bigger in size, compare to previous studies.

As the dataset contains missing values, we identify sales with and without imputation, following preceded literatures of Nakamura and Steinsson (2008), Eichenbaum et al. (2011), and Kehoeand Midrigan (2015) to indentify temporary price changes:

- In the case of no imputation: missing values breaks the time series and the only changes between consecutive observations are considered regular price changes.
- In the case with imputation: if the next available observation after the missing value is within one week and there is no change in price before and after the missing observation (consider the series {2, n.a., 2}), the missing observation is imputed with the price which it stands between. In this example, the imputed series is {2, 2, 2}. Regular price changes are then identified on the imputed series in the same way as in the case of no imputation.

	Posted Price Changes		**Regular Price Changes**	
	Mean frequency	**Mean size**	**Mean frequency**	**Mean size**
No imputation				
Food	0.0541	16.4063	0.0407	16.6806

Foodstuff	0.0760	15.3539	0.0573	15.2670
Meals and drinks out	0.0088	27.3317	0.0073	26.6384
With imputation				
Food	0.0474	16.5623	0.0100	17.3854
Foodstuff	0.0684	13.9225	0.0135	13.9130
Meals and drinks out	0.0075	29.6366	0.0007	27.4682

Table 6: Price Changes - Regular vs Posted price

Table 6 reports the frequency and size of both posted price changes and regular price changes, with and without imputation. Although it is obvious that the frequency of posted price changes is higher than the frequency of regular price changes (because temporary sales are excluded in the latter), we find that the sizes of changes in regular prices changes are not significantly smaller. This indicates that the sizes of temporary sales is not much bigger that of regular changes in price. This observation is similar is both the case of no imputation and with imputation, with the difference is that the frequency of regular price changes significantly reduced when we imputed the missing data.

	Mean Frequency			**Mean Size**		
	Overall	**positive**	**negative**	**overall**	**positive**	**negative**
No imputation						
Food	0.0541	0.0301	0.0241	16.4063	17.1744	15.4464
Foodstuff	0.0760	0.0405	0.0355	15.3539	15.4569	15.2365
Meals and drinks out	0.0088	0.0046	0.0042	27.3317	29.2669	25.2236
With imputation						
Food	0.0474	0.0276	0.0198	16.5623	17.3528	15.4616

Foodstuff	0.0684	0.0373	0.0312	13.9225	14.5217	13.2063
Meals and drinks out	0.0075	0.0035	0.0039	29.6366	33.9677	25.7475

Table 7: Price Changes - Positive vs Negative changes

We further examine the changes in posted prices and compare the frequency of positive changes and negative changes in Table 7. Overall, both the frequency and size of positive changes in larger than that of negative changes. This result is consistent between the two case of with and without imputation. In the long term, when prices generally increase more than decrease, we can expect to observe inflation tendencies.

IV. Inflation Nowcasting

1. Food, Foodstuffs and related services

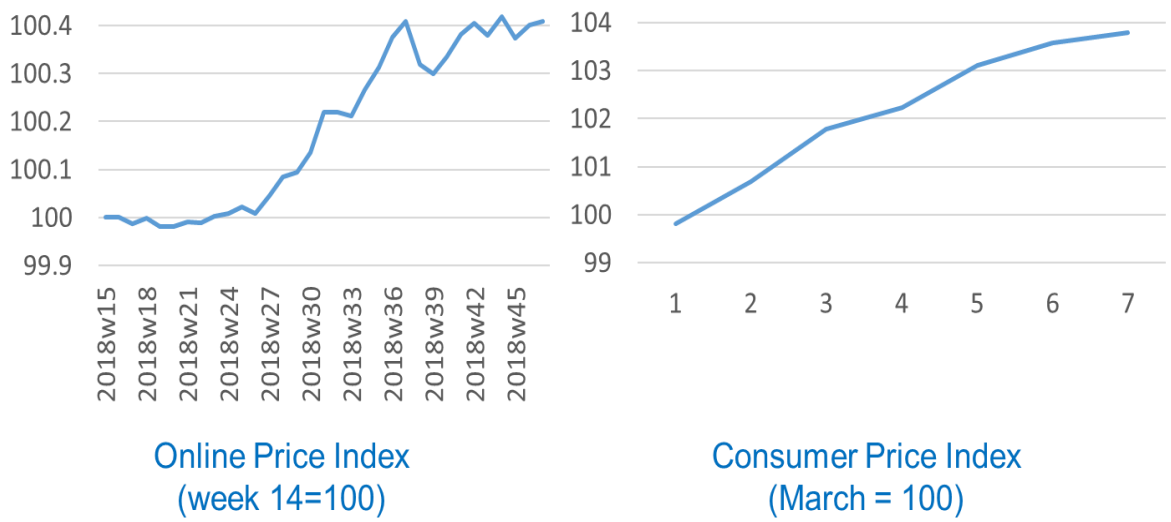

Figure 2: OPI (left) and CPI (right) – Category of Food, foodstuffs and related services

There are upwards changes in Online Price Index (OPI) for *Food, foodstuffs and related services* during the time span from Week 15, 2018 (from April 9th to April 15th) to week 47 (November 19th to November 25th). This trend is similar to the changes of CPI in *Food, foodstuff and related services* from April to the end of October. The size of changes differs greatly, however. While CPI record a change of nearly 4 percent in from April to October, OPI increases for only 0.4 percent. This shows that posted food prices do increase, but for much less amount than in traditional markets.

2. Garments, Hats and Footwear

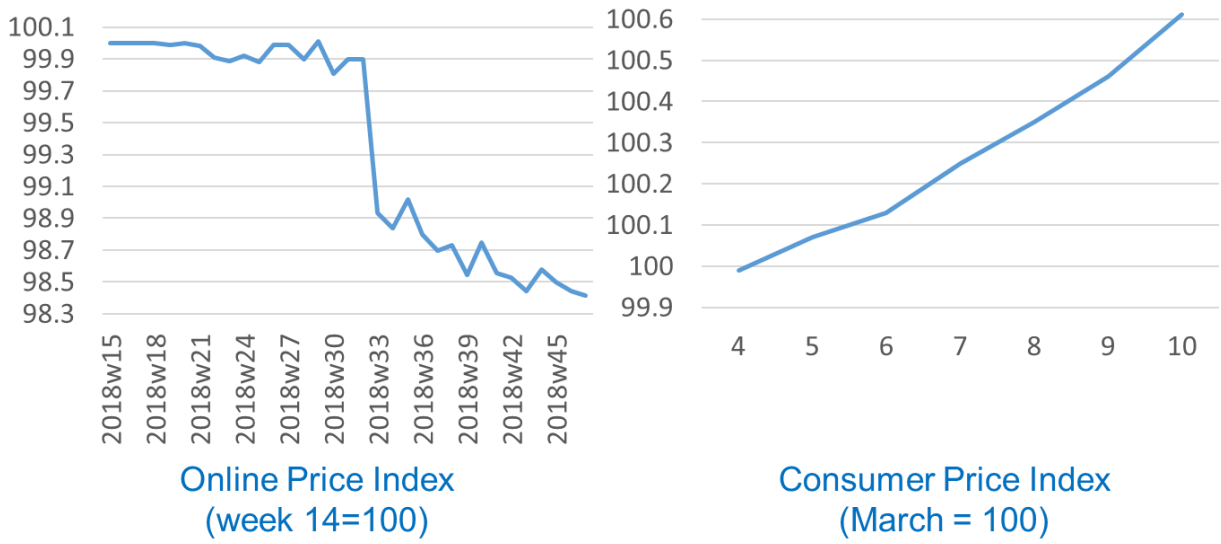

Figure 3 OPI (left) and CPI (right) – Category of Garments, Hats and Footwear

Unlike the case of *Food, foodstuff and related services*, the Online Price Index of subcategory *Garments, hats and footwear* shows a decrease of 1.5 percent during this period, especially during week 32 (From August 6th to August 12th), when the Online Price Index decreased nearly 1 percent. This trend contrasts with equivalent CPI during the same time period, where CPI for this subcategory increase for 0.6 percent. This implies that online sellers' behavior is independent from traditional sellers' pricing.

3. Household Equipment and Appliance

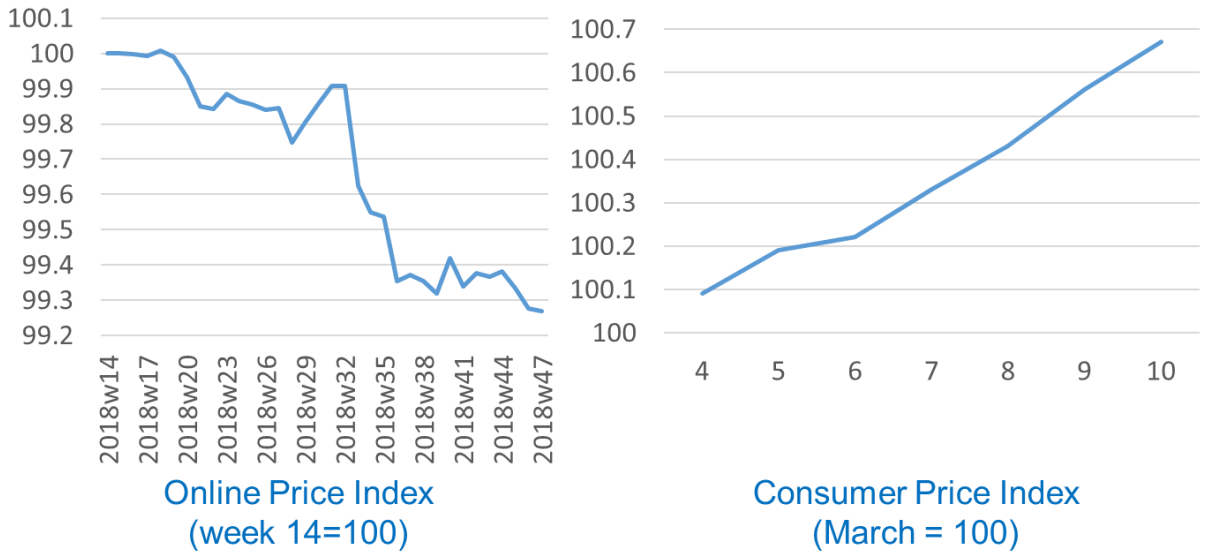

Figure 4 OPI (left) and CPI (right) – Category of Household equipment and appliance

For subcategory Household equipment and appliance, OPI decreased approximately 0.7 percent, while and CPI increased for 0.6 percent. The Online Price Index for this subcategory briefly increased (nearly 0.02 percent) during the time of week 29 to week 31 (From July 16th to August 5th), then fell heavily in the next two weeks (from August 6th to August 19th).

4. Culture, Entertainment and Tourism

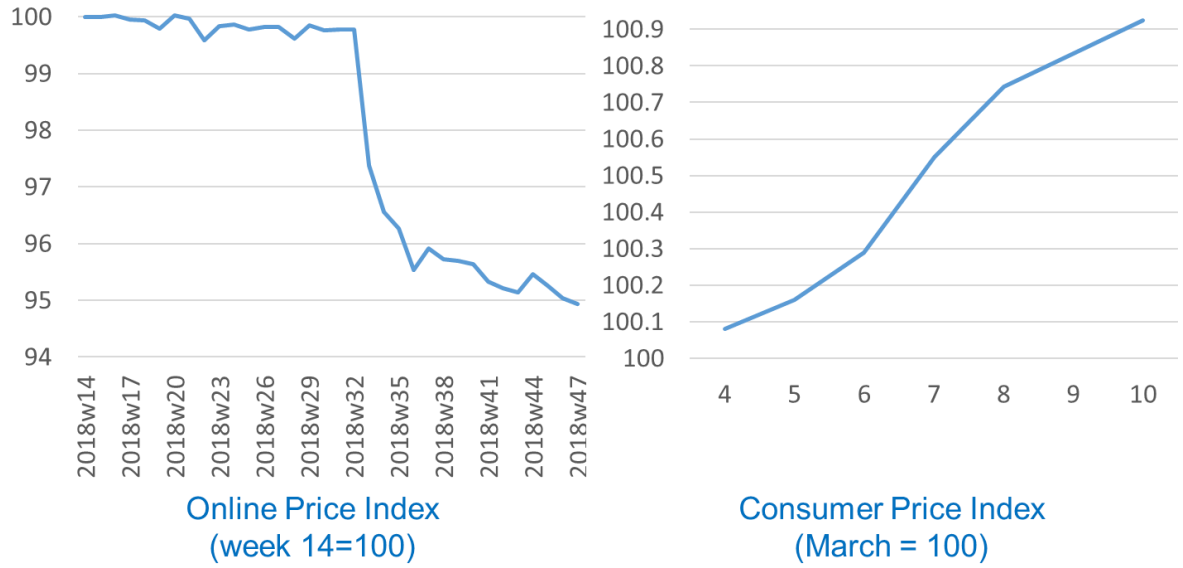

Figure 5 OPI (left) and CPI (right) – Category of Culture, entertainment and tourism

For *Culture, entertainment and tourism*, the results are similar to that of subcategory *Garments, hats and footwear* and *Household equipment and appliance*. The Online Price Index for this subcategory gradually decreased in the observed weeks, except for week 32-33 (from August 6th to August 19th), where it dropped approximately 4 percent. On contrary, CPI of the same group slowly increased for a total of 0.8 percent in the same time period.

V. Discussion and Conclusion

The internet offers limitless opportunities to study the pricing behaviors by enabling researchers with a large amount of data with near real-time updates, in an economic environment that allows for instant updates and prices comparison between different sellers.

Using the subset of data from the categories of Food, foodstuffs and related products, we study the characteristics of online sellers in Vietnam. Initial results show that in Vietnam, prices seem to change less frequent but with larger sizes, compare to similar studies. Upwards price changes seem to be concentrated on Mondays and

decrease in the day after, which is the sign of sellers changing their price dynamically according to demands from buyers. Our attempts to eliminate temporary sales (or V-shaped sales) from our time series of prices show that temporary sales is bigger in size than regular price changes, but not by significant margins. Overall, the sizes and frequencies of positive price changes are larger than that of negative price changes, which is a signal of inflation. These results are consistent when we apply imputation methods to missing data.

We also match the data collected with the Consumer Price Index basket structure provided by the Vietnam General Statistics Office to prototypes a similar price index to the CPI, called the Online Price Index (OPI). We calculate our price indexes for 4 main categories in the CPI basket structure: (a*) Food, foodstuffs and related services*; (b) *Garments, hats and footwear*; (c) *Household, equipment and appliance*; and (d) *Culture, entertainment and tourism* in the time span from April to November 2018. Comparison with CPI of the same period shows that price changes in online markets are much smaller than changes in CPI of corresponding subcategories. While online prices of *Food, foodstuffs and related services* increase during the time of our studies, the online indexes for 3 other categories reported a decrease in price, in contrary to the increasing trend of correspondent CPI.

It is evident from these mismatched comparisons between online prices the traditional price indexes that online data can provide us with new perspectives on inflation. While online data can provide us with unprecedented knowledge, we will need more data to be collected over time to produce more critical insights. On the other hand, in order to calculate better online price indexes, we need to find a better proxy for the number of goods purchased and extract more precise weights for each sub-category. One possible solution is to extract weights from the Vietnam Household Living Standard Survey. Another method is to collect the history of online transactions or the numbers of views/traffic on each item sold online.

References

Gorodnichenko, Yuriy, Viacheslav Sheremirov, and Oleksandr Talavera (2018). "Price setting in online markets: Does IT click?" In: Journal of the European Economic Association 16.6, pp. 1764–1811.

Gorodnichenko, Yuriy and Oleksandr Talavera (2017). "Price setting in online markets: Basic facts, international comparisons, and cross-border integration". In: American Economic Review 107.1, pp. 249–82.

Kehoe, Patrick and Virgiliu Midrigan (2015). “Prices are sticky after all”. In: Journal of Monetary Economics 75, pp. 35–53.

Cavallo, Alberto (2013). “Online and official price indexes: measuring Argentina’s inflation”. In: Journal of Monetary Economics 60.2, pp. 152–165.

Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo (2011). “Reference prices, costs, and nominal rigidities”. In: American Economic Review 101.1, pp. 234–62.

Blose, Laurence E (2010). “Gold prices, cost of carry, and expected inflation”. In: Journal of Economics and Business 62.1, pp. 35–47.

Klenow, Peter J and Benjamin A Malin (2010). “Microeconomic evidence on price-setting”. In: Handbook of monetary economics. Vol. 3. Elsevier, pp. 231–284.

Nakamura, Emi and Jón Steinsson (2008). “Five facts about prices: A reevaluation of menu cost models”. In: The Quarterly Journal of Economics 123.4, pp. 1415–1464.

Adrangi, Bahram, Arjun Chatrath, and Kambiz Raffiee (2003). “Economic activity, inflation, and hedging: the case of gold and silver investments”. In: The Journal of Wealth Management 6.2, pp. 60–77.

Thông cáo báo chí về một số nội dung cập nhật trong phương án tính chỉ số giá tiêu dùng thời kỳ 2015-2020. URL : http://www.gso.gov.vn/default_en.aspx?tabid=515&idmid=&ItemID=15520.

Other publications from VEPR:

VMM18Q4 Viet Nam Quarterly Macroeconomic Report, Quarter 4 - 2018

VMM18Q3 Viet Nam Quarterly Macroeconomic Report, Quarter 3 - 2018

VMM18Q2 Viet Nam Quarterly Macroeconomic Report, Quarter 2 - 2018

PD-07: Evaluation of the Law on Public Debt Management of Vietnam and some Policy Implications

PD-06: Minimum Wages in Viet Nam: Preliminary Observations

PD-05: A brief on impacts of Brexit on the World Economy and Vietnamese Economy

WP-18: Development characteristics of SME sector in Vietnam: Evidence from the Vietnam Enterprise Census 2006-2015

WP-17: The 19th National Congress of the Communist Party of China: Preparation for a New Era

CONTACT US

Viet Nam Institute for Economic and Policy Research
University of Economics and Business, Viet Nam National University, Ha Noi
Address: Room 707, E4 Building
144 Xuan Thuy strt, Cau Giay dist
Ha Noi, Viet Nam
Tel: (84-24) 3 754 7506 - 704/714
Fax: (84-24) 3 754 9921
Email: info@vepr.org.vn
Website: www.vepr.org.vn