Using Online Prices for Inflation Estimation and Pricing Behavior Research

A Case Study from Vietnamese Multichannel Retailers during the COVID-19 Pandemic

A Research Paper presented by:

Hoa Ngo
(Vietnam)

in partial fulfilment of the requirements for obtaining the degree of
MASTER OF ARTS IN DEVELOPMENT STUDIES

Major:

Economics of Development

Members of the Examining Committee:

Matthias Rieger
Trong-Hoai Nguyen

The Hague, The Netherlands
November 2021

Disclaimer:

This document represents part of the author's study programme while at the International Institute of Social Studies. The views stated therein are those of the author and not necessarily those of the Institute.

Inquiries:

International Institute of Social Studies
P.O. Box 29776
2502 LT The Hague
The Netherlands

t: +31 70 426 0460
e: info@iss.nl
w: www.iss.nl
fb: http://www.facebook.com/iss.nl
twitter: @issnl

Location:

Kortenaerkade 12
2518 AX The Hague
The Netherlands

Contents

List of Tables

List of Figures

List of Acronyms

BPP	Billion Price Project
CPI	Consumer Price Index
GSO	Government Statistics Office (of Vietnam)
ONS	Office for National Statistics (of the UK)
OPI	Online Price Index
ILO	International Labour Office
CPG	Consumer Packaged Goods
SKU	Stock Keeping Unit
HCMC	Ho Chi Minh City
MOIT	Ministry of Industry and Trade (of Vietnam)
URL	Uniform resource locator
COVID-19	Coronavirus disease 2019
JHU CSSE	Center for Systems Science and Engineering at Johns Hopkins University
VIF	Variance Inflation Factor

Acknowledgment

This research paper is the result of a learning journey that I feel fortunate to receive so much support and help along the way. First and foremost, I would like to express my sincerest gratitude to my supervisors. To Professor Matthias Rieger for always challenging me to strive for research excellence and being there when I need analysis advice, and to Professor Hoai Nguyen, for his critical feedback and guides to better structure my paper. At the same time, this research could not have been done without the help from young and enthusiastic intellectuals that helped me a great deal in building the web scraper tool: to Hieu Nguyen and Linh Bui from Vietnam Institute for Economic and Policy Research (VEPR), who kindly shared the codebase to the scrapers used in their study back in 2019 so I saved a substantial amount of time not having to build our own scraper from scratch; and to my technical assistant Ly Nguyen, who developed the tool all by herself, and always worked hard to make sure the data got collected as expected despite so many changes during this study.

In addition, I would like to extend my sincere appreciation to my family and my dog for their patience and understanding. Nothing I wish more than now is that the pandemic will be contained soon so I can go home to see them. Finally, I could not have completed this paper without my friends' continuous support and cheers, Ngoc Tran and Joost Peters, which gave me the mental strength to go forward in this journey.

Abstract

This thesis examines the potential of using online pricing data from multi-channel retailers for economic research during the pandemic time. It describes the method of collecting data through web-scraping techniques in two of the largest retailers in Vietnam during its first "true" wave of COVID-19 in 2021 and discusses the benefits and challenges of this approach. Data were collected daily across 167 days and from 2,398 product items, for a total of 396,335 observations. Despite limitations such as the short time frame of the research and the fluctuations in the number of data points gathered during the lockdown, the thesis shows that the online price index is capable of tracking the inflation dynamics during the pandemic. The approach can be helpful when price data cannot be collected in person due to lockdowns. Regressions of price dynamics on pandemic variables indicate that the pandemic trajectory, including the total number of vaccinations and the lockdown measures, correlates strongly with the discounting benefits consumers can enjoy, particularly on essential products like foods. Still, no evidence for the correlation between pandemic variables and inflation has been found within the scope of this research.

Keywords

Inflation; online price index; e-commerce; multi-channel retailers; pricing behavior; pandemic; Vietnam.

Chapter 1
Introduction

The global COVID-19 pandemic started in early 2020, yet until now it still has profound impacts on the world in many different aspects. Particularly in Vietnam, where its people only experienced the "true" wave of the pandemic since May 2021 (Tan, 2021), the pandemic and three-month complete lockdown in its two key cities has posed huge challenges not only for its citizens and businesses but also for policy-makers. The confrontation the policy-makers face is not limited to adjusting measures to cope with the pandemic but also about managing resources and tracking economic indicators. One indicator that has not been easy to collect and measure during the pandemic in Vietnam is inflation, due to difficulties in collecting data through traditional price surveys. At the same time, timely information on the supply and demand of consumer goods – reflected through price movements as per neoclassical microeconomics theories – can help the government manage resources better and ensure access equality by regulating the distribution of goods.

To address the uncertainty and inefficiency of traditional consumption price surveys, recent literature has suggested using an alternative source of data – online price data from e-commerce websites collected via web-scraping technique. In an attempt to help tackle the aforementioned problems posted to policy-makers during crisis time, this paper attempts to examine the possibility of using online prices from multi-channel retailers to track inflation during the COVID-19 pandemic. It also assesses the short-term impact of the pandemic on the consumer price level.

1.1 An alternative of price data source for economic research

Inflation has been an economic indicator of popular interest, as it reflects the status of a country's economy (Barro, 2013) and plays a critical factor in the planning and decision-making process in government bodies, firms, and individuals. One of the most well-known estimators of inflation is the Consumer Price Index (CPI), which is the indexed number that "measure changes over time in the general level of prices of goods and services that households acquire (use or pay for) for the purpose of consumption" as defined by ILO *et al.* (2020, p.1). Officially, there are two basic methods to collect price data: (1) local collection of consumption in stores, vendors, and markets, and/or (2) central collection, typically for goods and services with "centrally regulated or centrally fixed prices", which can be obtained directly to statistical head office (ILO *et al.,* 2020, p.12-13). While centrally collected price data is limited to only a few regulated goods and services such as healthcare or transportation, most products in the CPI basket are still collected locally with a survey process involving elaborate logistics and effort in fieldwork management, quality control, and validation (ILO *et al.,* 2020, p.86-95). The most common frequency of a CPI survey in this collection method is monthly, although some Oceanic countries and small economies collect quarterly or annually (Moody's Analytics, 2021).

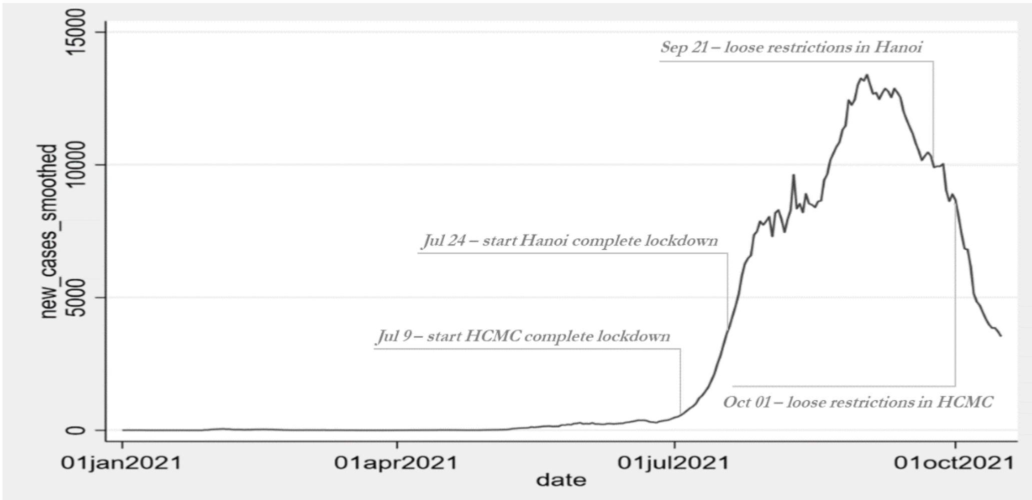
As we enter the twenty-first century – the technology era, Groves (2011) highlighted that using data from the Internet may lessen fieldwork burdens. The research process for price collection, unsurprisingly, should not be an exception. Many studies in the past decade have suggested an innovative substitute for traditional price collection methods: scraping price data from e-commerce platforms and constructing an Online Price Index (OPI) based on that data (Cavallo, 2012; Cavallo and Rigobon, 2016; Griffioen, de Haan and Willenborg, 2014; Nygaard, 2015; Breton, Swiel and O'Neil, 2015). Though this method has inherent challenges such as representativeness issues and a higher probability of missing data, the enormous advantages could compensate for that. OPI is a quick and low-cost method, which can be done at a higher frequency to reflect the change in the price level. Besides, this method can solve CPI's measurement problems as it publicizes detailed information on all products being sold by the retailers. This thesis argues that OPI's advantages become pertinent in such times of crisis like the ongoing COVID-19 pandemic.

1.2 The potentials of pandemic research with online price data in Vietnam

Vietnam is not an exception when it comes to inflation estimation using the CPI method. The Vietnamese CPI is measured with a national price survey conducted by the General Statistics Office (GSO) on a monthly basis. This practice incurs a high cost for human resources and logistics but does not yet meet the demand of policy-makers and firms for faster and more accurate data to make timely decisions (Nguyen, D. H. and Bui, 2019). Notably, during its latest strict lockdown restrictions since July 2021, the inflexibility of this traditional data collection method has become more apparent. Since May 2021, the country has struggled to contain the spread of the Delta-variant *(Figure 1)* and was forced to bring its key cities under complete social distancing, which requires almost all citizens to stay home unless it is essential to go outside (VietnamPlus, 2021). Households were assigned tickets to do grocery shopping at a designated time each week to ensure physical distancing at typically crowded places like fresh markets or supermarkets. Inter-district traveling was only allowed with a permit from the authority. The peak came around late August and early September when the government had to deploy military troops to enforce the strictest measures in Ho Chi Minh City (HCMC). During this time, no citizen was allowed to leave home for 14 days, even for food and grocery shopping (Reuters, 2021). As a result, activities like traditional fieldwork to collect price data encountered impassable barriers. A GSO price survey requires data collection at physical shopping points, including both traditional trade channels (such as wet markets, grocery stores, etc.) and modern trade channels (like supermarkets, convenience stores, etc.) (GSO, 2019, p.9-10). Hence, not only did the traveling restrictions hinder this data collection practice, but the physical distancing and vendors' willingness to support data collection might also be a problem. The matter was particularly challenging for traditional trade selling points, where most products do not have a price tag and the selling areas are commonly small. Consequently, the traditional trade vendors might refuse to cooperate to minimize the interaction with the price researchers. Before this wave, the national CPI used to be published monthly on the GSO's official website with detailed insights into the movement of each product category in the consumer basket. Nonetheless, by September 2021, the most recent announcement was from 29th Jun 2021 as a combined update of the first half of 2021, but not for each month separately. Under these difficult circumstances, online price data emerges as the top easy-access source when the price change is under research interest.

In addition to that, the development of the Vietnamese e-commerce landscape also offers an excellent opportunity for using online price data. Vietnamese e-commerce market ranked as the 28th largest e-commerce market in the world, behind the Finnish market but ahead of the Danish (Statista as cited by eCommerceDB, 2021). Its total e-commerce revenue is nearly 6.5 billion USD in 2020, with an estimate of 12% compound annual growth rate for the period of 2020-2024. Deloitte Vietnam (2019) attributes this digital retailing boom to the fast-growing middle class, representing the expanding population of tech-savvy Vietnamese youths. Furthermore, since the start of the COVID-19 pandemic in 2020, together with occasional lockdown restrictions, the demand for digital channels for shopping has been increasing significantly (Nielsen Vietnam, 2020; Kantar Worldpanel, 2020). The players in the Vietnamese e-commerce sector extend beyond online-only retailers like Shopee, Lazada, Tiki, and also include modern trade retailers. In 2021, most of the key modern trade retailers in Vietnam have had their own website or mobile app, or third-party mobile app that allows online product ordering, namely Coopmart, Vinmart, Big C, etc. This indicates an abundance of opportunities to use their price data to experiment with the OPI method.

Figure 1. Number of COVID-19 cases in Vietnam and lockdown timeline



Source: JHU CSSE (Dong, Du and Gardner, 2020)

Motivated by the usage of new data sources to address the uncertainty and challenge of traditional price research during crisis times, this study attempts to use online price data from multi-channel retailers in Vietnam to (1) track inflation via consumption price changes during the COVID-19 pandemic and to (2) anticipate crisis consumer demand and supply through the short-term impact of the pandemic on the consumer price.

This study makes three main contributions to the literature. Firstly, it contributes to the discussion of using online prices to track inflation in Vietnam. By the time this research was conducted, there was only one existing OPI research by Nguyen, D. H. and Bui (2019), but with potential representativeness bias because of the data stemming from online-only retailers. By focusing on the data source from multi-channel retailers, which represent a more significant share of Vietnamese consumer usage[1], this research has the potential to reflect a price movement closer to the real world. Secondly, it contributes to the general literature of using online prices to track inflation but from the perspective of the pandemic. Though Cavallo has initiated online data usage for inflation tracking and inspired various researchers to explore this approach over the last ten years, Alvarez and Lein's study (2020) were presumably the only research conducted during the coronavirus health crisis up to now. While their study covered the first lockdown phase in Switzerland when the information about the virus was still novel, this study, however, will provide insights into the construction of online price index under a different scenario – Vietnamese second nationwide lockdown, yet the most severe lockdown ever in a country used to be praised on its excellent containment of the virus in 2020 (Pollack *et al.,* 2021; Grant, 2021). Thirdly, this paper also makes the first empirical contribution of the short-term impact of the pandemic on consumer prices in Vietnam. There have been studies into how price changes during the pandemic, such as Balleer *et al.*'s research (2020) with German firms, which revealed supply and demand shocks on the

[1] To be discussed in more detail at Subsection *2.1 Tracking inflation with online prices*

pricing decision, or Alvarez and Lein's (2020), which also highlighted the Norwegian online price index reduction after the lockdown. However, to the best of my knowledge, all these studies were mainly qualitative and observation-based and did not yet go the distance to identify to what extent the pandemic impact on the price.

This research paper is structured as follows: Chapter 2 reviews the literature on the conceptualization of the OPI method and the relationship between the pandemic and the consumer price; Chapter 3 describes the methodology and the research process; Chapter 4 presents and discusses analysis findings, and finally, Chapter 5 concludes and considers the implications for future research.

Chapter 2
Literature review

2.1 Tracking inflation with online prices

The review of the empirical literature using online prices to construct price index aims to discuss the differentiation in the OPI strategies and its meaning on the context and results, based on which to shape the suitable strategy for this research.

When it comes to the OPI literature, Alberto Cavallo is the one who pioneered and has made the most contribution. Also motivated by the uncertainty of official CPI statistics, where the Argentinian government was under public suspicion of CPI manipulation during 2007 and 2015, Cavallo started to collect daily online prices from the largest supermarket across Latin America to compare and reflect the price change level from 2008 onwards (Cavallo, 2012; Cavallo and Rigobon, 2016). This research later demonstrated the price change level for online and offline data was almost consistent in seven Latin American countries except for Argentina. Official CPI figures from the Argentinian government show only an 8% annual inflation rate during 2007 – 2011 compared to 20% rate derived from Cavallo's research. Surprisingly, Cavallo's result was more in line with household expectations and estimation from domestic economists (Cavallo, 2012). Cavallo's works have motivated and inspired the usage of big data to reflect timelier and more efficient price movement. A number of studies followed his method to (1) extend the geographic and product coverage for the OPI method experiment and (2) evaluate the performance of this method against traditional CPI in more detail. The former purpose attracted both official government agencies like those from the UK, Brazil, the Netherlands, and Norway (Breton, Swiel and O'Neil, 2015; Nygaard, 2015; Griffioen, de Haan and Willenborg, 2014; da Silva *et al.,* 2019) as well as independent researchers (Banerjee, Singhal and Subramanian, 2018; Nguyen, D. H. and Bui, 2019; Alvarez and Lein, 2020). These works indicated mixed results and implications, which could be attributed to their different data collection strategies and will be further discussed in the next paragraph. Meanwhile, the latter purpose was well addressed in studies from Antonakakis *et al.* (2016), Harchaoui and Janssen (2018), Aparicio and Bertolotto (2020). These studies all suggested superior characteristics of the OPI method over the traditional CPI method in terms of less inflation persistence and higher timeliness. These advantages become even more critical during the COVID-19 pandemic, when most economies are under lockdown to some different extent, making it difficult to do traditional price collection (Alvarez and Lein, 2020; Cavallo, 2020). *This* argument could apply to Vietnam, where though the e-commerce market is enormous, but it takes up only around 3.8% of the whole retailing market. The majority of retail transactions are still taking place through offline channels. In the Vietnamese offline retail sector, while traditional trade channels are still dominant, a quick expansion trend is observed among modern trade channels both in terms of market transaction volume and value (Deloitte Vietnam, 2019). Deloitte Vietnam's annual consumer survey showed an increasing shopping preference at modern trade channels among urban consumers, from 44% votes for modern trade in 2018 to 57% in 2019 (2019, p.22). Among the choice for modern trade channels, the choice for supermarkets emerges as the majority. Regarding the debate between choosing online-only or multi-channel retailers for the OPI in the Vietnamese context, apparently, the latter outweighs the former includes the summary of these key empirical studies constructing and working around the OPI. Though all aforementioned studies used the price data scraped from online websites, the strategies upon which retailers, categories, and products to extract data, varied from one study to another.

Table 1. Major literature constructing and working around the online price index

#	Author	Year	Coverage	Retailer type	Product category
1	Cavallo	2012	5 Latin American countries	Multichannel	Food, CPG
2	Cavallo and Rigobon	2016	25 countries	Multichannel	Food, CPG
3	Griffioen *et al.**	2014	Netherlands	Multichannel	Clothing
4	Nygaard*	2015	Norway	Online-only and multichannel	Personal care
5	Breton *et al.**	2015	UK	Multichannel	Food
6	Antonakakis *et al.*	2016	8 countries (Argentina, Brazil, China, Japan, Germany, South Africa, UK, US)	Multichannel (using data from BPP)	Food, CPG
7	Banerjee *et al.**	2018	India	Online-only	Food
8	Harchaoui and Janssen	2018	The US	Multichannel (using data from BPP)	Food, CPG
9	da Silva *et al.**	2019	Brazil	Online-only and multichannel	Airfare and consumer electronics
10	Nguyen and Bui	2019	Vietnam	Online-only	Food, clothing, household appliance, entertainment
11	Alvarez and Lein	2020	Switzerland	Multichannel	Food, clothing, electronics, furniture, heating oil
12	Aparicio and Bertolotto	2020	10 countries (Australia, Canada, US and 7 European countries)	Multichannel (using data from BPP)	Food, CPG

** Authors of the study are price statisticians / statistical agencies*

Firstly, when it comes to choosing which retailers to scrape data from, the retailers' selling channel is an interesting aspect though it seems not to be a topic of debate. In his first official study about the OPI for Latin American countries, Cavallo (2012) did not clearly explain why he chose supermarkets' websites, i.e. multi-channel retailers. Other follow-up studies show mixed choices of online-only and multi-channel retailers without a clear rationale, so it seems that it might be based on the availability of the data. Statistics Norway's study, in which they scraped the data from both types, emphasized different price movements and pricing strategies between these two retailers (Nygaard, 2015), so it is crucial to decide upon this matter. Later in their paper to introduce the Billion Price Projects (BPP), which collect online price data on a large scale from more than 50 countries, Cavallo and Rigobon (2016) argued upon their choice of multi-channel retailers that though the price data is scraped online, it is also the products they sell at physical stores with about 72% with the same price (Cavallo, 2017). Hence, pricing data from multi-channel retailers make up a more representative of retail transactions than online-only ones. This argument could apply to Vietnam, where though the e-commerce market is enormous, but it takes up only around

3.8% of the whole retailing market.[2] The majority of retail transactions are still taking place through offline channels. In the Vietnamese offline retail sector, while traditional trade channels are still dominant, a quick expansion trend is observed among modern trade channels both in terms of market transaction volume and value (Deloitte Vietnam, 2019). Deloitte Vietnam's annual consumer survey showed an increasing shopping preference at modern trade channels among urban consumers, from 44% votes for modern trade in 2018 to 57% in 2019 (2019, p.22). Among the choice for modern trade channels, the choice for supermarkets emerges as the majority. Regarding the debate between choosing online-only or multi-channel retailers for the OPI in the Vietnamese context, apparently, the latter outweighs the former regarding representativeness.

Nguyen, D. H. and Bui (2019) were the first researchers experimenting with constructing the OPI for the Vietnam market. Still, their study mainly focused on online-only retailers, probably because modern trade retailers did not invest much into their online channel yet back in 2019 when their study started. This choice of retailers might contribute to their final results, which only show robust synchronization with the official CPI index in food-related services, but insignificant correlation for other categories. Nonetheless, as the boom of digitalized multi-channel retailers has emerged in the past year, it is inarguably a good time to reconstruct this index for Vietnam. While Coopmart, Big C, and Vinmart are the biggest retailers in Vietnam with 4.9%, 4.2%, and 1.9% respectively out of the total CPG value market share (Kantar Worldpanel, 2020), only Coopmart and Vinmart have built their own e-commerce website and have the ability to control their online selling channel autonomously. Hence, this research uses online pricing data from the website of Coopmart, which represents the Southern market of Vietnam, and Vinmart represents the North.

Secondly, the choice of product categories seems to depend a great deal on the researchers' objectives, which appear to be quite distinct between statistical agencies and independent researchers. For price statisticians, their primary objective is to improve the efficiency and accuracy of the index by increasing the e-commerce share into the basket, so they tend to experiment with categories of specific characteristics first to explore the potential. Statistics Netherlands experimented with the clothing industry because this category is considered "hard to measure" due to its seasonality and fashion effects (Griffioen, de Haan and Willenborg, 2014, p.2), while Stats New Zealand approached the consumer electrics, cars, and rent category (Bentley and Krsinich, 2017). Quite the contrary, Statistics Norway chose personal care products as these products are "homogenous and long-lived" (Nygaard, 2015, p.10). On the other hand, independent researchers' objective for constructing the OPI is to build a timelier and more convenient way to capture the dynamism of inflation, based on which they can conduct other macro studies further. Hence, they tend to choose retailers that cover the majority of consumer expenses like grocery and CPGs to begin with (Cavallo, 2012; Banerjee, Singhal and Subramanian, 2018; Nguyen, D. H. and Bui, 2019; Cavallo and Rigobon, 2016). As the main objective of this research is closer to the latter group, the products of this research's target will also represent a big portion in the national CPI basket, namely grocery and CPGs. The products used in our data are from four out of eight main categories specified by GSO, which include (1) Food - where this study only cover two subcategories as Dried food and Other food, (2) Beverage, (3) Household products and (4) Others products & services.

For the purpose to create an alternative way to quickly proxy inflation, most independent researchers validated the efficiency of their OPI by comparing its price movement with the country's traditional CPI (Cavallo and Rigobon, 2016; Cavallo, 2012; Nguyen, D.

[2] Total value of Vietnamese retailing market is estimated at 170 billion USD (Mordor Intelligence, 2021) while the value of its e-commerce market is at 6.5 billion USD (Statista as cited by eCommerceDB, 2021)

H. and Bui, 2019; Banerjee, Singhal and Subramanian, 2018; Aparicio and Bertolotto, 2020; Harchaoui and Janssen, 2018). The traditional CPI is considered the benchmark for OPI, though it is not entirely correct all the time, for example the case of the Argentinian government's manipulation during 2007 – 2015 (Cavallo, 2012). The comparison of the OPI against CPI in Cavallo's work (2012) mainly was to confirm the doubt on the Argentinian case when compared with other Latin countries. However, the context in Vietnam is different, when the uncertainty on the CPI results is objective as a result of the pandemic impact as discussed in subsection *1.2 The potentials of pandemic research with online price data in Vietnam.* In addition, the comparison is also barely possible due to the lack of data, as nearly six months of data collection for this OPI research allows for only two data points for quarterly updated CPI. Simultaneously, there is also the risk of consumer basket changing during the lockdown with the evidence from multi-country studies (Cavallo, 2020; Gautier, Ulgazi and Vertier, 2020), as consumers tend to pantry stock goods for their "basic needs" when the future is uncertain (Sheth, 2020, p.281). Hence, looking at the movement of the price index at the categorical level is more meaningful than aggregated price index during this particular time. For this research paper, the OPI result will be analyzed at the category level and will not be compared with the CPI published by GSO. Instead, the empirical result will reflect the qualitative insights collected from news and social listening during the data collection time.

2.2 The conceptualization of the factor model impacting the consumer price

Though the topic of how the pandemic impacts the consumer price has caught the attention of and has been widely covered by government researchers, such as those from the US Bureau of Labour (Mead, 2020) or Vietnam Directorate of Market Surveillance (Luu and Tran, 2021), all of them only contributed in terms of observations about price fluctuations during the pandemic, none have managed to quantify the size of this impact. One of the main reasons could be the lack of a dynamic data source representing the price movement. To build the conceptual model of factors impacting the consumer price, this subsection will firstly review the classical theory determining the market price and observations to hypothesize the relationship between the pandemic and consumer price; and then, based on additional empirical evidence to identify other pricing factors.

2.2.1 Classical economic perspective about market price determination

The most well-known price theory is proposed by classical economists, which theorizes that market price for a perfectly competitive and not regulated product is determined by the equilibrium point of its supply and demand curve (Pindyck, Rubinfeld and Rabasco, 2013). According to this theory, any market elements that impact the customer demand or producer supply will shift either or both curves, which identify a new equilibrium point for the price and the quantity sold for the product. The COVID-19 pandemic has been observed to notably impact both the supply and demand for consumer goods in various ways, through which the price of goods can be changed accordingly.

When it comes to demand, it seems that the COVID-19 pandemic increases the demand for certain consumer products but only in the short term. At the start of the pandemic, it was all over Western media about irrational consumer behavior about hoarding and stockpiling for products considered "essential". NielsenIQ (2020) reports a sudden global rise in sales for medical supplies such as rubbing alcohol, antibacterial wipes, antiseptics, facemasks, etc. In addition to healthcare products, reports from Accenture and McKinsey show the same increasing trend in the consumption of non-perishable food and personal hygiene products like toilet paper (Arora *et al.,* 2020; Wright and Blackburn, 2020). Sheth (2020, p.281) called these hoarding and stockpiling behaviors as "a common reaction to managing the uncertainty of the future supply of products for basic needs". However, which products are considered to be essential might vary based on consumers' culture and living style. For Vietnamese specifically, necessities could be healthcare supplies and non-perishable food like instant noodles, as these products witnessed an abnormal increase in consumption spending here in 2020, during the first wave of the pandemic (Kantar Worldpanel, 2020; Nielsen Vietnam, 2020). Simultaneously, as the economic conditions and lifestyle change, the necessity of certain product types could change accordingly and impact the demand elasticity (Bijmolt, Van Heerde and Pieters, 2005). For example, non-perishable food might not be considered essential before the pandemic because fresh food is abundant and usually considered "better natural quality" in Vietnam. The same thing could be true for medical supplies like facemasks or alcohol rub. Nonetheless, under the lockdown context, this perception can change. The lockdown comes with the possibility of less frequent visits to the grocery store, which is synonymous with more groceries should be bought in one trip, and food bought should be kept longer in the pantry/fridge. This theoretically boosts the necessity of non-perishable food and makes its demand less elastic, while doing the opposite for products of less importance to survive the lockdown like beverages or household products.

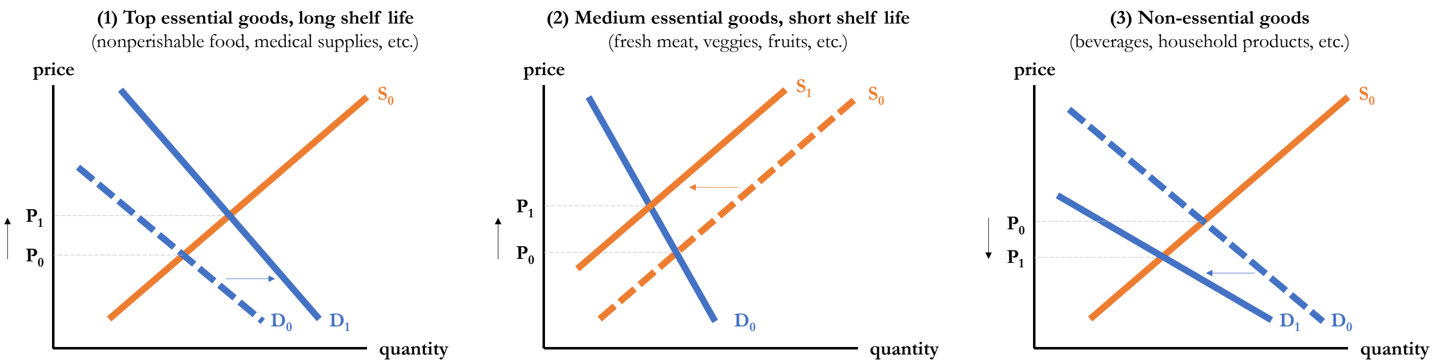
In addition, the pandemic might also impact the demand for consumer goods in a different direction regarding consumers' indispensable income. Widnyana and Widyawati (2020) demonstrated the empirical evidence from Bali that the pandemic decreased the consumption per capita for consumers goods significantly after the first wave of the pandemic. The pandemic, together with company closing and business abruption, put pressure on the citizens' employment status, income, and purchasing power (Blustein *et al.,* 2020; Joshi, Bhaskar and Gupta, 2020), accordingly forcing them to manage their spending in a "priority scale of needs" (Widnyana and Widyawati, 2020, p.465). This could be the same case for Vietnam, when an online survey conducted in July 2021 by Statista and InFocus indicated that 67% of the respondents got impacted by the pandemic, either to suffer a job loss or a temporary salary reduction (Nguyen, M. N., 2021). With the demand for certain essential consumer products on the rise while the purchasing power is diminishing, it is expected that demand for non-essential products might drop off.

Meanwhile, the supply side is mainly under the pandemic measures' oppression. The COVID-19 pandemic is reported to interrupt the whole supply chain for consumer goods through the physical lockdown that restricts laborers' movement, close down the production facilities, and disrupt the distribution logistics all over the world (Aday and Aday, 2020; Mahajan, 2020; Widnyana and Widyawati, 2020). India and Bali witnessed a drop in supplies for FMCG products after its first COVID-19 wave (Widnyana and Widyawati, 2020; Mahajan, 2020). The issue of labor shortage might not impact too much the supply for consumer goods in Vietnam's case, as policy-makers prioritized measures to ensure continuous economic activities by suggesting alternative labor arrangement models to businesses, namely "3 tại chỗ" (3 activities at the same location: employees work, eat, and rest in one place) and/or " 1 cung đường – 2 địa điểm" (1 route – 2 locations: employees travel on only one route connecting their home and their workplace) (MOIT, 2021). Nonetheless, in another aspect, during the peak of the pandemic around August – September 2021, authorities in Hanoi and HCMC disrupted severely the last stage of the supply chain – the distribution of goods – with measures restricting traveling and inter-district delivery. This policy affected the most on short shelf-life products, especially fresh groceries (VnExpress, 2021), because of their nature requiring higher frequency to refill the inventory. The supply for longer shelf-life products should not expect the immediate impact from the pandemic but could experience it in the longer term if the lockdown measures remain the same.

In summary, the impact of the pandemic on supply and demand for consumer goods depends substantially on which category that goods belong to. Under the classical economic perspective, goods that are considered essential like health supplies or non-perishable food (rice, noodles, canned food, etc.) will expect a short-term surge in their prices. This is the consequence of anticipated higher and less elastic consumer demand following the increasing uncertainty about the future. Prices for goods of short shelf-life and medium necessity level like fresh food (vegetables, fruit, meat, etc.) also expect to rise due to the limited supply. Last but not least, goods like beverages or household products might see a price reduction mainly due to the more elastic and decreasing demand from the consumers' limited purchasing power and lower level of necessity in consumers' perceptions. These hypotheses are visualized and can be found in *Figure 2*.

However, it is worth highlighting the assumptions used by this price theory, which might not always be accurate in practice. Firstly, the market following the hypotheses above assumes there is no authority intervention. In a study on pricing at French supermarket chains, Berardi *et al.* (2017) found that though local demand and competition do have a role in influencing prices but only at a small scale, while the major factor is "retail groups' bargaining power" at a national level. Besides, if any ceiling or floor price is applied, price should not fluctuate when supply or demand change, but instead might cause the scarcity or abundance of resources. Though information on any ceiling/floor price for consumer goods in Vietnam is unknown, a government attempt to supervise and stabilize prices during the pandemic is recorded to ensure essential goods are affordable and accessible for its citizens (MOIT as reported by Chau, 2021). While to what extent the Vietnamese government influences the pricing of consumer goods is yet to be known, it is undoubtedly a factor that can hinder the price fluctuations during the pandemic. Secondly, this theory assumes the market is perfectly competitive, while most consumer products nowadays, packaged and branded, belong to the oligopolistic market where product quality and branding from both manufacturers and retailers also play a role in determining the price. Last but not least, Morgenroth (1964, p.17) pointed out that a critical deficiency of this theory is that it assumes that the market knowledge of businesses is complete, which might not always be the case. This is especially true during crisis times when the future is unpredictable, and there is no relevant historical data to support business decisions going forwards.

Figure 2. Hypotheses of the pandemic's impact on the price of consumer products



Note: the level of necessity here is defined under the pandemic context only

2.2.2 Factors impacting consumer price from empirical evidence

Even though the classical price theory is a meaningful model that helps understand how supply and demand can impact the price, it still has specific deficiencies as previously mentioned, and subsequently, is insufficient to explain the dynamism of consumer price change. To achieve a more thorough understanding of various factors that may influence the consumer price change, this section is dedicated to discussing empirical research on pricing determinants in modern trade retailers, and seeing where it is relevant to include such factors into our pricing model.

The research by Shankar and Bolton (2004) was the first to use scanner pricing data of six types of products[3] in supermarkets across five cities in the US to determine the factors impacting their pricing. Their results indicated that six factors reflecting (1) competitiveness, (2) product category, (3) chain characteristics, (4) store characteristics, (5) branding, and (6) customer characteristics, were the most statistically significant elements explaining the variance in supermarket pricing. Later, Ellickson and Misra's study (2008), with multi-sources of data from store managers' reports to census data on supermarkets and customer demography, also confirmed the importance of these factors except for the branding and category factors, but only due to the limitations of their data source. These studies help to shape the conceptual framework to build the price change model and confirm the previous reasoning under the classical economic theory. Suppose all things remain the same except for the severity of the pandemic, the category factor[4] and customer factor, under the health crisis impact, is linked directly to the anticipated supply and demand for consumer goods as discussed under subsection *2.2.1 Classical economic perspective about market price determination*. The considerations could be tricky for the chain factor, i.e. how the retailer positioning impacts the pricing strategies. Indeed, Coopmart and Vinmart target slightly different income segments that Vinmart's price tends to be higher than Coopmart's. Still, in this research, they represent two distinct geographical markets with different situations of the pandemic. The variable defining each retailer will be included in the model, but its impact does not solely reflect chain positioning but might also interact with customer preferences or different pandemic measures in each region.

On the other hand, there will be certain limitations for this research paper as there are a few elements we cannot include in our pricing model. The competitor factor, as explained by Shankar and Bolton (2004), connotes the competitor price and competitor's deal frequency. However, due to the limitation of the research coverage, which experiments with only two retailers and assumes each retailer represents one region in Vietnam. Therefore, though evidence indicated competitor factor is a paramount element in supermarket pricing strategies (Shankar and Bolton, 2004), it is hard to measure within the context of this study and therefore will be excluded from the consumer pricing model. In addition, while brand factor also plays a role in pricing and the branding information could be extracted from the product name in theory, hundreds of different brands of the products will complicate the model in practice. Therefore, for the simplicity of the price model, this research paper will opt out this factor. Last but not least, the store factor, which reflects the store size and assortment, might not be a critical element for the case of this research's context. Coopmart's pricing policy is reported to be the same for all stores since the price is decided by the Union (Nguyen, H. T. *et al.,* 2020). For Vinmart, the pricing policy is yet known to us, but after initial checking and comparing three random stores of Vinmart in Hanoi, we found about 100% of the products share the same price among these stores. The data collector used in this study only allows to scrape from one store at one specific location from each of these

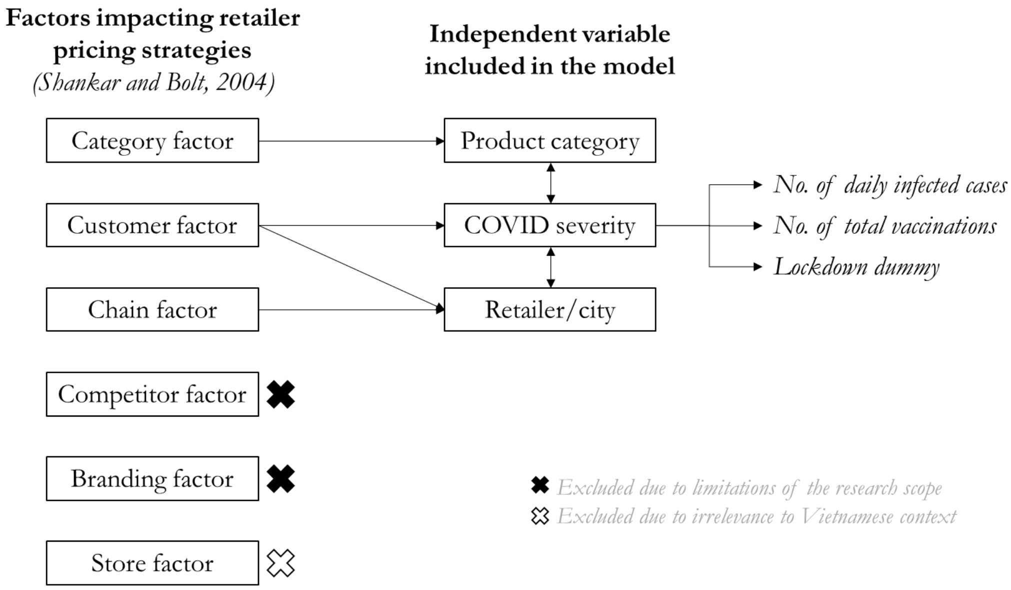
[3] The products used for this research were "spaghetti sauce, bathroom tissue, liquid bleach, ketchup, mouthwash, and frozen waffles" (Shankar and Bolton, 2004, p.29)

[4] The category factor is defined as the storability and necessity of the product (Shankar and Bolton, 2004)

two retailers, but based on their price consistency, it seems that the choice of the store does not have any impact on the pricing level.

While almost all studies on pricing factors are rooted in the Western market, this paper will contribute the first empirical evidence from an emerging market during the pandemic. The conceptual framework of the price model used in this research to explain the price change can be found in *Figure 3*. The explanatory variables include product category, retailer (Coopmart or Vinmart, but also reflect city HCMC or Hanoi), and COVID statistics data to reflect the pandemic severity. Three types of data will be used to proxy the pandemic, namely the number of daily infected cases, the number of total vaccinations from the real-

Figure 3. The conceptual model of factors explaining the price change in this study



time dataset from the Center for Systems Science and Engineering at Johns Hopkins University (JHU CSSE) (Dong, Du and Gardner, 2020), and the lockdown dummy which indicates the days that the city is in strict lockdown following the government directive. However, it is highly likely to exist the multicollinearity between these three variables as vaccinations can substantially help to reduce the infection rate by providing direct protection against the virus and reducing transmission (Lipsitch and Dean, 2020), and as the number of daily infections goes down together with more vaccinated citizens, the government can have the confidence to loosen the lockdown. Therefore, this model will only experiment separately with each out of these three pandemic variables.

Last but not least, there is also a high potential for endogeneity in this model due to the omission of variables for government intervention as well as competitor and branding factors. This problem of endogeneity emerges when excluded variable(s) is correlated with existing variables and, accordingly, might misspecify the model, bias the beta-coefficients, and distort the causal inference (Roberts and Whited, 2013, p.494-499; Wilms *et al.,* 2021, p.3-5). In this research, if there is little or no relationship between government intervention, competitor factors or brand factors with the other three exploratory variables of the model, the effect of those omitted variables estimated will be entirely absorbed into the error terms, the beta-coefficients are consistently estimated, and endogeneity should not be a concern. However, it would be a problem if there exists any correlation. Let's take excluded competitor factor and included the pandemic severity variable as an example: if the more severe the pandemic becomes, other competitors raise the price high to respond to the demand and supply shocks, and because of this, the retailers of interest might have more motivation to increase the price. In this case, if the competitor factor is not presented, the coefficients for pandemic severity might exaggerate its actual impact on the consumer price change. In reality, Bach Hoa Xanh – another big retailer in the Southern market – did increase their price for some perishable food items to a great extent during the peak of the recent wave, which faced a wave of public criticism (Nguyen, Q., 2021). Commenting on this news, both Coopmart and Vinmart promised that they would not increase the price [on top of the manufacturers' price?] during this challenging time (Vietnambiz, 2021). However, it is still uncertain if this claim was the retailers' real commitment or only a marketing trick to boost their brand image while their prices still increase slowly because of the competitor factor. If the latter is the case, endogeneity might be expected, though at a small scale because of public effort in consumer price stabilization. Besides this example, there are still other possibilities for correlation that might be unknown between these omitted variables and the included variables in our model, so the endogeneity problem is worth keeping in mind when we interpret the results for any causal inference later on.

Chapter 3
Methodology and data collection

As this research focuses on employing the innovative source of online data to answer two questions on price research during the pandemic, the first section of this chapter will explain the process, how the data is collected and validated. The second section describes how the online price index is computed, and the last section describes how our pricing model is constructed.

3.1 The data

3.1.1 Data collection

This research uses the web scraping method to collect the data from supermarket websites because of its easy accessibility and real-time price update. This scraping technique could be performed with different tools, but we used Python Selenium and BeautifulSoup library because of its available open-source. The web scrapers' mechanism is basically guiding the tool to log into the predefined website every day, browsing through pages of products from categories to categories, screening through all available products, then recording all the data which matches predefined variables into a data file for that specific day. The variables collected include identifier code (as defined by its unique URL), product name, discounted price, non-discounted price, category name.

The duration of the scraping is from May 1st to Oct 14th 2021, with the attempt to cover nearly the whole cycle of the biggest wave of COVID-19 in Vietnam up to this moment. The web scrapers are built to get the data from two supermarkets in Vietnam: a Coopmart store located in HCMC, and a Vinmart store situated in Hanoi. Since the structure of each website is different, we had to customize our scraping tool for each store, and if the retailers change the structure of their website, the scrapers also got to be adjusted accordingly. In June 2021, Coopmart changed their product classification structure on their website, and in September, Vinmart migrated to a new website with a new brand name (Winmart), which all took us more effort to revise the script for the scraping tool. That is also the reason why Nygaard (2015) suggested taking a cost-benefit analysis before starting to make sure the amount of data collected at the end is worth all the effort and time paid to build, test, and adjust the tool. It is also worth mentioning that during this pandemic time, we had to stop the scraping of some stores or change the store of target several times as the previously chosen ones were forced to close for virus disinfection[5], resulting in the widely varied number of products collected. This will be discussed in further detail under subsection *3.1.2 Data validation*. In addition, due to the different strategies for product naming, assigned identifier code, and listed stock-keeping units (SKUs) at these two retailers, it is almost impossible to identify and match the same product at both retailers. This limits our capacity to compare the price movement between the retailers as well as between cities for identical products.

[5] The detailed scraping timeline with which stores to be scraped could be found in *Appendix 3. Data collection at Coopmart (HCMC) supermarket* and *Appendix 4. Data collection at Vinmart (Hanoi) supermarket*

3.1.2 Data validation

Though the prominent advantage of this method is its quick and cheap turnover to collect data, the massive amount of data collected sets additional challenges for data cleaning and validation (Nygaard, 2015).

For data cleaning, we first encountered a number of products with abnormal pricing changes and may have been incorrectly uploaded from the retailer's side. We had to alert ourselves if the price change is outside the normal range (more than ±500%) and make a judgment if the change is sensical or not and correct it if needed. Most of the time, the reason is lacking some zeros in the price[6]. Meanwhile, other cases can be the change in the quantity of products sold. For example, previously, Pepsi was sold at the price for a bunch of six cans, but after a while, the retailer decided to break down and sell a single can of Pepsi at one time to make it more flexible for consumers. Below are some examples of the significant changes that we identified:

Product name	Normal price	Defective price	Judgment and solutions
Vinamilk yogurt 100g	6200 (92%)	62 (3%)	Missing zeros → multiply the price with 100
Coop Select napkin	12500 (99%)	63000 (1%)	Errored pricing once → adjust with the major price
Sting strawberry bottle	56900 (89%)	10000 (11%)	Adjust the quantity sold per order (from 1 to 6 bottles) → remove because cannot justify the price change per bottle is due to inflation or due to the split price

Secondly, due to the overlapping classification structure of the retailers, a high number of data are duplicated. For example, a frozen bag of shrimp can be classified as both "Seafood" and "Frozen products", intended to make it easier for consumers to browse according to their shopping behavior. The considerable number of duplications realized started in mid-June when Coopmart changed its website structure. Luckily, a product classified into two different categories still has only one URL. Therefore, we can identify if it is duplicated and then remove the duplicates from the data.

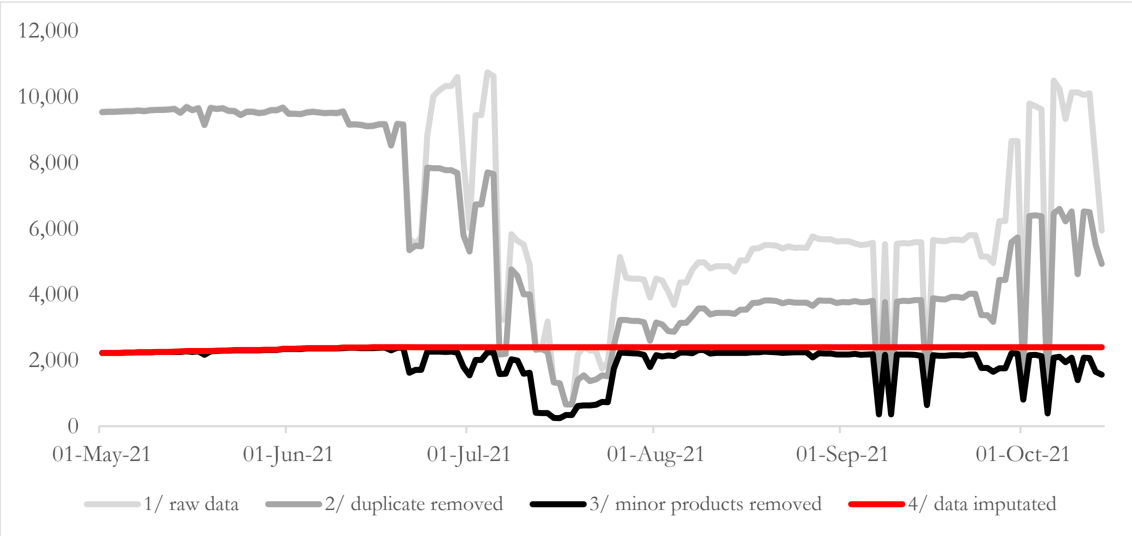
Last but not least, data missing during the scraping is inevitable and also a big challenge because of heavy reliance on technology such as website structure change, unexpected blackout, etc. (Griffioen, de Haan and Willenborg, 2014; Breton, Swiel and O'Neil, 2015) or just because of the unavailability of some specific products from time to time. It is suggested to apply imputation techniques to solve this problem (Breton, Swiel and O'Neil, 2015; Nygaard, 2015; Cavallo, 2012). Though ILO *et al.* (2020, p.16) suggests amputating data by referencing the average price change of either that product or similarly "comparable" categories, Cavallo (2012, Appendix B1) argues the most rational way to fill in the missing pieces is using its last available price because of these gaps usually last for only several days. However, the imputations done during normal times should not be the same as during crisis times. The number of products scraped every day in this study varies substantially because of the changes of lockdown measures, so before doing this imputation step, we take one more validation step: keeping only the products that are available more than 60% of the time, then

[6] Vietnamese currency is particularly large, the smallest note in use is 1,000 VND. A cup of Starbucks coffee would cost around 50,000 VND at the minimum. Missing some zeros can make a big change to the price.

removing the rest. With this approach, we can make sure the imputation is only done where necessary and expect to retain only major products, which we hypothesize as more "essential" products that the supermarkets are keen on keeping even during crisis times. However, the risk associated with this approach is that possibly the products eliminated are the products of such high demand when the pandemic hits, which supermarkets could not supply any more. This should also be taken into considerations upon reviewing the final index results.

In summary, to validate the data, we have gone through four steps to make sure the final set of data is usable and representative: (1) cleaning the data, (2) removing the duplicate values, (3) removing products that are available less than 40% of the time, and (4) imputing the missing data. *Figure 4* below shows how the number of products collected changes after each phase of data validation. Though the total number of products scraped in around 6,000 on average, the final number achieved after the validation is just above 2,000 products. It is worth pointing out that there are a considerable number of products which feature imputed prices in July because of the supermarkets' temporary closing during the lockdown, which might not correctly reflect the price index movement during this short period.

Figure 4. Total number of products collected after each data validation stage



3.2 Index computation

To build the price index for each category, we first started with getting the index at the product level, then took the geometric mean for the category. The base price used is the first day we begin this research: May 1st 2021. Particularly, we achieve the unit price index *PI* for the product *i* at time *t* by taking the quotient of its non-discounted at time *t* over its non-discounted on May 1st, 2021 (t_0):

$$PI_{i,t} = \frac{P_{i,t}}{P_{i,t_0}}$$

Next, we take the geometric mean of price index at time *t* for all products *i* in the category *j* to achieve the price movement of this category at time *t* :

$$PI_{j,t} = \sum_{i=0}^{n} (PI_{i,t})^{\frac{1}{n}}$$

Our data collected both discounted and non-discounted prices for the products, and which price to use remains a big challenge. Breton *et al.* (2015, p.3) did not consider the impact of discount in their OPI, while Nygaard (2015) only took into account the actual price sold (after discount) without giving the detailed rationale for their choice. Hence, this research will present the index from both types of price and examine how these indices respond to the general trend in Vietnam as reported on social media.

3.3 Empirical strategy for estimating the pandemic's effect on consumption prices

To understand the retailers' pricing behavior during crisis time, we use linear regressions to explore the impact of the pandemic on the unit price index. As discussed in subsection

2.2.2 Factors impacting consumer price from empirical evidence, our model includes variables reflecting the pandemic severity, plus the product category and the city/retailer where the product is scraped from. Our model is constructed as below:

$$P_{ij,t} = \beta_0 + \beta_1 Covid_t + \beta_2 Category_{ij} + \varepsilon_{ij,t}$$

with $P_{ij,t}$ is the pricing strategy for the product *i* from city/retailer *j* at time *t*
$Covid_t$ indicates the severity of the pandemic in Vietnam at time *t*
$Category_{ij}$ is the category of the product *i* from city/retailer *j*

For the pricing strategy as the dependent variable, we will look at both the original price change of the product and the discounting pattern to reflect two different types of pricing: manufacturer and retailer pricing, respectively. The original price change is employed in this model to avoid the mixed impact from retailers' unexpected promotion schemes and reflect better the retail price set by manufacturers. In this way, we expect to better understand the real pricing decisions under the impact of the pandemic, rather than such reasons as storage planning problems from the retailer. In the meantime, retailers' discount and promotion patterns reflect retailers' pricing strategies. Nonetheless, as the discount percentage with truncated data from below with the minimum value is zero, tobit regression is applied for the latter dependent variable with the lower-censoring limit of zero, while OLS regression is employed for the former. Both models will be controlled with time-fixed effect variables, including the calendar week of the year and the calendar day of the week.

The pandemic severity will be proxied through either one of the following three variables: (1) the log of daily infected cases[7] ($lnCases_t$), (2) the log of total vaccinations[8] ($lnVax_t$) – both of which are achieved from JHU CSSE's real-time database[9], and (3) the lockdown dummy ($Lockdown_t$) indicating if the city where the retailer is located under a strict lockdown or not. Since multicollinearity is highly likely among these three variables, we will experiment with three different models with these pandemic variables separately.

[7] 7-day average statistics. During the time frame of this research period, there is no zero value for number of daily cases so there is no need for data transformation to be usable for the log.
[8] There are some days that the number of total vaccinations is missing from JHU CSSE database, simply because Vietnamese Ministry of Health did not report the statistics on that day. These missing data is imputed with the nearest previous data.

Chapter 4
Results

4.1 Descriptive statistics

Our scraper tool has collected nearly one million unique data points for 13,460 products from two supermarkets in Vietnam over 167 days, from May 1st to Oct 14th 2021. After the data validation process, the final database is left with only approximately 40% of its original size, due to the substantial fluctuations in the number of products collected per day due to the pandemic and its subsequent policy (*Table 2*). Only 16.67% of the final database is imputed data according to the price data of the closest previous day.

It is also worth highlighting that the final database is highly skewed towards Coopmart in terms of contributions. The contribution split is 80% from Coopmart against only 20% from Vinmart (*Table 3*). This could result from either the different number of products sold among these retailers or their different online listing policies. No matter what is the case, the representativeness aspects should be taken into careful consideration, that the analysis results are more likely to demonstrate the trend in the south of Vietnam rather than the whole country.

Table 2. Summary statistics of the panel data size after different phases of data validation

	No. of data points	**No. of products**
Initially collected as raw data	1,132,579	13,460
After removing duplicates	942,137	13,460
After removing minor products	330,248 *(35%)*	2,398 *(17.8%)*
After imputation	396,335 *(42%)*	2,398 *(17.8%)*

Note: Figures in bracket shows the percentage out of the number of unique data points (after removing duplicates)

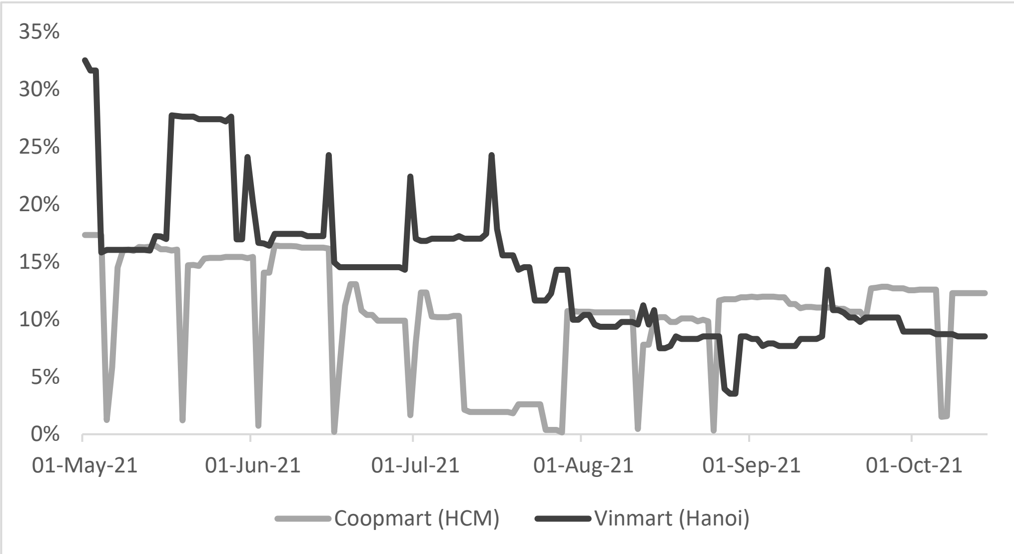
Table 3. Summary statistics of the final database by retailers and categories

Product category	**No of. data points at Coopmart (HCMC)**		**No of. data points at Vinmart (Hanoi)**		**Total**
Dried food	52,429	76%	16,922	24%	69,351
Other food	62,053	66%	32,064	34%	94,117
Beverage	21,456	68%	10,012	32%	31,468
Household products	159,610	88%	20,809	12%	180,419
Others products & services	20,966	100%	14	0%	20,980
Total	316,514	80%	79,821	20%	396,335

Note: Figures in percentage show the contribution from each retailer towards each category

In the database, the discounting pattern from the retailers is also an interesting aspect to look at. *Figure 5* demonstrates the daily number of products got discounted by each retailer, and it appears that each retailer has a different discounting scheme. Most of the time, Coopmart has about 15% of their product under some discounting scheme, and about twice a month, there would be briefly one or two-day break when there is almost no discount at all. For Vinmart, the policy is quite the opposite: before the strict lockdown, Vinmart always offers discounts for around 17~18% of the products in the basket, but twice a month, there would be a brief surge of up to 25% for the number of discounted products. After Hanoi went under lockdown, the number of discounted products in Vinmart reduced to only 7~8% and lost the fortnightly pattern.

Figure 5. Percentage of discounted products out of total number by retailer



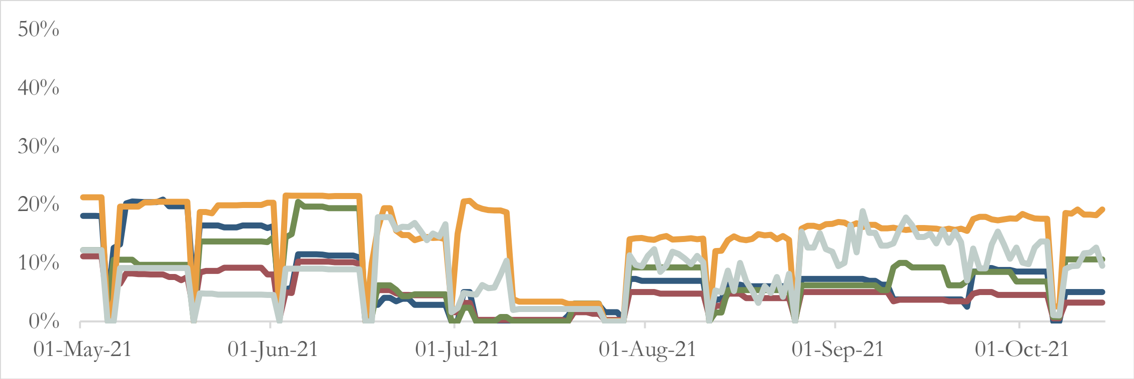
Regarding which products to get the discounted price, household products are more likely to be included in the promotion scheme, with the most frequent discounting pattern in both retailers (*Figure 6*). Non-food and beverage products are the category most likely to experience the highest discount percentage, around 20% on average (*Table 4*), which is probably due to its nature of lower turnover rate. On the other hand, food is the category with the lowest percentage of discounted products, and fewer food products, especially dried food in Hanoi, got discounts since the lockdown was imposed.

Table 4. Discount percentage by product category

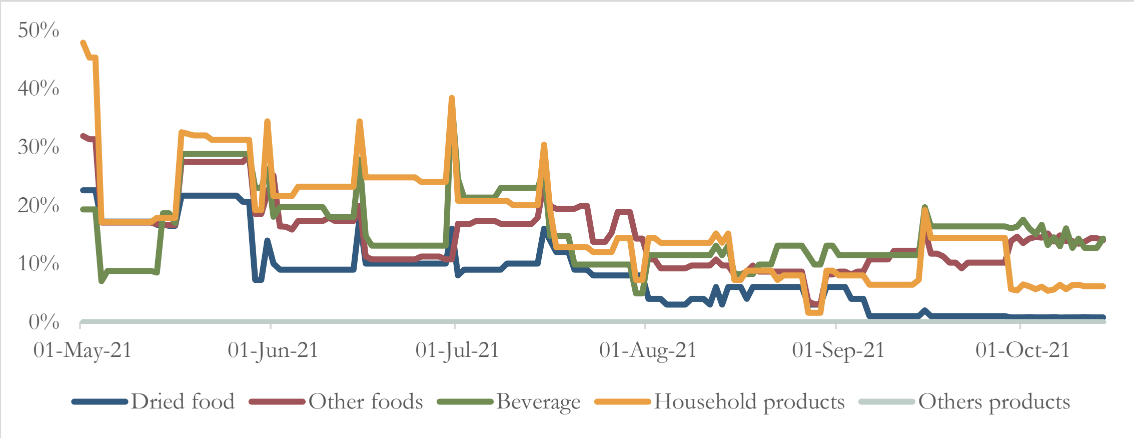
Product category	No. of discounted products	Average discount percentage	Standard deviation
Dried food	5,131	15%	8%
Other food	7,621	14%	6%
Beverage	3,167	16%	7%
Household products	26,854	19%	9%
Others products & services	1,817	23%	7%

Figure 6. Percentage of discounted products out of total number by category

a. Coopmart (HCMC)



b. Vinmart (Hanoi)



4.2 Online price index

Figure 7 demonstrates the movement of the price index based on the actual price sold (including discounted price) versus the original non-discounted price. For both price indexes, the price for dried food leads the inflation during this research time. The price of dried food peaked just before the lockdown loosened up in both cities. By October 14th 2021, the actual price sold on average increased nearly 7% compared to its price at the start of May, while the inflation is only about 3% for non-discounted prices. Apart from that, the movement trends of these two price types show nothing alike. The movement of actual price sold has a wider variance, with short-term periodical surge matching the promotion pattern from Coopmart[10], especially for the prices of household products. On the other hand, the movement of non-discounted prices is closer to real life. According to GSO's press release in September 2021, the change in CPI compared to the previous month is relatively small, oscillating around ±0.5% change for food, beverage, and other household products (GSO, 2021), hence the more discreet movement of the non-discounted price index is more reliable to reflect the inflation than the swinging index based on the actual price sold.

If based on non-discounted prices, food products appear to increase the price more quickly than other categories ever since the start of the lockdown. After the pandemic's peak in mid-August when no citizens in HCMC were allowed outside, the increasing price index

[10] The trend match promotion trend from Coopmart rather than Vinmart because 80% of the products in the database is from the former retailer.

for dried food (including grains, flours, noodles, seasoning, and sauce) accelerated, to more than 2% increase as compared to May. Regarding the level of perishability (

Figure 8), frozen food and seasoning/sauce are two non-perishable subcategories that increase their price the most during the research period. However, the retailers raised the price for frozen food when there were signals for a big new wave, up to 2.5% increase, while the rise of seasoning/sauce price is sharper after the pandemic peak with stricter measures than before.

Meanwhile, significant volatility was observed for the perishable food price index, especially for veggies and fruits. The price of meat, seafood, and egg, similar to the price of seasoning/sauce, also started to increase more quickly from early August; while that of

Figure 7. Price index based on real price sold and non-discounted price

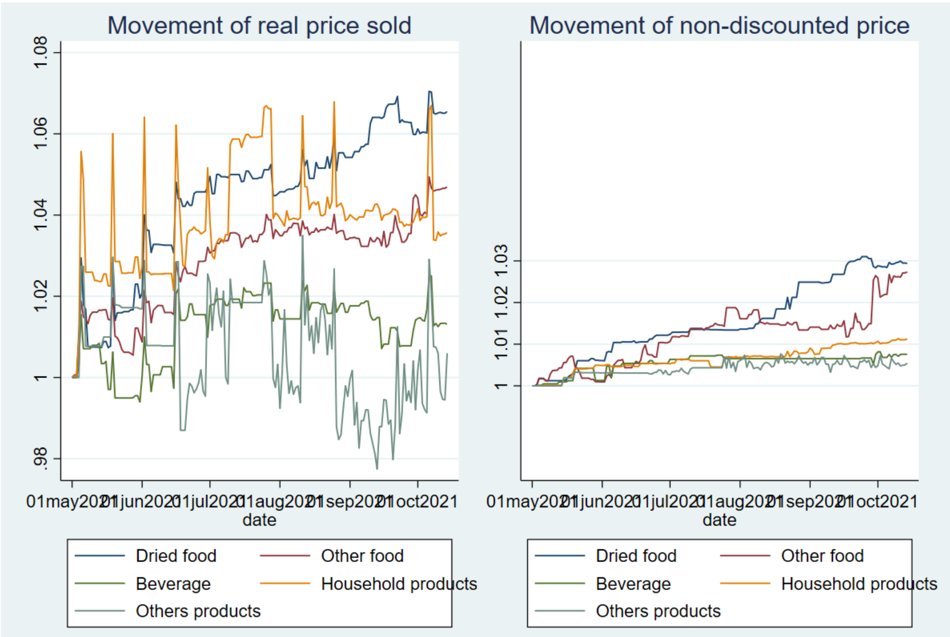
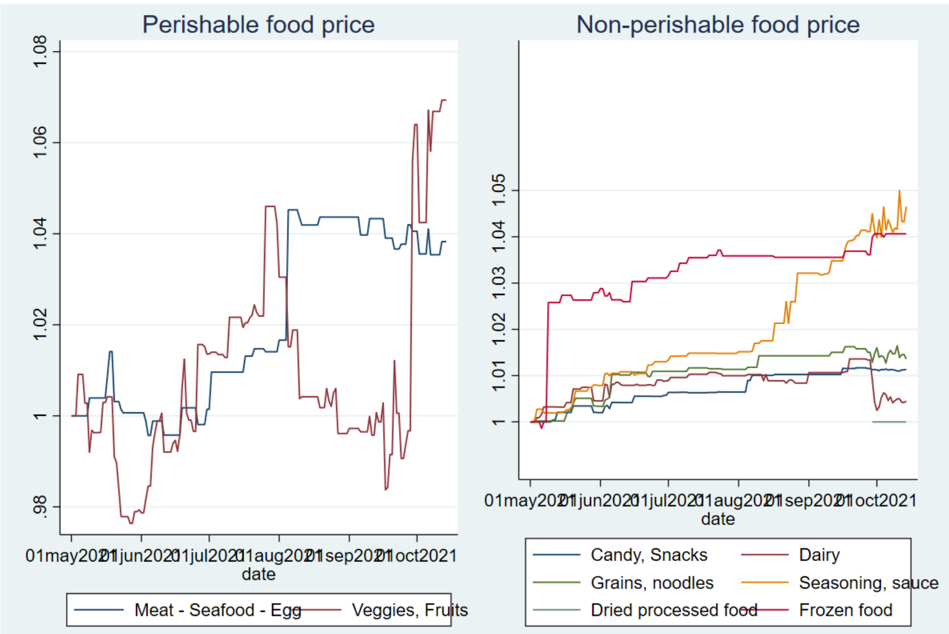


Figure 8. Non-discounted food price index based on the level of perishability



veggies and fruits fluctuated with increasing trend up to the peak of early August and then started decreasing towards the end of September. Veggies and fruits are more susceptible to the product seasonality, as the products got the price reduced more than 50% are mostly summer fruits like "sap" avocado, watermelon, and dragon fruit (*Table 5*). Nonetheless, this subcategory is in fact the one with the highest increasing prices product, that the products have the price more than doubled are all green veggies (*Table 6*), which could be attributed to the difficulty of the logistics for short shelf-life products during the lockdown (Kinh Te Do Thi, 2021).

Table 5. Veggies and fruits that got more than 50% price reduction

Product name	No of days with more than 50% price reduction
Bơ sáp / "Sap" avocado	93
Bưởi năm roi / "Nam roi" pomelo	13
Chuối sứ / "Su" banana	8
Củ cải trắng / Daikon	8
Dưa hấu không hạt / Seedless watermelon	28
Dưa hấu đỏ / Red watermelon	69
Lòng mứt / Sapote	10
Thanh long ruột đỏ / Red dragon fruit	95
Đậu đũa / Green beans	1

Table 6. Products that have their price more than doubled

Product name	No of days with more than double price
Bắp cải trắng / White cabbage	22
Cải ngọt / Bok choy	85
Mồng tơi / Malabar spinach	16
Trái bầu / Calabash	24

4.3 Retailer's consumer pricing strategy

4.3.1 Regression model for consumer price change

On the contrary, Hanoi experience significantly higher price inflation than HCMC across all product categories, with the category most vulnerable to inflation being non-dried food.

Table 7 demonstrates the regression results of the consumer price change model based on the regressand as the original non-discounted price change, in three different ways to proxy the pandemic severity.

The R-squared for all three models is under 0.02, suggesting that the regression models explain little about the dynamic price change for each product (R-squared < 0.01). This does not merge as a surprise since the model lacks the regressors that can reflect the critical determinants of consumer pricing, such as government intervention, competitor, and brand factors. While the main objective of this research is to examine the economic impact of the pandemic on consumer price, the paper will focus only on discussing the beta-coefficients achieved in the model and its significance. The impact of the pandemic variables on the product-level price index is relatively small and insignificant, suggesting that there is not enough evidence to conclude that the dynamic price change we observe from this research is a consequence of the pandemic.

Across all these models, dried food products in Ho Chi Minh City show more susceptibility to higher inflation than the products of other categories during the pandemic as they might be of higher necessity during crisis times, and higher demand over other food because of their longer shelf life. On the contrary, Hanoi experience significantly higher price inflation than HCMC across all product categories, with the category most vulnerable to inflation being non-dried food.

Table 7. Results of consumer price change OLS regression

Dependent variable: Non-discounted price change in percentage at the product level

Exploratory variables	Cases (I1)	Vax (I2)	Lockdown (I3)
Pandemic severity			
$lnCases_t$	-0.017		
$lnVax_t$		-0.053	
$Lockdown_j$			-0.049
$Category_{ij}$ *(HCM – Dried food as base)*			
Hanoi - Dried food	1.543**	1.542**	1.540*
HCM - Other food	-0.653***	-0.653***	-0.651***
Hanoi - Other food	2.139***	2.138***	2.135***
HCM - Beverage	-0.438	-0.439	-0.436
Hanoi - Beverage	1.655*	1.654*	1.650*
HCM - Household products	-0.298***	-0.298***	-0.296***
Hanoi - Household products	1.513**	1.513**	1.512**
HCM - Other products	(omitted because of collinearity)		
Hanoi - Other products	(omitted because of collinearity)		
Constant	99.935***	100.578***	99.885***
N =	396,335	396,335	396,335
R-squared	0.012	0.012	0.012

*Note: */**/*** means coefficient is significant at the 0.1 / 0.05 / 0.01 level*

4.3.2 Regression model for discount patterns

We also found a significantly different discounting pattern between different categories in each city. Compared to the beta-coefficients earned in all three models, the discounting percentage is much lower in food products than other products. Substantially, household products, in general, will have up to 13~15% higher discount percentage than food products in both cities. These results show consistency with what we observed in descriptive statistics in subsection *4.1 Descriptive statistics.* shows the results from tobit regression analysis for retailers' discounting behaviors. On the contrary to the price change model, the discount model finds evidence that the lockdown and the number of vaccinations have a significant impact on retailers' discount behaviors, but there is not yet any evidence for the number of infected cases. Ceteris paribus, a strict lockdown that limits traveling imposed decreases 10.5% of the discount benefit for a certain product as indicated in model D3. Meanwhile, model D2 suggests that 100% increase in the number of total vaccinations can cause the retailers to increase the discount percentage by 2.694%. Overall, the less severe the pandemic gets as perceived through the total vaccinations and the lockdown measure, the more discount benefits the consumers can enjoy from the retailers.

We also found a significantly different discounting pattern between different categories in each city. Compared to the beta-coefficients earned in all three models, the discounting percentage is much lower in food products than other products. Substantially, household products, in general, will have up to 13~15% higher discount percentage than food products in both cities. These results show consistency with what we observed in descriptive statistics in subsection *4.1 Descriptive statistics*. Vinmart Hanoi seems to give more discount than Coopmart HCMC on average, and the discount for dried food in Coopmart tends to be approximately 6.6% higher than that for other food, while it is the opposite for Vinmart, with dried food have around 10% lower discount than the other food.

Table 8. Results of discount percentage tobit regression

Dependent variable: Discount percentage at the product level

Exploratory variables	Cases (D1)	Vax (D2)	Lockdown (D3)
Pandemic severity			
$lnCases_t$	-0.397		
$lnVax_t$		2.694***	
$Lockdown_j$			-10.451***
$Category_{ij}$ *(HCM – Dried food as base)*			
Hanoi - Dried food	2.73	2.731	1.413
HCM - Other food	-6.599***	-6.598***	-6.604***
Hanoi - Other food	12.571***	12.573***	11.179***
HCM - Beverage	0.587	0.59	0.591
Hanoi - Beverage	14.876***	14.877***	13.476***
HCM - Household products	13.949***	13.949***	13.995***
Hanoi - Household products	16.067***	16.068***	14.706***
HCM - Other products	4.400*	4.401*	4.339*
Hanoi - Other products	41.85	41.86	40.088
Constant	-37.678***	-74.149***	-38.458***
N =	396,335	396,335	396,335
Pseudo R-squared	0.023	0.023	0.023

*Note: */**/*** means coefficient is significant at the 0.1 / 0.05 / 0.01 level*

4.4 Discussion of findings

4.4.1 Consumer online price index during the pandemic

The online price index derived from this study suggests that the price movement for the discounted price and non-discounted price follows different patterns. Still, the index based on non-discounted prices is more reliable and shows closer movement to inflation expectation in real life.

Food is the category most susceptible to higher inflation rather than other consumer products during the pandemic in Vietnam. This result somehow resonates the similar findings from the first lockdown phase of Switzerland in Alvarez and Lein's study (2020, p.9-10) where while most consumer categories like housing, energy, household products, and recreation experience deflation, only the food and beverage categories experience a rise in its price index. The food subcategories that got the most inflated price during the research period are frozen food, seasoning/sauce, meat/seafood/eggs, and veggies/fruits. The reasons for increasing prices could differ for each subcategory based on their index pattern concerning the pandemic timeline. The price index of frozen food rose by 2.5% when the number of daily infected cases signaled an upcoming new wave, indicating this subcategory is foreseen to be of high consumer preference during a lockdown, possibly due to its increasing necessity and long shelf life, so consumers can stockpile these products for a long time while waiting for the lockdown to end. Meanwhile, seasoning/sauce prices only rose after the pandemic peak when stricter measures prevented consumers from going outside for grocery shopping. On the other hand, perishable products like meat, seafood, eggs, veggies, and fruit face different problems; despite their similar necessity for daily consumption, their shelf life is short and needs continuous supply. As the supply chain was disrupted because of travel restrictions, these products got higher price in response to more expensive logistic costs; some veggies even doubled their price, like cabbage, bok choy, spinach, and calabash. Nonetheless, the categorical price index saw a plunge from August towards the end of September for veggies and fruits, possibly due to the seasonality shift mainly for summer fruit products.

The result of this study implies that the online price data from multi-channel retailers could be of use as an alternative way to track inflation in Vietnam during times of data uncertainty. However, the price index should be looked at from the categorical level instead of the total level, because firstly, the product database does not cover the whole consumption basket, and secondly, there is the possibility of consumption basket change (Cavallo, 2020; Gautier, Ulgazi and Vertier, 2020) due to the shifted consumer behavior in such a black swan event like the COVID-19 pandemic. Besides, pandemic research with online retail data could pose additional challenges to data representativeness as the number of products on the shelf could fluctuate significantly in a short period, and the insights behind the substantial number of missing products remain a puzzle awaiting the revelation in the future research.

4.4.2 The economic impact of the pandemic on the consumer price

The regression results on the price change (*Table 7*) and price discounting (*4.3.2 Regression* model for discount patterns

We also found a significantly different discounting pattern between different categories in each city. Compared to the beta-coefficients earned in all three models, the discounting percentage is much lower in food products than other products. Substantially, household products, in general, will have up to 13~15% higher discount percentage than food products in both cities. These results show consistency with what we observed in descriptive statistics in subsection *4.1 Descriptive statistics.* shows the results from tobit regression analysis for retailers' discounting behaviors. On the contrary to the price change model, the discount model finds evidence that the lockdown and the number of vaccinations have a significant impact on retailers' discount behaviors, but there is not yet any evidence for the number of infected cases. Ceteris paribus, a strict lockdown that limits traveling imposed decreases 10.5% of the discount benefit for a certain product as indicated in model D3. Meanwhile, model D2 suggests that 100% increase in the number of total vaccinations can cause the retailers to increase the discount percentage by 2.694%. Overall, the less severe the pandemic gets as perceived through the total vaccinations and the lockdown measure, the more discount benefits the consumers can enjoy from the retailers.

We also found a significantly different discounting pattern between different categories in each city. Compared to the beta-coefficients earned in all three models, the discounting percentage is much lower in food products than other products. Substantially, household products, in general, will have up to 13~15% higher discount percentage than food products in both cities. These results show consistency with what we observed in descriptive statistics in subsection *4.1 Descriptive statistics*. Vinmart Hanoi seems to give more discount than Coopmart HCMC on average, and the discount for dried food in Coopmart tends to be approximately 6.6% higher than that for other food, while it is the opposite for Vinmart, with dried food have around 10% lower discount than the other food.

Table 8) demonstrate interesting insights into the economic impact of the pandemic on the pricing decision at the manufacturer and retailer level.

For the manufacturer level, we use the regression model with the original price change of consumer products. The results suggest insufficient evidence for pandemic's impact on manufacturer's pricing strategies, at least within the time bound of this research. On the contrary, discounting price patterns reflect more about the retailers' pricing decisions. Its results show the empirical evidence that the less severe the pandemic gets, the more discount benefits the consumers can enjoy from the retailers. However, this is only the case when looked through the lenses of the vaccination figures and the lockdown decision. All things remain the same, a strict lockdown that limits traveling decreases 10.5% of the discount benefit for a specific product. Meanwhile, 100% increase in the number of total vaccinations can cause the retailers to increase the discount percentage by 2.694%.

This gap might highlight the different levels of price sensitivity between manufacturers and retailers in response to external factors. The pricing set by the manufacturers might take more time to re-adjust that we might not be able to capture it within just only five months and a half. Meanwhile, the retailers can enjoy more versatility in adjusting their pricing, but more likely through changing the discount percentage to make sure their prices look stable and keep a good brand image to the public.

Consistently, we also found that more essential products during this COVID-19 wave, like food, are more vulnerable to price inflation and have less chance of discounting than products of less necessity like household products. Within the food category itself, there is a distinct differentiation in preference between Hanoi and HCMC: dried food is more susceptible to inflation and tends to get less discount than other food in HCMC, but the case is the opposite in Hanoi. One simple explanation for this is different local preferences and different perceptions of which category is more essential than the other during the crisis. However, it could also be that the condition of lockdown measures is also very different between the two cities that its impact on the local product necessity goes in two distinct directions. HCMC suffered a longer and stricter phase of lockdown that for 14 days straight, no one was allowed to go outside. This might make citizens in HCMC perceive that dried food should be of more importance during the lockdown, because of its longer shelf life while they cannot go to the store as often as before. Meanwhile, the lockdown situation is not that strict in Hanoi and people here were still allowed to go to markets or supermarkets periodically. This means that Hanoians get better access to fresh food than HCMC, and there is less need to store non-perishable food. Our pricing model might capture the timespan of the lockdown difference in these two cities, but the qualitative difference in terms of lockdown strictness is challenging to measure and include, and this discrepancy in the result between the HCMC and Hanoi might well reflect this.

Chapter 5
Conclusion

This research paper studied the usage of online price data from multi-channel retailers for economic research. The study had two main objectives. Firstly, it investigates whether and how online price data could be used to track inflation during times of data uncertainty like the COVID-19 pandemic. Secondly, this research also analyses the economic impact of the pandemic on consumer pricing strategies. The final dataset consists of the pricing information from 2,398 products over 5.5 months at two of the largest retailers in Vietnam, which totaled up to 396,335 observations.

The study results suggest that online price data from multi-channel retailers could be used as a quick and easily-accessed source to track the movement of the price index in Vietnam. The results from the consumer pricing model also suggest that the pandemic severity impacted the retailers' discounting patterns, especially for essential products like food products. When a city goes under lockdown, the consumers enjoy fewer discount benefits, but the effect is reversed as more people get vaccinated. However, we have not yet found any evidence that the pandemic impacts inflation, based explicitly on the original price set for the products (prices before discount). The time span of the research reaches only five months and a half, which might not yet demonstrate the pricing sensitivity on the manufacturers' side, and call for research in the future to observe for a more extended period to achieve more complete results. Nonetheless, this study still provides some empirical evidence for policy-makers to boost further vaccination plans and measures to contain the spread of the virus to stabilize the actual prices of essential products that consumers have to pay for in supermarkets (through discounting benefits). The decision to impose a lockdown should be the last resort.

The study adds to the existing literature on using web-scraped pricing data for inflation tracking and on the impact of the pandemic on the demand and supply of consumer products, through the understanding of its direct impact on price. Nonetheless, the lack of official CPI for comparison purpose, together with a short research timeframe, make it challenging to generate concrete evidence for the validity of the generated online price index. Besides, representativeness is still a problem that needs to be carefully considered when constructing the online price index in crisis times, as the data collected does not stand for most products in the consumer basket and the number of products available can fluctuate vastly during the crisis. Last but not least, the impact of the pandemic on consumer price could explain partially how it impacts the consumer demand and supply through a deduction on classical theories, yet to what extent is still unknown. Thus, future research can explore further to enhance the more comprehensive understanding of pandemic factors that might influence consumer price.

Appendices

Appendix 1. Our scraper tool

The Python script for our scraper tool (for both Coopmart and Vinmart) can be found in this Github link: https://github.com/lynth29/e-commerce-sites-scraping. The script is applicable for the retailers' website structure as of October 14th, any change onwards will need further adjustment in the scripting accordingly.

Appendix 2. Variable definition

Table 9. The definition of variables used in this research

Variable name	Value description	Data source
Price	Actual price sold in VND	Web-scraper tool
Old price	Non-discounted price in VND	
Original price change	The price index at time *t* is calculated as its non-discounted price at time *t* divided by the base non-discounted price at the first day the data was collected (t_0 = May 1st, 2021) $Original\ price\ change_t = \left(\frac{Old\ price_t}{Old\ price_0}\right)\%$	
Discount	The discount percentage that retailers offer for one product based on its original price $Discount_t = \left(1 - \frac{Price_t}{Old\ price_t}\right)\%$	
CPI category	*Classified following GSO's guide for CPI research in the period 2020 – 2025 (2019)* 1: Dried food (incl. grains, noodles, flour, seasoning, sauce) 2: Other food 3: Beverage 4: Household products 5: Other products	
City	1: Coopmart / Ho Chi Minh City 2: Vinmart / Hanoi	
Cases	The number of daily infected cases in Vietnam (last-7-day average)	JHU CSSE (Dong, Du and Gardner, 2020)
Vax	The total number of vaccinations in Vietnam	
Lockdown	1: Under strict lockdown (limiting out-of-home traveling) 0: Not under strict lockdown (limiting out-of-home traveling) → *Lockdown = 1 for HCMC between Jul 9th and Sep 30th, for Hanoi between Jul 24th and Sep 20th*	Announcement of official COVID-19 measures

Appendix 3. Data collection at Coopmart (HCMC) supermarket

Table 10. Timeline of store scraping in Coopmart (HCMC)

Date	No. of days	Store in target
May 1st – Jul 11th	72	D7 / Tân Phong
Jul 12th – Jul 18th	7	Closed for store disinfection
Jul 19th	1	D10 / Vạn Hạnh Mall
Jul 20th – Jul 24th	5	D. Thủ Đức / Phạm Văn Đồng
Jul 25th – Oct 14th	82	D7 / Tân Phong

Figure 9. Number of products collected from Coopmart (HCMC)

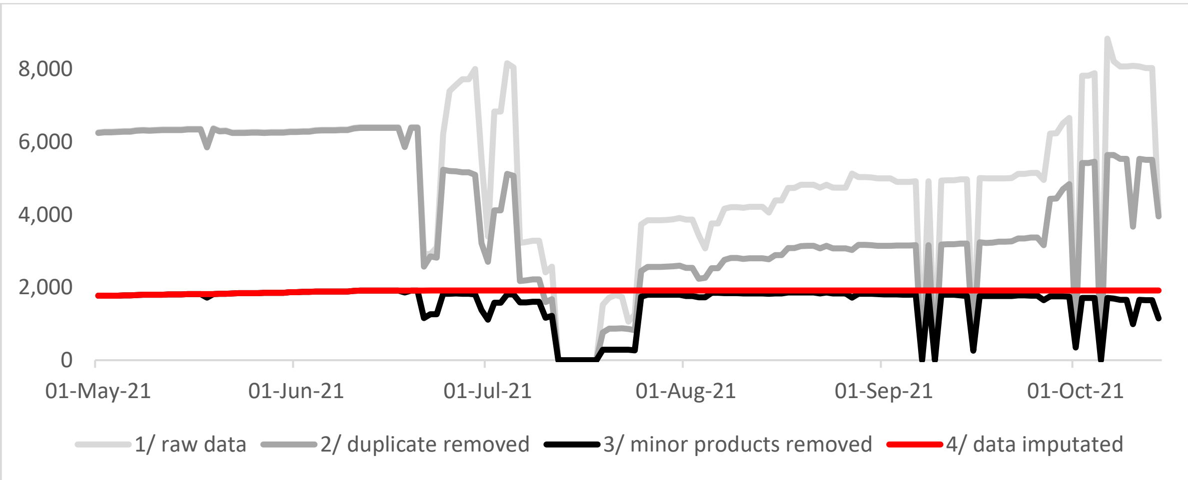


Figure 10. Number of products by discounting scheme in Coopmart (HCMC)

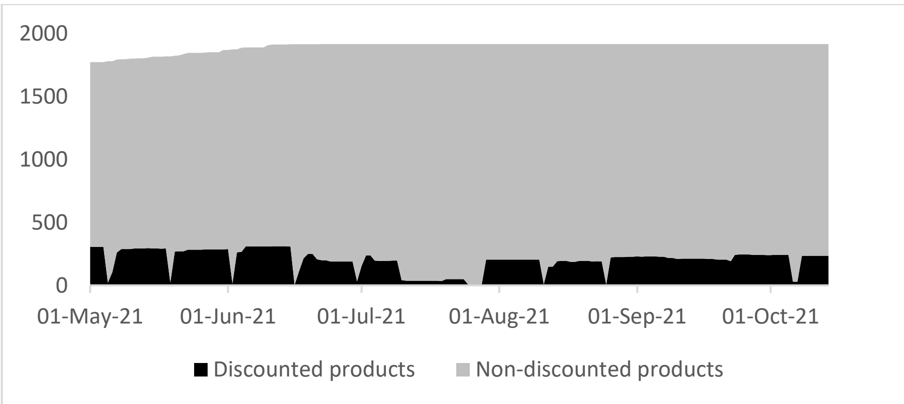


Table 11. Discount percentage at Coopmart (HCMC) by product category

Product category	No. of discounted products	Average discount percentage	Standard deviation
Dried food	3776	15%	7%
Other food	2907	14%	5%
Beverage	1599	15%	6%
Household products	23396	19%	9%
Others products & services	1811	23%	7%

Appendix 4. Data collection at Vinmart (Hanoi) supermarket

Table 12. Timeline of store scraping in Vinmart (Hanoi)

Date	No. of days	Store in target
May 1st – Jul 5th	66	D. Hai Bà Trưng / Bách Khoa
Jul 6th – Jul 7th	2	Closed for store disinfection
Jul 8th – Jul 30th	23	D. Hai Bà Trưng / Bách Khoa
Jul 31st – Aug 3rd	4	Đống Đa / Trung Tự
Aug 4th – Aug 7th	4	Cầu Giấy / Dịch Vọng
Aug 8th – Aug 9th	2	D. Hai Bà Trưng / Bạch Đằng
Aug 10th – Sep 23rd	45	D. Hai Bà Trưng / Minh Khai
Sep 24th – Sep 28th	5	Website maintenance (Winmart)
Sep 29th – Oct 14th	16	D. Hai Bà Trưng / Minh Khai

Figure 11. Number of products collected from Vinmart (Hanoi)

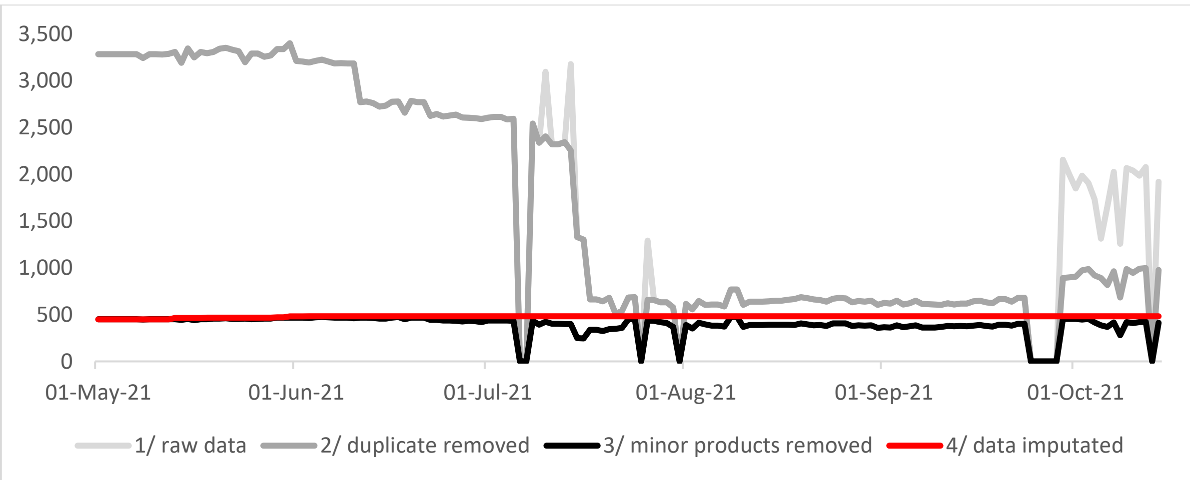


Figure 12. Number of products by discounting scheme in Vinmart (Hanoi)

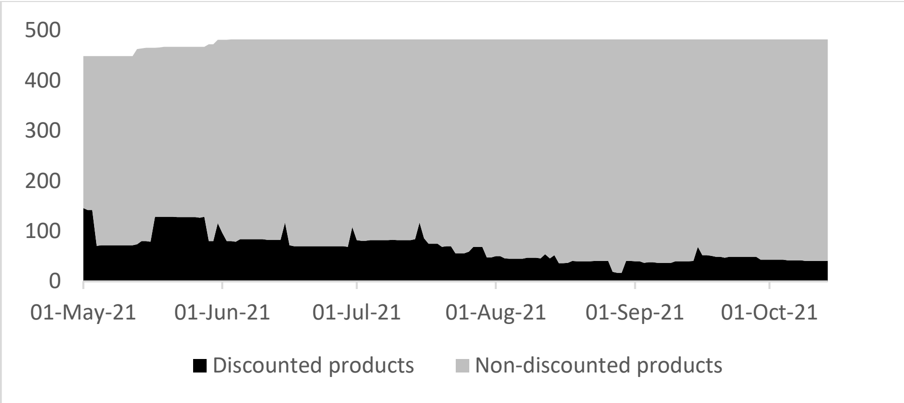
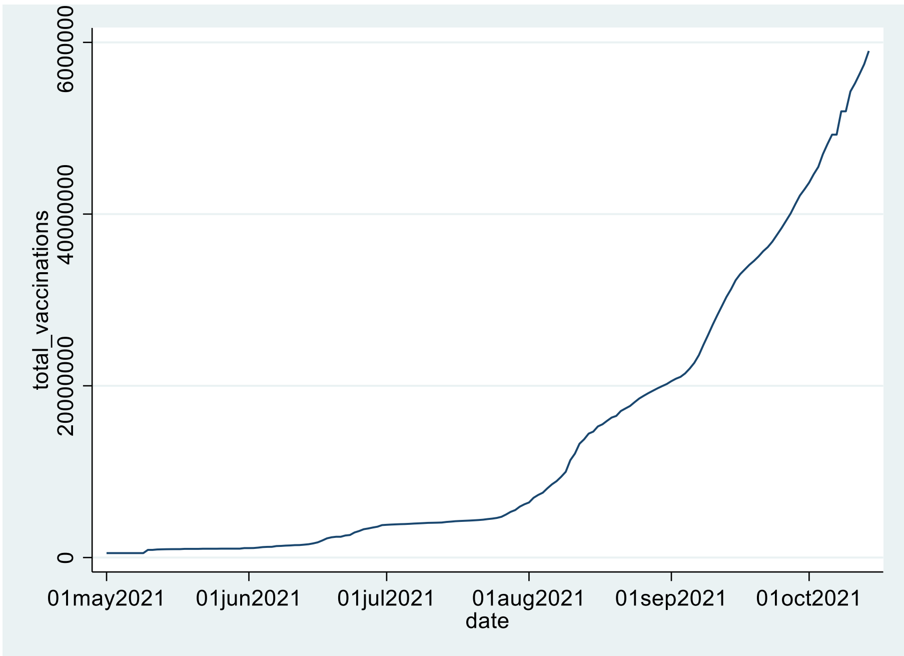


Table 13. Discount percentage at Vinmart (Hanoi) by product category

Product category	No. of dis-counted products	Average discount percentage	Standard deviation
Dried food	1355	17%	10%
Other food	4714	14%	6%
Beverage	1568	17%	8%
Household products	3458	18%	7%
Others products & services	6	23%	0%

Appendix 5. Vaccination statistics in Vietnam

Figure 13. Total number of vaccinations in Vietnam



Source: JHU CSSE (Dong, Du and Gardner, 2020)

References

Aday, S. and Aday, M.S. (2020) 'Impact of COVID-19 on the food supply chain', *Food Quality and Safety,* 4(4), pp. 167-180. doi: 10.1093/fqsafe/fyaa024.

Alvarez, S.E. and Lein, S.M. (2020) 'Tracking inflation on a daily basis', *Swiss Journal of Economics and Statistics,* 156(1), pp. 1-13.

Antonakakis, N., Cunado, J., Gil-Alana, L.A. and Gupta, R. (2016) 'Is inflation persistence different in reality?', *Economics Letters,* 148, pp. 55-58.

Aparicio, D. and Bertolotto, M.I. (2020) 'Forecasting inflation with online prices', *International Journal of Forecasting,* 36(2), pp. 232-247.

Arora, N., Charm, T., Grimmelt, A., Ortega, M., Robinson, K., Sexauer, C. and Yamakawa, N. (2020) *A global view of how consumer behavior is changing amid COVID-19.* McKinsey and Company. Available at: https://www.mckinsey.de/~/media/McKinsey/Business%20Functions/Marketing%20and%20Sales/Our%20Insights/A%20global%20view%20of%20how%20consumer%20behavior%20is%20changing%20amid%20COVID%2019/20200707/covid-19-global-consumer-sentiment-20200707.pdf (Accessed: Jun 14, 2021).

Balleer, A., Link, S., Menkhoff, M. and Zorn, P. (2020) 'Demand or supply? Price adjustment during the Covid-19 pandemic', *CESifo Working Paper,* .

Banerjee, R., Singhal, N. and Subramanian, C. (2018) 'Predicting food price inflation through online prices in India', *Economic and Political Weekly,* 53(23), pp. 132-135.

Barro, R.J. (2013) 'Inflation and Economic Growth', *Annals of Economics and Finance,* 14(1), pp. 121-144.

Bentley, A. and Krsinich, F. (2017) *Towards a big data CPI for New Zealand.* 15th Meeting of the Ottawa Group on Price Indices. Eltville, Germany: 15th Meeting of the Ottawa Group on Price Indices, .

Berardi, N., Sevestre, P. and Thébault, J. (2017) 'The determinants of consumer price dispersion: Evidence from French supermarkets'*The econometrics of multi-dimensional panels* Springer, pp. 427-449.

Bijmolt, T.H., Van Heerde, H.J. and Pieters, R.G. (2005) 'New empirical generalizations on the determinants of price elasticity', *Journal of Marketing Research,* 42(2), pp. 141-156.

Blustein, D.L., Duffy, R., Ferreira, J.A., Cohen-Scali, V., Cinamon, R.G. and Allan, B.A. (2020) 'Unemployment in the time of COVID-19: A research agenda', *Journal of vocational behavior,* 119, pp. 103436. doi: https://doi.org/10.1016/j.jvb.2020.103436.

Breton, R., Swiel, N. and O'Neil, R. (2015) *Using Web Scraped Data to Construct Consumer Price Indices.* pp. 9.

Cavallo, A. (2020) 'Inflation with covid consumption baskets', *NBER Working Paper Series,* . doi: 10.3386/w27352.

Cavallo, A. (2017) 'Are online and offline prices similar? Evidence from large multi-channel retailers', *American Economic Review,* 107(1), pp. 283-303.

Cavallo, A. (2012) 'Online and official price indexes: Measuring Argentina's inflation', *Journal of Monetary Economics,* 60(2), pp. 152. doi: 10.1016/j.jmoneco.2012.10.002.

Cavallo, A. and Rigobon, R. (2016) 'The billion prices project: Using online prices for measurement and research', *Journal of Economic Perspectives,* 30(2), pp. 151-178.

Chau, A. (2021) *Tình hình cung ứng hàng hóa tiếp tục ổn định trong trạng thái "bình thường mới".* Available at: http://moit.gov.vn/tin-tuc/thi-truong-trong-nuoc/tinh-hinh-cung-ung-hang-hoa-tiep-tuc-on-dinh-trong-trang-thai-binh-thuong-moi-.html (Accessed: Oct 26, 2021).

da Silva, L.T., de Oliveira, I.L., Dantas, T.M. and Miranda, V.G. (2019) *Studies of new data sources and techniques to improve CPI compilation in Brazil.*

Deloitte Vietnam (2019) *Retail in Vietnam: Navigating the digital retail landscape.* Available at: https://www2.deloitte.com/vn/en/pages/consumer-business/articles/vietnam-consumer-retail-2019.html (Accessed: Jun 14, 2021).

Dong, E., Du, H. and Gardner, L. (2020) 'An interactive web-based dashboard to track COVID-19 in real time', *The Lancet infectious diseases,* 20(5), pp. 533-534.

eCommerceDB (2021) *eCommerce Market Vietnam - Data, Trends, Top Stores.* Available at: http://ecommercedb.com/en/markets/vn/all?q=%2Fen%2Fmarkets%2Fvn%2Fall (Accessed: Sep 23, 2021).

Ellickson, P.B. and Misra, S. (2008) 'Supermarket Pricing Strategies', *Marketing Science,* 27(5), pp. 811-828.

Gautier, E., Ulgazi, Y. and Vertier, P. (2020) 'Inflation and households' inflation expectations during the Covid-19 pandemic', *Banque De France, Eco Notepad,* 171.

Grant, B. (2021) *Vietnam was ranked 2nd for successfully handling the coronavirus pandemic, and its contact tracing was so good it barely had to lock down.* Available at: https://www.businessinsider.com/vietnam-coronavirus-measures-among-best-in-world-contact-tracing-masks-2021-2 (Accessed: Nov 10, 2021).

Griffioen, R., de Haan, J. and Willenborg, L. (2014) *Collecting clothing data from the Internet.* Statistics Netherlands, .

Groves, R.M. (2011) 'Three eras of survey research', *Public opinion quarterly,* 75(5), pp. 861-871.

GSO (2021) *Thông cáo báo chí về tình hình giá tháng 9, Quý III và 9 tháng đầu năm 2021.* Available at: https://www.gso.gov.vn/du-lieu-va-so-lieu-thong-ke/2021/09/thong-cao-bao-chi-ve-tinh-hinh-gia-thang-9-quy-iii-va-9-thang-dau-nam-2021/ (Accessed: Nov 5, 2021).

GSO (2019) *Phương án điều tra giá tiêu dùng thời kỳ 2020 - 2025.* Hà Nội: Nhà Xuất Bản Thông Tấn. Available at:(Accessed: Mar 06, 2021).

Harchaoui, T.M. and Janssen, R.V. (2018) 'How can big data enhance the timeliness of official statistics?: The case of the U.S. consumer price index', *International Journal of Forecasting,* 34(2), pp. 225-234. doi: 10.1016/j.ijforecast.2017.12.002.

ILO, Eurostat, United Nations Economic Commission for Europe, OECD, World Bank and IMF (2020) *Consumer Price Index Manual: Concepts and Methods.* Washington, DC: International Monetary Fund.

Joshi, A., Bhaskar, P. and Gupta, P.K. (2020) 'Indian economy amid COVID-19 lockdown: A prespective', *J.Pure Appl.Microbiol,* 14, pp. 957-961.

Kantar Worldpanel (2020) *FMCG Consumer Survey Mar 2020.* Available at:(Accessed: Jun 01, 2021).

Kinh Te Do Thi (2021) *Giá rau xanh tại TP Hồ Chí Minh tăng vọt vì dịch Covid-19.* Available at: https://kinhtedothi.vn/gia-rau-xanh-tai-tp-ho-chi-minh-tang-vot-vi-dich-covid-19-425260.html (Accessed: Nov 5, 2021).

Lipsitch, M. and Dean, N.E. (2020) 'Understanding COVID-19 vaccine efficacy', *Science,* 370(6518), pp. 763-765.

Luu, Q. and Tran, T. (2021) *Tình hình cung ứng và giá cả hàng hóa tại các tỉnh, thành phố phía Nam ngày 26/7.* Available at: https://dms.gov.vn/tin-chi-tiet/-/chi-tiet/tinh-hinh-cung-ung-va-gia-ca-hang-hoa-tai-cac-tinh-thanh-pho-phia-nam-ngay-26-7-35888-4.html (Accessed: Oct 25, 2021).

Mahajan, Y. (2020) 'Impact of coronavirus pandemic on fast moving consumer goods (FMCG) sector in India', *Journal of Xi'an University of Architecture & Technology,* 12.

Mead, D. (2020) 'The impact of the COVID-19 pandemic on food price indexes and data collection.', *Monthly Labor Review,* .

MOIT (2021) *TP HCM đưa ra 4 phương án sản xuất cho doanh nghiệp.* Available at: http://moit.gov.vn/tin-tuc/doanh-nghiep/tp-hcm-dua-ra-4-phuong-an-san-xuat-cho-doanh-nghiep.html (Accessed: Oct 26, 2021).

Moody's Analytics (2021) *Consumer Price Index (CPI) - Country / Jurisdiction List* . Available at: https://www.economy.com/indicators/consumer-price-index-cpi (Accessed: Jun 14, 2021).

Mordor Intelligence (2021) *Vietnam Retail Market | 2021 - 26 | Industry Share, Size, Growth.* Available at: https://www.mordorintelligence.com/industry-reports/retail-industry-in-vietnam (Accessed: Oct 21, 2021).

Morgenroth, W.M. (1964) 'A Method for Understanding Price Determinants', *Journal of Marketing Research,* 1(3), pp. 17-26.

Nguyen, D.H. and Bui, T.T.L. (2019) *Vietnam E-commerce Platforms: Pricing Behavior and Inflation Nowcasting*. Available at: http://vepr.org.vn/533/ebook/vietnam-e-commerce-platforms-pricing-behavior-and-inflation-nowcasting/27214.html (Accessed: Mar 05, 2021).

Nguyen, H.T., Nguyen, M.N., Pham, M.D., Tran, M.T., Nguyen, T.V. and Santural, J.R.J. (2020) 'Comparative analysis of price policy of Coopmart and Big C in Vietnam retail industry', *International Journal of Advanced Educational Research,* 5(3), pp. 38-42.

Nguyen, M.N. (2021) *Vietnam: COVID-19 impact on employment status 2021*. Available at: https://www.statista.com/statistics/1200315/vietnam-covid-19-impact-on-employment-status/ (Accessed: Oct 26, 2021).

Nguyen, Q. (2021) *Bách Hóa Xanh tăng giá bán hàng là đi ngược với đạo lý của thời cuộc*. Available at: https://vov.vn/kinh-te/thi-truong/bach-hoa-xanh-tang-gia-ban-hang-la-di-nguoc-voi-dao-ly-cua-thoi-cuoc-875734.vov (Accessed: Nov 12, 2021).

Nielsen Vietnam (2020) *Vietnam Retail Audit Mar 2020*. Available at:(Accessed: Jun 01, 2021).

NielsenIQ (2020) *Key Consumer Behavior Thresholds Identified as the Coronavirus Outbreak Evolves*. Available at: https://nielseniq.com/global/en/insights/analysis/2020/key-consumer-behavior-thresholds-identified-as-the-coronavirus-outbreak-evolves-2/ (Accessed: Jun 14, 2021).

Nygaard, R. (2015) 'The use of online prices in the Norwegian Consumer Price Index', *Statistics Norway,* .

Pindyck, R.S., Rubinfeld, D.L. and Rabasco, E. (2013) *Microeconomia*. Pearson Italia.

Pollack, T., Thwaites, G., Rabaa, M., Choisy, M., van Doorn, R., Le, V.T., Duong, H.L., Dang, Q.T., Tran, D.Q., Phung, C.D., Ngu, D.N., Tran, A.T., La, N.Q., Nguyen, C.K., Dang, D.A., Tran, N.D., Le, M.S., Pham, Q.T., Vu, D. and Exemplars in Global Health (2021) *Emerging COVID-19 success story: Vietnam?s commitment to containment*. Available at: https://ourworldindata.org/covid-exemplar-vietnam (Accessed: Nov 10, 2021).

Reuters (2021) *Vietnam deploys troops to enforce COVID lockdown in largest city*. Available at: https://www.reuters.com/world/asia-pacific/vietnam-deploys-troops-enforce-lockdown-largest-city-2021-08-23/ (Accessed: Oct 17, 2021).

Roberts, M.R. and Whited, T.M. (2013) 'Endogeneity in empirical corporate finance1'*Handbook of the Economics of Finance* Elsevier, pp. 493-572.

Shankar, V. and Bolton, R.N. (2004) 'An empirical analysis of determinants of retailer pricing strategy', *Marketing Science,* 23(1), pp. 28-49.

Sheth, J. (2020) 'Impact of Covid-19 on consumer behavior: Will the old habits return or die?', *Journal of Business Research,* 117, pp. 280-283. doi: 10.1016/j.jbusres.2020.05.059.

Tan, H.Y. (2021) 'Vietnam's 4th wave of Covid-19 is its first 'real wave', after earlier blips ', *Straits Times,* May 22, .

Vietnambiz (2021) *Bách Hoá Xanh tuyên bố tăng giá bán không phải để kiếm lời và màn đáp trả cực gắt từ Saigon Co.op, VinMart.* Available at: https://vietnambiz.vn/bach-hoa-xanh-tuyen-bo-tang-gia-ban-khong-phai-de-kiem-loi-va-man-dap-tra-cuc-gat-tu-saigon-coop-vinmart-20210715131126293.htm (Accessed: Nov 12, 2021).

VietnamPlus (2021) *Toàn cảnh thực hiện Chỉ thị 16 để phòng chống COVID-19 tại TP.HCM.* Available at: https://www.vietnamplus.vn/toan-canh-thuc-hien-chi-thi-16-de-phong-chong-covid19-tai-tphcm/727133.vnp (Accessed: Sep 10, 2021).

VnExpress (2021) *Thực phẩm tươi sống tìm lối thoát trong đại dịch - VnExpress.* Available at: https://vnexpress.net/thuc-pham-tuoi-song-tim-loi-thoat-trong-dai-dich-4336326.html (Accessed: Oct 26, 2021).

Widnyana, I.W. and Widyawati, S.R. (2020) 'Supply of Consumer Goods, Per Capita Consumption due to Covid-19 Pandemic', *Economics Development Analysis Journal,* 9(4), pp. 458-467.

Wilms, R., Mäthner, E., Winnen, L. and Lanwehr, R. (2021) 'Omitted variable bias: A threat to estimating causal relationships', *Methods in Psychology,* 5, pp. 100075. doi: https://doi.org/10.1016/j.metip.2021.100075.

Wright, O. and Blackburn, E. (2020) *How COVID-19 will permanently change consumer behavior, Accenture.* Available at: https://www.accenture.com/_acnmedia/PDF-134/Accenture-COVID19-Consumer-Behaviour-Survey-Research-PoV.pdf#zoom=40 (Accessed: Jun 14, 2021).