Spring 5-4-2022

Grocery Sales and the Pandemic

Joy A. Salentine
Northern Illinois University, z1843650@students.niu.edu

NORTHERN ILLINOIS UNIVERSITY

Grocery Sales and the Pandemic

A Capstone Submitted to the

University Honors Program

In Partial Fulfillment of the

Requirements of the Baccalaureate Degree

With Honors

Department Of

Economics

By

Joy Salentine

DeKalb, Illinois

Spring 2022

University Honors Program
Capstone Faculty Approval Page

Capstone Title (print or type)

Grocery Sales and the Pandemic

Name (print or type) **Joy Salentine**

Faculty Supervisor (print or type) **Virginia Wilcox**

Faculty Approval Signature [signature]

Department of (print or type) **Economics**

Date of Approval (print or type) **May 3, 2022**

Date and Venue of Presentation **NIU ECONference, April 13, 2022**

Check if any of the following apply, and please tell us where and how it was published:

- ☐ Capstone has been published (Journal/Outlet):

- ☐ Capstone has been submitted for publication (Journal/Outlet):

Completed Honors Capstone projects may be used for student reference purposes, both electronically and in the Honors Capstone Library (CLB 110).

If you would like to opt out and not have this student's completed capstone used for reference purposes, please initial here: _______ (Faculty Supervisor)

Abstract

The focus of this paper is to examine the impact of COVID-19 on grocery sales in the United States. My hypothesis is that periods with higher COVID-19 severity, measured through the number of COVID-19 deaths, will show higher grocery sales, all things equal, due to the increased demand during the unusual circumstances of the pandemic. Using data from the Centers for Disease Control and Prevention, the U.S Department of Agriculture, and the Bureau of Economic Analysis, I perform a regression analysis to test my hypothesis. The results confirm my hypothesis by showing that there is a positive and significant effect of severity on sales. However, the effect is inelastic and relatively small. My results are consistent with related literature. These results are important for understanding the changes that can occur during times of uncertainty in an otherwise stable industry.

Keywords: COVID-19, pandemic, grocery, shift in demand

Grocery Sales and the Pandemic – Joy Salentine

Spring 2022 ECON 492

Introduction

Grocery is an important industry in the economy; everyone needs to eat. COVID-19 proved to a be a large economic disturbance in many ways, but it seems to have had some positive impacts in the grocery industry. How has the pandemic affected grocery sales? My hypothesis is that periods with higher COVID-19 severity, shown through the number of COVID-19 deaths, will show higher grocery sales, all things equal. With panic buying, greater transition to working from home, and restaurant closures, there was a greater demand for groceries in the United States.

It's important to investigate how the pandemic has affected the grocery industry to understand the shifts in demand and changes in consumer behavior that have occurred. The impacts made from the pandemic could also signal some permanent changes to the industry itself. Since the grocery industry plays a role in everyone's life, it's important to understand the shifts and trends that are happening, especially through unusual circumstances such as a global pandemic. More complete understanding can help to explain future data that might look different due to the changes that are occurring presently.

Literature Review

Since COVID-19 is a more recent event in history, there is a more limited source of literature that has studied its impact on the grocery industry specifically. In one early response to the shifts in the food industry, Hobbs (2020) discussed the shocks to the food supply chain in Canada for both supply and demand. Most notably, there were increases in demand due to the consumer behavior of panic buying. Further pressure on demand came from restaurant closures redirecting consumers to buy food elsewhere, especially with increased working from home. As for the supply side, Hobbs (2020) noted the potential for labor shortages and difficulties in transporting goods to uphold supply. This journal article is important for understanding the changes that the food industry started experiencing with the pandemic, which helps to make sense of results in further studies.

One initial study by Baker et al. (2020) examined the response of household spending during the onset of the pandemic. To track consumer spending, they used observational data at the transaction level from SaverLife, which is a Fintech company that helps users save money and records their daily spending through linked cards and accounts. The empirical study used a sample of 4,735 users that had updated their accounts in March 2020 and that had transactions in the preceding months to get a complete panel dataset. The study focuses on the weeks starting February 26, 2020, through March 31, 2020, and looks at spending in different categories.

Baker et al. (2020) used OLS regression and had the weekly spending as the dependent variable. In their models, they found that even though many areas of household consumption were decreasing in the second half of March 2020, grocery sales remained elevated—especially in areas with shelter-in-place protocols. They also concluded that the response in spending to the pandemic was partially influenced by the severity of the pandemic in the state. This is consistent with consumer behavior of stock piling and the closure of restaurants that occurred at the beginning of the pandemic.

Goolsbee and Syverson (2021) conducted another study analyzing the impact of the pandemic on the economic slowdown. They used data from the SafeGraph panel of mobile phone use that collects data on the number of visits to areas specified by address through phone location. The data they used from this panel is the weekly number of visits to a sample of over two million businesses that would indicate a measure on economic activity from March 1, 2020, through May 16, 2020, as well as monthly visits for preceding months. In their OLS regression models, they used transformations of visits per day for businesses as the dependent variable to test the impact of policy and COVID-19 fears on the number of visits.

With consumers decreasing visits during the pandemic, Goolsbee and Syverson (2021) concluded that the economic slowdown was driven more by fear of infection rather than policy restrictions. They also found that lockdown orders did shift consumers to visiting “essential” businesses, like grocery stores instead of restaurants. An almost 30% decrease in visits to bars and restaurants was paralleled with about a 27% increase in visits to non-restaurant food and drink businesses. While this study did not examine expenditure, it points out the changing behavior of consumers during the pandemic. There was increased focus on more essential businesses with the fear about COVID-19 playing a significant role in the shift of consumer behavior.

Another study from the Netherlands by Baarsma and Groenewegen (2021) looked at the effects of the pandemic in the realm of online grocery. They acquired panel data from an online-only Dutch grocer named Picnic, with data from 2018, 2019, and the beginning of 2020 used as a control group and data from the end of February 2020 onward as the treatment group. The dependent variables of their OLS regression models came from this data, including the sales per online order. They also used data from the Dutch National Institute for Public Health and the Environment for the number of COVID-19 hospital admissions and Google search data on keywords to represent local severity and perception on the pandemic to see how they affected the demand for online grocery.

Baarsma and Groenewegen (2021) found evidence in their regression models that increased numbers of hospital admissions led to higher demand for online grocery and to orders with higher sales. They concluded that local hospital admissions and perception of the pandemic affected the purchasing behavior for online grocery. Even though this study focuses on the online grocery segment, this finding still supports that there was a shift in consumer behavior that showed increases in grocery sales.

Another study from Chang & Meyerhoefer (2021) examined the changes in demand for online grocery in Taiwan. They used a combination of data from an e-commerce food platform named Ubox, COVID-19 case numbers from the CDC, and keyword search data from Google and media sources to create a set of panel data. The data starts from the first confirmed case in Taiwan near the end of January 2020 to the beginning of April 2020. Additionally, data from prior years of the same week range was used for control groups. Their regression models analyzed the impact of the number of cases and keyword search indexes on the dependent variables of sales and customer levels for the online shopping platform.

Chang & Meyerhoefer (2021) concluded that an additional COVID-19 case resulted in a 5.7% increase in sales, as well as a 4.9% increase in the number of customers. Increased keyword searches online and increased mentions in the media also had a positive impact on sales. This study confirmed the increase in online demand that the pandemic created. The results are specific to Taiwan, but the underlying principles are still relevant in other countries; Increased cases and increased awareness of the state of the pandemic through keyword searches and the media led to higher demand for online grocery.

Overall, these articles support the theory that the demand for groceries experienced a great shock from the pandemic. Each of the studies gives a slightly different perspective into the situation. While Hobbs (2020) did not perform an empirical study, she discusses the underlying shocks in demand from the side of both supply and demand. Baker et al. (2020) emphasize the changing behaviors of consumption in response to the pandemic, which shows increased grocery spending at the onset of the pandemic. In contrast, Goolsbee and Syverson (2021) examine the change in consumer visits that are primarily driven by fear of infection in the initial months of the pandemic. Baarsma and Groenewegen (2021) look at the online side of grocery and find that perception and severity of the pandemic play a role in changing consumption in grocery as well. Chang & Meyerhoefer (2021) also examine the increase in online demand for grocery due to the number of cases and awareness of the pandemic. All the studies support the theory that COVID-19 had increasing effects on demand for groceries.

Additionally, each of the four empirical studies used similar methods to come to their conclusions. All four used various models of OLS regression to get their results. However, they all used very different sources of data for their empirical work. Baker et al. (2020) used transaction level data from a Fintech company. This provided very granular data that was useful in measuring the immediate responses of consumers during the initial stages of the pandemic. Goolsbee and Syverson (2021) had a very different approach. They used data from the use of cellular phones to track the number of visits to different businesses. While they did not use any metrics on expenditure, visits to stores gave another angle in looking at changes in economic activity and consumer behavior. Baarsma and Groenewegen (2021) used data from an online grocer to specifically address the demand and expenditure of online grocery. Chang & Meyerhoefer (2021) used data from an online food platform in another country to assess changes in demand. The data varies across the studies, but all the data comes together to help tell the story of consumer behavior for grocery in response to the pandemic.

Since there is still limited research available on the specific impacts of the pandemic on grocery, my project will continue to build on the studies that have been done. To summarize the findings above, consumers were changing their demand in food retail due to panic buying and more working from home, increasing their spend on grocery while decreasing spend elsewhere, shifting their visits to essential businesses due to COVID-19 fears, and purchasing more online with higher expenditure (Hobbs, 2020; Baker et al., 2020; Goolsbee & Syverson, 2021; Baarsma & Groenewegen, 2021; Chang & Meyerhoefer, 2021). My hypothesis that higher COVID-19 severity will result in higher grocery sales, all things equal, seems to be consistent with the findings in these studies. Additionally, my project looks at a different time frame than the studies that I have noted here. The focus of these studies was the onset of the pandemic; I have looked at a longer timeframe into 2021, which will provide further insight into the lasting effects of the pandemic over time.

Theory

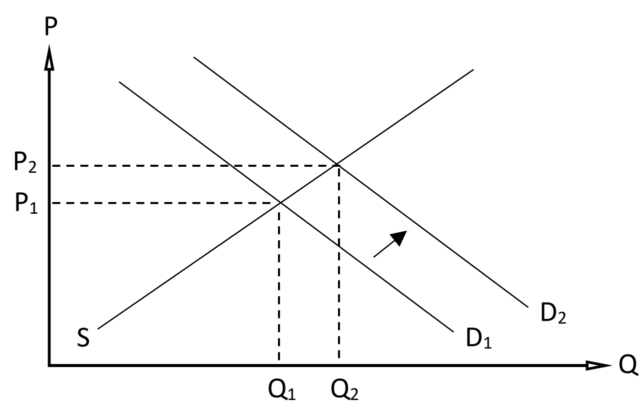
The demand for groceries is generally stable because everyone must buy food. However, COVID-19 caused some key shifts in the demand for groceries. In the beginning of the pandemic, many restaurants started to close. This caused the demand for groceries to increase because people needed to supplement eating out with eating at home. Also, many workers were required to stay home or work remotely during peak periods of the pandemic, which created additional need for food at home versus food away from home. The pandemic forced a shift in tastes toward buying more food to prepare at home.

Another factor aiding the shift in demand for groceries was expectations of the future. All the uncertainty surrounding the pandemic caused a great amount of panic buying. People were unsure what the future conditions would be, which led them to buy large quantities compared to what they normally would buy. When the pandemic had peaks in severity and uncertainty, the demand curve also would have shifted to the right because of these consumer expectations.

With this theory of demand for groceries increasing during the pandemic, I will be testing the hypothesis that periods with a higher COVID-19 severity indicate higher grocery sales, all things equal. Since an increase in demand would result in larger quantities being purchased and upward pressure on price, the sales would increase from initial prices (P_1) multiplied by initial quantities (Q_1) to increased prices (P_2) multiplied by increased quantities (Q_2) as shown in Figure 1. This would result in increased sales during those periods of higher severity. This hypothesis addresses whether increases in sales can be determined by the severity of the pandemic given the changes in demand.

Figure 1

Increase in Demand



Data

To track the sales in grocery over the pandemic, I used data from the U.S. Department of Agriculture, which lists data collected by Information Resources Inc. (IRI). The dataset provided by IRI shows the weekly retail food sales from proprietary scanner data. The scanner data reports point-of-sale revenues and quantities for products in the stores. This data is collected from a representative sample of food retailers. The data is aggregated into 10 product categories and is reported for 39 states by week. Alaska, Delaware, Hawaii, Idaho, Iowa, Kansas, Nebraska, New Jersey, North Dakota, Mississippi, and Montana are not included. Data is not available for all states due to the limitations of what IRI reports, but the samples that are included are representative at the state level. The five product categories that I used are Dairy, Vegetables, Fruits, Grains, and Meats/Eggs/Nuts. The timeframe of the data is October 2019 through November 2021, which covers a large portion of the pandemic.

As an indicator of COVID-19 severity, I used data on the number of COVID-19 deaths provided by the Centers for Disease Control and Prevention (CDC). The CDC provides daily numbers of deaths by state from January 2020 up to the present. I rolled up this daily data into weekly data and joined it to the sales data to perform regression analysis. For the combined dataset, I started from March 2020 (the beginning of the pandemic) and used data through November 2021, which is the latest available data for sales. This gave 91 weeks per state starting from the week ending March 8, 2020, through the week ending November 28, 2021, for a total of 3,549 observations for each food category. Each observation was also assigned its respective season (fall, winter, spring, or summer) according to the date.

Additionally, I used data from the Bureau of Economic Analysis (BEA) that gives income per capita and population. The BEA provides quarterly data on population and income per capita based on mid-quarter estimates for each state. Since this data is not as granular as the sales and severity data, I applied the quarter estimates to their corresponding weeks in the combined dataset. I divided the sales by the population of the corresponding state to get sales per capita.

Empirical Model

To test the hypothesis that periods of higher COVID-19 severity indicate higher grocery sales, all things equal, I applied regression analysis with the following empirical model: $sales\ per\ capita_{it} = \beta_0 + \beta_1 severity_{it} + \beta_2 income\ per\ capita_{it} + \beta_3 winter_{it} + \beta_4 spring_{it} + \beta_5 summer_{it} + \alpha_2 state_i + \alpha_3 week_t + \mu_{it}$. The variables in the model are defined in Table 1 and Table 2 contains their summary statistics.

Grocery sales per capita by state is the dependent variable in the model, which is the weekly sales divided by the population of the corresponding state. Since the data is broken out by product category, I ran several regressions with each using the sales for a different category. The five categories of sales used are Dairy, Vegetables, Fruits, Grains, and Meats/Eggs/Nuts.

The main independent variable in the model is the severity variable. To represent severity, the model uses the number of COVID-19 deaths in hundreds. Regression on the model is used to identify if there is a relationship between grocery sales and the severity of the pandemic—indicated by the coefficient on the severity variable. I expect this coefficient to be positive to support my hypothesis because that would indicate that higher severity during the pandemic increases sales.

There are also several control variables in the model. The first is income per capita in USD 1000. This is to account for differences in income across states that could affect sales. The next control variable is season, which is represented through the dummy variables of winter, spring, and summer. The season could have an impact on the availability of products, such as fruit or vegetables. The next two control variables (state and week) also represent vectors of dummy variables; Since the data is broken out by state and week, these are accounted for in the model as well. There are 38 dummy variables for state and 90 dummy variables accounting for the weeks.

Results

The regression results for the models of the five food category sales are summarized in Table 3. In each model, the severity variable has a p-value of less than 0.05, meaning that it is significant at the 5% level across all categories. Additionally, the coefficient variables of severity all have a positive coefficient, which indicates a positive relationship with sales. After calculating the elasticities by dividing the severity mean by the sales per capita mean and then multiplying by the coefficient of severity, the models give the following results: an increase in the number of COVID-19 deaths by 10% increases the dairy sales per capita by 0.01%, the fruit sales per capita by 0.04%, the vegetable sales per capita by 0.03%, the grain sales per capita by 0.03%, and the meats/eggs/nuts sales per capita by 0.02%. This supports my hypothesis that increases in the severity show increases in sales, all things equal. However, the impact is inelastic and relatively small for a 10% increase in deaths across all categories. The elasticities do vary slightly across categories with the effect being the most inelastic for dairy and meats/eggs/nuts compared to fruits, vegetables, and grains. Increasing the number of COVID-19 deaths has a significant, but inelastic, impact on the grocery sales.

The control variables have ranging effects on the dependent variable. Increasing income has a negative impact on the sales per capita for dairy, grains, and meats/nuts/eggs, while it has a positive impact for vegetables. Income does not have a significant impact on fruit sales. The variables for summer and spring have negative impacts on sales across sales of all categories except for meats/eggs/nuts. The variable for winter has negative impacts on sales for dairy and fruits but has a positive impact on vegetables. It does not have a significant impact on sales for grains or meats/eggs/nuts. The dummy variables for state and week are included in all models as well.

The F-statistics for all five models are very good. The significance value for each F-statistic is very close to zero. This indicates that the joint effect of all the variables together in the model is significant. Further, each model has a very high R-squared value, ranging from 0.94 to 0.97. This means that the independent variables explained a large portion of the variance on the dependent variable. However, such high values could also indicate a problem within the model, such as multicollinearity or missing explanatory variables.

Conclusion

Overall, my regression analysis on the pandemic and grocery sales confirmed my hypothesis; Higher severity in COVID-19 causes higher grocery sales, all things equal. This is shown through my regression models reflecting the sales across five different food categories. With the number of COVID-19 deaths indicating severity, there is a positive effect of severity on sales across all categories studied, although the effect is inelastic. This is what I expected given the theory of increasing demand. These results are consistent with prior research conducted on the topic of grocery sales and the pandemic. The general findings have shown that consumers responded to increases in severity during the pandemic with increased demand and expenditure in the grocery realm. Therefore, my findings align with the related literature.

One weakness in my research was compiling data at such a granular level. Since the sales and severity data was weekly, it made finding data for the control variables more difficult. The smallest level of data available for income was quarterly, which is not ideal because that means there was less variation to account for in that variable. Having data at a weekly level was very good for having enough observations, especially due to the shorter timeframe of less than two years for the pandemic, but it made the control variables less effective and more difficult to account for in the models.

One implication for this research is how the grocery industry responds in states of emergency. I did not cover the supply side of this issue in this research, but the increasing demand on groceries during events, such as a pandemic, creates problems for the supply in the grocery industry. It’s important that there is enough food supply to provide for the extra demand in times of hardship.

References

Baarsma, B., & Groenewegen, J. (2021). Covid-19 and the demand for online grocery shopping: Empirical evidence from the Netherlands. *De Economist*, *169*(4), 407–421. https://doi.org/10.1007/s10645-021-09389-y

Baker, S. R., Farrokhnia, R. A., Meyer, S., Pagel, M., & Yannelis, C. (2020). How does household spending respond to an epidemic? consumption during the 2020 COVID-19 pandemic. *The Review of Asset Pricing Studies*, *10*(4), 834–862. https://doi.org/10.1093/rapstu/raaa009

Chang, H. H., & Meyerhoefer, C. D. (2021). Covid-19 and the demand for online food shopping services: Empirical evidence from Taiwan. *American Journal of Agricultural Economics*, *103*(2), 448–465. https://doi.org/10.1111/ajae.12170

Goolsbee, A., & Syverson, C. (2021). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, *193*, 104311. https://doi.org/10.1016/j.jpubeco.2020.104311

Hobbs, J. E. (2020). Food Supply Chains during the Covid-19 pandemic. *Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie*, *68*(2), 171–176. https://doi.org/10.1111/cjag.12237

Table 1

Variable Definitions

Variable Name	Definition
Sales	Weekly sales per capita from a sample of food retailers by state across food categories (Dairy, Vegetables, Fruits, Grains, Meats/Eggs/Nuts)
Severity	Number of weekly COVID-19 deaths in hundreds
Income	Income per capita by state in USD 1000
Winter	1 if the week is in winter, 0 otherwise
Spring	1 if the week is in spring, 0 otherwise
Summer	1 if the week is in summer, 0 otherwise
State	State corresponding to each observation
Week	Week corresponding to each observation

Table 2

Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Sales – Dairy	3.24	0.60	1.97	6.86
Sales – Fruits	0.83	0.15	0.54	2.00
Sales – Vegetables	1.07	0.25	0.57	2.96
Sales – Grains	2.62	0.51	1.81	6.98
Sales – Meats/Eggs/Nuts	4.47	0.82	2.76	8.77
Severity	1.86	3.14	-1.35	37.05
Income per Capita	59.24	8.91	42.70	83.98
Winter	0.16	0.37	0	1
Spring	0.30	0.46	0	1
Summer	0.29	0.45	0	1

Table 3

Regression Results

Model:	Dairy	Fruits	Vegetables	Grains	Meats/Eggs/Nuts
Intercept	3.747 p=0.000***	0.900 p=0.000***	1.098 p=0.000***	2.971 p=0.000***	5.180 p=0.000***
Severity	0.00179 p=0.031*	0.00173 p=0.000***	0.00195 p=0.000***	0.00389 p=0.000***	0.00478 p=0.002**
Income per Capita	-0.0133 p=0.000***	0.000486 p=0.554	0.00341 p=0.017*	-0.0122 p=0.000***	-0.0121 p=0.014*
Winter	-0.526 p=0.000***	-0.0836 p=0.000***	0.0503 p=0.001**	-0.0381 p=0.170	0.0747 p=0.160
Spring	-0.437 p=0.000***	-0.136 p=0.000***	-0.173 p=0.000***	-0.113 p=0.000***	0.379 p=0.000***
Summer	-0.417 p=0.000***	-0.135 p=0.000***	-0.147 p=0.000***	-0.0695 p=0.004**	0.246 p=0.000***
State Included	Yes	Yes	Yes	Yes	Yes
Week Included	Yes	Yes	Yes	Yes	Yes
R^2	0.97	0.95	0.95	0.96	0.94
F-statistic	839.2	499.7	473.3	607.5	431.7
Prob>F-stat	0.000	0.000	0.000	0.000	0.000
Observations	3,549	3,549	3,549	3,549	3,549