Three Essays on Risk Analysis in Agri-Food Supply Chains

A Thesis

Presented in Partial Fulfilment of the Requirements for the

Degree of Master of Science

with a

Major in Applied Economics

in the

College of Graduate Studies

University of Idaho

by

Zhengliang Yang

Approved by:

Major Professor: Xiaoxue Du, Ph.D.

Committee Members: Andres Trujillo-Barrera, Ph.D.; Patrick Hatzenbuehler, Ph.D.

Department Administrator: Christopher McIntosh, Ph.D.

August 2022

Abstract

This thesis consists of three separate chapters, which cover upstream and downstream of the agri-food supply chain. In the first chapter, I revisit the price and volatility transmission mechanism among natural gas, fertilizer, and corn markets between 2011 and 2021. The results turn out natural gas price returns in the short-term are significantly affected by its lagged returns from itself and corn markets, and it will be affected by its lagged return sand fertilizer markets. In the second chapter, I explore the role of the collective reputation of online platforms in agri-food live streaming. The case study uses Taobao Live as an example and I explore the relationship between the collective reputation of online platforms and live streaming revenue. The main result shows that small agri-business owners who want to adopt live streaming to make money should launch their live streaming on a platform with a collective reputation. The third chapter analyzes the photo content in an online review platform. It studies the impact of photo content differences between business owners and customers on restaurant ratings. By constructing a new dataset and building the Latent Dirichlet Allocation model to find the latent of each photo, I found out the photo content in specific topics (the Dish, Protein Food, Dessert) has a positive impact on restaurant rating. It means that business owners can boost their rating by increasing the number of photos regarding these topics.

Acknowledgements

At first, I would like to sincerely thank my three committee members, Dr. Du, Dr. Trujillo-Barrera and Dr. Hatzenbuehler for their dedicated support and guidance. Without your help and recommendations, I can't make it to have my idea about my thesis, give presentation on AAEA meeting, and get offers from Ph.D programmes.

Second, I would like to express the special thanks to Dr. Liang Lu. He makes me realize my career goal and always encourage me to think novel and important research questions.

Third, I would like to appreciate myself for not being depressed and stay healthy both mentally and physically during the pandemic years. And I also want to appreciate the efforts I've been done in last two years.

Finally, I would like to thank my friends and parents. I want to give my special thanks to Ruizhuo Yu for her generous funding and support; Wenqi Wang, Han Sun, and Xiang He for their listening ears; Connor Allen for his kind help and brilliant ideas.

Table of Contents

List of Tables

List of Figures

Chapter 1

Price and Volatility Transmissions between Natural Gas, Fertilizer, and Corn Markets: A Revisit

1.1 Introduction

Commodity markets have experienced intensive price volatility in recent years, especially during the recent months of 2021. Some commodity prices like fertilizer, which were to "rise moderately by 3% "in 2020, subsequently increased sharply in 2021 and reached near-record high prices, achieving unseen high level that since the 2008-09 global financial crisis(WorldBank, 2020). Other commodities have experienced similar price volatility as the fertilizer; e.g. global food prices steeply falling in June and July 2021, and then surging to their highest level in a decade by November, or within three months(Alcorn, 2021). Natural gas prices in the U.S also experienced dramatic fluctuation, reaching their highest price since the 2005 - 06 winter average, at a 7-year high of record global prices(EIA, 2021). These severe price changes not only increase the cost of risk management but also have a negative impact on the economic recovery and growth rate of some countries, especially underdeveloped countries(Jacks et al., 2011).

The goal of this paper is to investigate the price and volatility transmission mechanisms between natural gas, fertilizer, and US corn markets during recent ten years (from 2011 to 2021). A novel study by Etienne et al. (2016) identifies the price volatility transmission mechanism between natural gas, ammonia, and corn prices from 1994 to 2014. They find significant correlations between fertilizer and corn prices and a weak relationship between those markets and the natural gas market. They also find that a unidirectional impact of lagged conditional volatility of fertilizer prices (specifically ammonia prices) positively affects the conditional volatility of corn markets.

However, given the recent intensive price volatility of the last few years, it is not certain whether such price transmission mechanisms among the three markets remains the same or has changed. Moreover, understanding these mechanisms assists the government in making targeted policies and possibly help companies manage potential related risks.

To achieve the objective, a vector error correction model (VECM) and Multivariate Generalized Autoregressive Conditional Heteroskedasticity (MGARCH) framework based on Etienne et al. (2016) is used, but with substantial changes. First, daily frequency data is adopted instead of weekly frequency data, enabling the capture of volatility spillover effects between different markets. Different levels of frequency data may yield inconsistent results, thus working with higher frequency data (i.e., daily) is recommended(Saghaian et al., 2018). Moreover, the sample size is expanded from around 1000 observations (in Etienne et al. (2016)âs research) to more than 2000.

Second, the exchange-traded funds (ETF) is used as a proxy for fertilizer prices. The World Bank Commodity Outlook does not provide US Midwest weekly ammonia prices nor daily frequency of fertilizer prices, thus there is a need to look for variables that can become approximate proxy for fertilizer prices. The ETF index is selected because it can reflect the daily performance and general price trend of the major fertilizer suppliers, and is sensitive to price changes of related raw materials commodities (e.g., natural gas prices). This index seems to reflect changes in fertilizer prices since the companyâs return is largely dependent on the productâs price, assuming a constant capacity in the short run. Moreover, the monthly trends regarding monthly fertilizer prices from the International Monetary Fund (IMF) world primary commodity database are inspected and find that the fertilizer ETF index and fertilizer prices share a similar trend(IMF, 2021), as figure 2 and figure 3 show. Thus fertilizer ETF is a plausible good proxy for fertilizer prices.

Lastly, the sample period is updated to include the latest date (30 November

2021), reflecting the current price transmission mechanism and providing a timely and practical implication.

Results obtained show several similarities and differences compared to Etienne et al. (2016). First, results show that natural gas prices are statistically and significantly affected by their own lagged and a corn lagged term; fertilizer markets are also statistically and significantly affected by natural gas markets. These results are consistent with results from Etienne et al. (2016). However, the relationship between fertilizer prices and corn prices is not found but find different results in that corn prices and the fertilizer markets are negatively affected by the natural gas prices (Etienne et al. (2016)âs results show the positive relation- ships). Third, in the long run, fertilizer is found to be the only statistically significant parameter between adjustment parameters. This is contrary to Etienne et al. (2016)âs results of strong statistical significant relationships for adjustment parameters and fertilizer (ammonia) or corn. Moreover, lagged conditional volatility of corn prices is found to affect the conditional volatility of the natural gas market but not vice versa.

The contribution of this paper lies in two parts: First, our work contributes to existing knowledge of price and volatility transmissions between natural gas, fertilizer,and corn markets by proving their relationship in recent time period. Second, we adopt a new methodology to gather higher frequency data for fertilizer to examine the mechanisms between these market prices in detail.

The rest of the paper organized as follows. The next section presents the recent literature on prices relationships in different markets. Section 3 and section 4 covers the methodology and data, respectively. Section 5 presents the results, and section 6 concludes the paper.

1.2 Literature Review

Recent studies during the past five years of the price relationships between commodity markets of food, energy (such as oil or natural gas), and fertilizer, can be divided into three categories: (i) energy prices and food prices (ii) fertilizer prices and food prices; (iii) energy prices and fertilizer prices. Previous literature has identified the price volatility linkage between energy price, fertilizer prices, and food price (e.g., Etienne et al. (2016)). However, more recent literature on how prices in these three markets are linked remains sparse.

First, some recent studies have pointed out the inter commodity price volatility transmission between energy prices and food prices, but its mechanism remains unclear. On the one hand, some literature points out their unidirectional relationship and proved such relationship in the short-run and long-run, respectively. Shahnoushi et al. (2017) showed that crude oil and gasoline prices have a significant positive impact on food price subgroups such as cereals and meats: with the rise in Di-Ammonium Phosphate (DAP) and Triple Superphosphate prices, the cereals, beverages, and vegetable oil prices increased. Taghizadeh-Hesary et al. (2018) argued that energy price (oil price) has a significant impact on food prices and further find out that the shares of oil prices in agricultural food price volatility are the largest, according to their results of impulse response functions. Similarly, Ji et al. (2018) showed the significant risk spillovers from energy (oil and natural gas) to agricultural commodities (maize, rice, soybean and wheat) by measuring the conditional value-at-risk (CoVaR) and delta CoVaR. The later research also conducted by Taghizadeh-Hesary et al. (2019) indicates that food prices will respond positively to any shock from oil prices. Particularly, oil price movement can explain 64.17% of food price variance. The research from Nwoko et al. (2016) revealed a unidirectional causality with causality running from oil price to food price volatility but not vice versa. Supported their conclusion, Siami-Namini et al. (2019) studied volatility transmission among oil price, exchange rate, and agricultural commodities prices and concluded that volatility in the agricultural commodity returns for most cases is affected by the volatility of the crude oil returns in the post-crisis period. A recent study by Dutta et al. (2021) investigated the correlation between energy price uncertainty and the Malaysian palm oil industry during the 2014 oil price decline and the COVID-19 outbreak. They concluded that oil market volatility negatively impacts palm oil prices and such impact intensified during 2014 and the COVID-19 outbreak.

For short-run or long-run relationships, Ibrahim (2015) studied a case from Malaysia and found that positive oil price exerts significant influences on inflation of food price in the short run, and that there is a significant relation between oil price increases and food price in the long run. Recent work by Radmehr and Rastegari Henneberry (2020) found that both in the short run and long run, food prices increase in response to an increase of energy prices. Chowdhury et al. (2021) found out that the relationship between energy prices and food prices is nonlinear and asymmetric: in the short run, food prices are only affected by positive changes in energy prices while in the long run, both positive and negative changes in energy prices impact food prices. However, Fowowe (2016) argued that structural break cointegration shows no long-run link between energy and food prices. Meanwhile, nonlinear causality tests show no short-run link between energy and food prices. Meyer et al. (2018) focuses on the effects of oil price changes on food prices in oil-exporting developing countries between 2001 and 2014 and find no long-run relation between oil price reduction and food prices. Similarly, Eissa and Al Refai (2019) adopted the nonlinear model to explore the dynamic relationship between oil prices and agricultural commodities (barley, corn and rapeseed oil) from 1990 to 2018 but do not find correlations in the long-run. Roman et al. (2020) also only find the short term linkage of crude oil prices occurred with food, cereal, and oil prices between January 1990 and September 2020.

Conversely, Rezitis (2015) used panel VAR methods and Granger causality tests, with results indicating bidirectional panel causality effects between crude oil prices and international agricultural prices; as well as between US exchange rates and international agricultural prices. Su et al. (2019) investigated causalities between oil and agricultural prices in the global market, and found a bidirectional positive causality between oil and agricultural products prices. The evidence from De Gorter and Just (2008) shows that different agricultural shocks can have different effects on oil price and that corn use in ethanol plays an important role in the impact of corn demand shocks on oil price.

Compared to the studies that explore the relationship between energy prices and food prices, recent research on the correlations between food prices and fertilizer prices are relatively few, despite fertilizer playing an essential role in agricultural production. Some research highlights the important role of fertilizer prices in agricultural commodities prices, but they arrive at mixed results. Dillon and Barrett (2016) found a negligible effect of fertilizer prices in local maize price determination once controlling for changes in global maize prices. However, Ismail et al. (2017) investigated the relationship between price volatility of food and fertilizer and found that fertilizer prices (urea) positively and significantly affect the mean prices of some agricultural commodities, such as rice and sugar. However, the volatility of fertilizer prices is only transmitted on specific products, such as sunflower oil. Kalkuhl et al. (2016) used an empirical model and conclude high fertilizer prices and price risk will substantially decrease the global supply response to higher crop prices.

Finally, some studies explored the relationship between energy prices and fertilizer prices. Chen et al. (2012) evaluate the effect of crude oil price on global fertilizer prices in both the mean and volatility and show that most fertilizer prices are significantly affected by the crude oil price while the volatility of global fertilizer prices and crude oil price from March to December 2008 are higher than in other periods. Results from Sanyal et al. (2015) showed that changes in oil and natural gas prices increased fertilizer prices from June 2007 to June 2008, suggesting that the volatility effects of oil and natural gas prices on fertilizer prices were significant. Wongpiyabovorn (2021), found that natural gas prices strongly influence both ammonia and urea prices during the pre-2010 period.

1.3 Methodology

Following the methodology in (Etienne et al., 2016)âs paper, a time-series approaches is used to examine the linkages between natural gas, fertilizer, and agricultural commodity prices. Several statistical properties are considered before modeling. First, time-series data with nonstationarity properties may result in spurious regression results when using normal approaches (OLS, for instance).. Therefore, a first-order difference is applied to address the problem. Second, cointegration may appear among nonstationary variables. Cointegration refers to nonstationary variables having a long-run relationship. Therefore, these variables are to be modeled considering both short-and long-run dynamics. Finally, time-varying and clustering volatility of time series data will violate the standard regression model assumption of being independently and identically distributed (*iid*). This problem is addressed with ARCH and GARCH models.

1.3.1 The VECM-MGARCH Model

In a similar path, a vector-error correction multivariate GARCH model (VECM-MGARCH) is considered in equation (1) in the case of cointegration, in order to examine the price and volatility spillovers between natural gas, fertilizer, and agricultural commodity prices:

$$\Delta P_t = \sum_{j=1}^{p} \beta_j \Delta P_{t-j} + \alpha ECT_{t-1} + \varepsilon_t \tag{1.1}$$

$$\varepsilon_t | I_{t-1} \ (0, H_t)$$

Where ΔP_t is a 3×1 vector, that includes the first difference of prices of natural gas, fertilizer, and corn prices, β_j is a 3×3 matrix of autoregressive coefficients that models the short-run dynamics between the three markets, p is the lag length; ECT is the error correction term that measure the three markets' long-run relationship; α is a 3×1 vector, indicating the the speed of adjustment response in long-run. ε_t is a 3×1 vector of error term. H_t is a variance-covariance matrix of error term the conditioning on past information I_{t-1}.

1.3.2 The BEKK Model

Next, the Baba-Engle-Kraft-Kroner (BEKK) model is used to specify the H_t. . The advantages of adopting the BEKK model (see Baba et al., 1990 and Engle and Kroner, 1995 for example) are two parts: (i) it directly estimates the conditional covariance matrix; (ii) it ensures the positive definiteness of the conditional covariance matrix and reduces the dimension of the parameter vector in H_t. Equation (2) is the BEKK model with one lag:

$$H_t = CC' + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'H_{t-1}B, \tag{1.2}$$

where C is a 3 Ã 3 lower triangular matrix that corresponds to the constant, $\varepsilon_{t-1}\varepsilon'_{t-1}$ are the squared lagged errors, A is a 3 Ã 3 matrix of ARCH parameters containing elements measuring the degree of innovation from one market to another, and B is a 3 Ã 3 matrix of GARCH parameters that shows the persistence in conditional volatility between two markets. After setting up the model, we can estimate them by maximum likelihood procedures (see Seo, 2007 and Serra et al., 2011).

1.4 Data

The dataset used comes from different sources, with some being the same sources as Etienne et al. (2016). The fertilizer price is represented by the Solactive Global Fertilizers/Potash Total Return Index, which tracks the performance of the largest and most liquid listed companies globally that are active in some aspect of the fertilizer industry. Such ETF index may reflect timely price changes of the fertilizer industry, and be a good proxy for fertilizer prices. Besides, non-energy commodities in commodity index have experienced lower volatility. This data is collected from Yahoo Finance.com. The ammonia price in the US Midwest is not used as fertilizer price because the dataset from the World Bank Commodity Markets does not provide it at the daily level. For agricultural market commodities, corn is considered, since it is a dominant and common global crop and one of most reliant crops on fertilizer. Yahoo Finance provides the history data. The natural gas prices are the rolling prices of futures contracts traded on the New York Mercantile of Exchange (NYMEX), obtained from the Energy Information Administration (EIA). The three datasets are merged at daily frequencies by matching the date and removing some null values. The sample period starts from 25 May 2011 to 30 Nov 2021, since the inception date of fertilizer ETF is 25 May 2011. There are 2,219 observations used in this study.

From Figure 1, which plots the three different prices series, we see several important patterns: First, natural gas prices show a general increase or decrease trend during some periods before Jan 2020. This indicates an increasing trend (May 2011-May 2012), constantly rising with some fluctuations (May 2012 - March 2014), sharply dropping (March 2014 - January 2016), increasing again (January 2016 - November 2018) and followed by descending (November 2018- January 2020), respectively. Second, corn prices remained relatively stable in some periods (November 2014 to December 2019), except for a boom from May 2012 to July 2012 and a drop from May

2013 to September 2014. Thirdly, similar to natural gas price fluctuation patterns, fertilizer prices show a increasing or decreasing patterns at certain times, such as falling from January 2011 to January 2014; and sharply increasing from July 2018 to November 2018.

We also observe a similar volatility pattern between fertilizer prices and natural gas prices after 2014, which is different from findings in Etienne et al. (2016). Finally, a significantly common upward-trending pattern is observed in natural gas and corn prices in January 2020. Previous research has explained that such price trends for natural gas may be the consequence of preceding warm winters; thus market players have been less optimistic and more cautious about future investments as they had already expected lower sales(Nyga-Łukaszewska and Aruga, 2020).

Figure 2 plots the one-month rolling coefficient of variation(standard deviation divided by mean) for natural gas, fertilizer, and corn prices, respectively. The coefficient of variation is used to compare the volatility of these price series. After adopting data at the daily level, several results are found to be different from Etienne et al. (2016). First, the range of coefficient of variation was significantly smaller than from Etienne et al. (2016). This may mean that estimates of the relationship between different markets may reflect more accurately the volatility of prices because of daily level data . Second, in contrast to the positive correlation of volatility of the three prices in previous research, Figure 2 suggests that they only show correlations in a specific period, such as November 2011 to January 2012 and January 2020 to January 2021.

For natural gas prices from December 2015 to January 2016, significant price volatility is experienced compared to other commodities. The coefficient of variation continued to decline during the pandemic of COVID-19. For the fertilizer price, its coefficient of variation shows peaks in December 2019 followed by three months of decline till March 2020. For corn prices, the coefficient of variation is stable (around

0.005), and its range of variation is the smallest among the three markets. Meanwhile, unlike the cornâs coefficient of variation in previous research, in this study it peaked only in July 2013 and quickly returned to a smaller value (0.005) within a month. Although the coefficient of corn prices fluctuates due to the pandemic, the degree of variation is less than 0.015.

For stationarity testing, The Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests are applied to the three price series. The lag length of the ADF test is selected according to the Akaike Information Criterion (AIC). All prices have been adjusted to logarithmic scales. Results show a failure to reject the null hypothesis of nonstationarity at the five percent significance level for any three log price series. All prices return at first differences are stationary at the five percent significant level.

Table 1 shows the descriptive statistics of price returns, multiplied by 100, for each of the three commodities. In panel (a), the average return of three commodities is observed to be similar to each other, all being close to zero. Although natural gas has the highest average price return, the average return difference between the natural gas and the other two commodities is less than 0.1. In contrast to Etienne et al. (2016), the price return of the fertilizer market is the most volatile commodity compared to other commodities, which is consistent with the patterns observed in figure 2. Results from the Jaque-Bera test and excess Kurtosis indicate that all return series may not be normally distributed. The statistically significant results in panel (b) and panel (c) are the same as the Etienne et al. (2016), which shows the rejection of the null hypothesis that no autocorrelation for both returns and squared returns.

To determine if a potential cointegrating relationship exists, the Johansen maximum likelihood test is applied with results in Table 2. Based on AIC, a maximum lag of 26 (26 days) is selected. Results suggest rejecting the null hypothesis of no cointegrating vector between three markets at the five percent significance level. However, one cointegrating vector exists for three commodity prices as the trace statistic is between the five percent critical value and one percent critical value. As a result, a vector error correction model (VECM) should be employed to account for the long-run relationship.

1.5 Estimation Results

Table 3 provides estimation results that evaluate the extent of the price of dependency and transmission across natural gas, fertilizer, and maize markets, using the same approach as the VECM-MGARCH model in Etienne et al.(2016) are presented. We choose three as the lag term based on AIC.

Panel (a) of Table 3 presents the estimated error correction term. Results suggest that the fertilizer prices are positively correlated with corn prices. This is consistent with the result from Etienne et al. (2016) which can be explained by the fact the farmers have an incentive to use more fertilizer when corn prices increase. As a result, it will increase the fertilizer demand and fertilizer prices. Besides, the negative shocks to the supply of the raw materials of fertilizer production (e.g., oil or natural gas) can also lead to higher fertilizer prices, resulting in lower corn production and upward pressure on corn prices. However, fertilizer prices are negatively correlated with the natural gas prices, which are different from the previous results of positive correlations between fertilizer prices (ammonia price) and natural gas prices. From the estimated parameters, in the long run, we can know that a one percent increase in corn price and natural gas price will result in a 0.45 percent increase and 0.10 percent decrease of fertilizer price, respectively[1]. By comparing the elasticity response between natural gas and corn, corn plays a more significant role in fertilizer price, confirming previous research results.

The last line of panel (b) in Table 3 presents long-run dynamics between the

[1]All prices are transformed in the logarithmic format, so the estimated parameters in long-run are the elasticity of prices.

three markets. First, the only statistically significant parameters are for fertilizer prices, which is contrary to the Etienne et al. (2016) result of strong statistical significant relation ships between adjustment parameters and ammonia and corn. Second, fertilizer prices are found to have a much lower speed in responding to the disequilibrium in long-run parity than the speed in Etienne et al. (2016) (0.0017 percent per day v.s. 32.6 percent per week). Third, the corn prices lack response to the long-run equilibrium after adopting daily frequency data, which may mean natural gas and corn commonly lead price changes in this three-commodity system and fertilizer prices are the only one making the adjustment to the disequilibrium.

The rest of panel (b) in Table 3 presents the short-run interactions of three return series. In column (1), natural gas price returns are significant negatively affected by its lagged returns from itself and corn markets, but not by lagged returns in the fertilizer market, which is in line with results from Etienne et al. (2016). In column (2), there is a short-run significant and negative effect of natural gas prices return on the fertilizer prices return, which is different from Etienne et al. (2016). The new result may be because fertilizer manufacturers cannot change their original plans for increasing production in the short run (especially in the daily horizon), so fertilizer supply increases and fertilizer prices decrease in the short term when natural gas prices rise in short-run. In column (3), no statistically significant correlations between corn prices return and natural gas is found, which is consistent with Etienne et al. (2016)âs result that natural gas prices have no impact on corn prices return. However, corn prices seem to negatively respond to changes in fertilizer (ammonia) prices, which is different from results in this study (e.g. no significant correlations are found). No statistical correlations between corn prices returns and natural gas or fertilizer prices returns are found, which may be due to some other impact from macroeconomic factors not captured by the model.

Panel (c) of Table 3 shows the results from MGARCH estimation. The diagonal elements in matrix A and B measure the volatility persistence of the three markets and how shocks originating in one market affect each oneâs conditional volatility, respectively. In matrix A, the significant and no-zero diagonal ($a_{ii} \neq 0$) terms show the strong own volatility spillover in all three markets, which is consistent with the results from Etienne et al. (2016). Similarly, the diagonal terms ($b_{jj} \neq 0$) in matrix B indicates that conditional volatility significantly depends on its own lagged volatility, consistent with the previous study.

The off-diagonal elements in two matrix panels (c) measure the cross-market volatility dynamics. Given the BEKK formulation in equation (2), the off-diagonal elements in matrix B measure the direct persistence of volatility. In matrix A, they measure how the effects of lagged innovation originated from one market directly on the other. Only b_{31} term is statistically significant, which means the conditional price volatility of natural gas can be affected by the lagged volatility in corn market but not the vice versa. This result is different from Etienne et al. (2016), where natural gas seems to behave completely independent from the other two markets. Moreover, shocks from natural gas fail to explain either corn or fertilizer volatility as in Etienne et al. (2016) This illustrates the results from using daily frequency data, which identifies potential price volatility relationships between corn and natural gas markets. Regarding the lagged innovations impact, a significant but negative effect is found (Etienne et al. (2016)âs result is positive) from fertilizer to the corn prices since $a_{23} \neq 0$. However, the reverse spillover from corn to fertilizer($a_{32} = 0$) is not found, in similar results as Etienne et al. (2016).

To interpret and compare the results from VECM-MGARCH, figure 4 presents the results from impulse response functions. Following the approach of Etienne et al. (2016) and Gardebroek and Hernandez (2013), this study generates the volatility impulse response functions to a shock originated in another market that increases its conditional volatility by 1 percent. Due to daily frequency data, the longer response steps up to 60 periods is set. The results from impulse response functions are consistent with the discussion in the previous part of this paper. Shocks generated in one market may have slight responses but not respond in longer periods. For example, the response of natural gas volatility to shocks from fertilizer market or corn markets are slightly influenced positively (corn markets) or negatively (fertilizer markets) at the beginning (day 4 or day 5) but back to 0 after the peak. It shows that as an indispensable material for production and life, natural gas price characterized as rigidity.

Except for comparing the results with Etienne et al. (2016), we also compared our results with several studies on the prices relationship between these three markets and found some consistencies and differences. First, our conclusion support the result from Eissa and Al Refai (2019) that energy prices (oil prices) and agricultural commodity prices (barley, corn and rapeseed oil) may not have long-run relationship. Contrary to our finding, however, Koirala et al. (2015) show that agricultural commodity and energy future prices are highly correlated and exhibit positive and significant relationship. The results from Fernandez-Perez et al. (2016) and Lucotte (2016) also show strong co-movement between oil and agricultural commodity prices instead of the unilateral relationship in our study. In addition, our findings point out that corn prices may be one of the factors influencing natural gas prices, which is different from recent findings from Ferreira et al. (2022) that indicating the influence of the other energy products price (e.g., the diesel) on natural gas prices. Second, our result is different from Dillon and Barrett (2016) that found minor effect of fertilizer prices on local corn price, instead, we do not find the significant impact from fertilizer prices on corn prices. Third, our result shows the natural gas prices has negative impact on fertilizer prices in short-run but not vice versa, which supports the result from Sanyal et al. (2015).

1.6 Conclusions

In this paper, the work from Etienne et al. (2016) is revisited to determine the latest perspective about price and volatility transmission between natural gas, fertilizer, and corn markets. Following a similar methodology (the Vector Error Correction Model and Multivariate Generalized Autoregressive Heteroskedasity), daily frequency data is used, enabling to capture daily volatility between these commodities. Moreover, the ETF index is used as a proxy for fertilizer prices. Contrary to previous findings that fertilizer and corn prices have significant relationships and correlations between natural gas prices and fertilizer (or corn markets), no linkage between the natural gas prices and fertilizer prices or the fertilizer markets and corn markets during 2011-2021 are found. Additionally, fertilizer prices are found to react only to deviation from the long-run parity. Regarding conditional volatility, the negative condition volatility from corn markets on natural gas is found but not vice versa, and no other effects from one market to the other are found.

This paper offers a comprehensive analysis of the association of natural gas, fertilizer, and corn markets so the results have practical implications for investors and policy makers. The short-run prices relationships showed by these three markets enable policymakers to develop specific policies and investors to predict trends in the prices of these commodities. Policymakers need to pay attention to the growing volatility among the three markets, especially smallholders who may not have the means and capacity to optimally balance their risks (Aderajew et al., 2020). Meanwhile, the price volatility transmission pattern in three markets enable policymakers to choose to take into account about the effects of price volatility in certain markets when formulating relevant policies. By referring such information, investors can make appropriate investment decisions and reduce losses caused by potential risks.

Although we have some different results from the Etienne et al. (2016), we do not know what factors caused such a difference in results, prior to Etienne et al. (2016)'s work, a large body of literature focused on exploring reasons for the price transmission. For instance, Gilbert (2010) argues that the agricultural price booms in 2006-2011 is attributed to highlight the demand from developing countries; Abbott et al. (2008) showed that crude oil prices, exchange rates, growing demand are driving force for increasing food price; Baek and Koo (2010) found that exchange rate play the key roles in determining the short- and long-run movement of U.S. food prices. All of these factors may have an explanatory role in our results, but we do not know exactly which factors are at play. Particularly, the COVID-19 epidemic has led to intricate price relationships in these commodity markets, making it difficult to explore the relational movements between these markets. As a result, this will be left to future research.

For the future work, a conceptual work on understanding the mechanism behind the price transmission, such as supply chain issues (Lu et al., 2021; Lu and Winfree, 2021; Reardon et al., 2021), market arrangements like the role of contracts and insurance (Du et al., 2017), comparing the price transmission pattern of other agricultural commodities in different countries(Bekkers et al., 2017) especially when price information is sparse (Hatzenbuehler et al., 2021), as well as adoption of big data methodologies (Lu et al., 2022) will be needed.

Table 1.1: Descriptive Statistics of Natural Gas, Fertilizer, and Corn Returns (2011-2021)

	Natural Gas Return	**Fertilizers Return**	**Corn Return**
Panel (a) Summary Statistics			
Observations	2,218	2,218	2,218
Mean	0.00	-0.03	-0.01
Std. Dev	3.27	6.67	1.86
Minimum	-18.05	-35.66	-25.2
Maximum	19.8	40.55	8.9
Skewness	0.03***	0.28***	-0.98***
Excess Kurtosis	6.40***	5.99***	17.1***
Jarque-bera	1071.21***	856.50***	19000***
Panel (b) Return correlation coefficients			
AC(1)	-0.0558***	-0.249***	0.0277
AC(2)	0.00305	-0.286***	-0.0387**
Ljung-Box(6)	11.03**	155.62***	13.38***
Ljung-Box(12)	21.85***	169.10***	19.76**
Panel (c) Squared return correlation			
AC(1)	0.141***	0.146***	0.00963
AC(2)	0.0710***	0.00525	0.0311**
Ljung-Box(6)	185.28***	66.74***	12.5**
Ljung-Box(12)	272.93***	106.79***	17.87
Panel (d) Unit root test results			
ADF	-49.73***	-60.69***	-45.76***
PP	-49.79**	-65.06***	-45.75***

Note: returns are calculated as $log(p_t/p_{t-1}) \times 100$. ADF and PP stand for augmented Dickey-Fuller and Phillips Perron tests, respectively. AC is autocorrelation coefficient. Ljung-Box is the test for autocorrelation of the returns and squared returns series and computed with 6 and 12 lags, respectively. t-statistics denoted as $***p{<}0.01, **p{<}0.05, *p{<}0.1$, respectively. âââ denoting the rejection of the null hypotheses (normality, no autocorrelation, unit root, non-stationarity, and conditional homoscedasticity at the 1% significance level).

Table 1.2: Johansen Maximum Likelihood Test of Cointegrating Relationship between Natural Gas, Fertilizer, and Corn Prices (2011-2021)

Maximum rank (K)	Log-likelihood	Eigenvalue	Trace statistic	5% critical value	1% critical value
0	13058.74		35.41	29.68	35.65
1	13069.28	0.01	14.33	15.41	20.04
2	13073.91	0.004	5.09	3.76	6.65
3	13076.45	0.002			

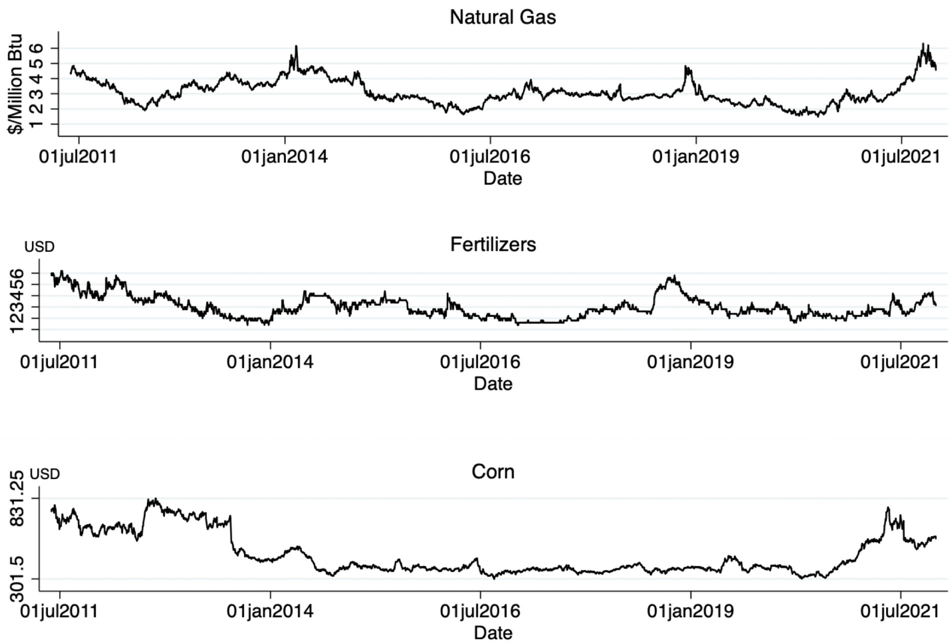
Note: the null hypothesis for each row is that the three prices have a maximum of K cointegrating vector (measured in logrithimic format). When the trace statistic reaches the crucial value, the null hypothesis is rejected.

Table 1.3: VECM-MGARCH Model Estimation Results(2011-2021)

Panel (a) Error Correction Term			
$ECT_t = .004P(NG)_t - 0.048***p(Fertilizer)_t + .026***p(Corn)_t + 0.104$			
Panel (b) Vector Error Correction Model			
	Natural Gas($i=1$)	Fertilizer($i=2$)	Corn($i=3$)
$\Delta p(NG)_{t-1}$	-0.709***	-0.0599*	0.00275
	(0.0201)	(0.0339)	(0.0115)
$\Delta p(NG)_{t-2}$	-0.338***	-0.0310	-0.0141
	(0.0200)	(0.0338)	(0.0115)
$\Delta p(Fertilizer)_{t-1}$	-0.0178	0.239***	0.00171
	(0.0205)	(0.0345)	(0.0117)
$\Delta p(Fertilizer)_{t-2}$	-0.0202	0.0712***	0.00882
	(0.0126)	(0.0213)	(0.00722)
$\Delta p(Corn)_{t-1}$	0.0848**	-0.0113	-0.619***
	(0.0351)	(0.0592)	(0.0201)
$\Delta p(Corn)_{t-2}$	0.0494	0.0141	-0.330***
	(0.0351)	(0.0592)	(0.0201)
ECT_{t-1}	0.0113	-0.0017***	-0.00274
	(0.0267)	(0.0451)	(0.0153)
Panel (c) Vector Error Correction Model			
$H_t = CC' + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'H_{t-1}B,$			
c_{1i}	0.47***		
	(0.088)		
c_{2i}	0.975	2.41***	
	(0.634)	(0.368)	
c_{3i}	0.142***	-0.063	0.247***
	(0.072)	(0.070)	(0.057)
a_{1i}	0.228***	-0.034	0.001
	(0.022)	(0.062)	(0.009)
a_{2i}	0.011	0.480***	-0.012**
	(0.014)	(0.040)	(0.007)
a_{3i}	0.058	0.015	0.202***
	(0.036)	(0.115)	(0.021)
b_{1i}	0.968***	0.007	0.001
	(0.006)	(0.024)	(0.003)
b_{2i}	-0.018	0.808***	0.005
	(0.011)	(0.035)	(0.004)
b_{3i}	-0.029***	0.013	0.966***
	(0.013)	(0.059)	(0.006)

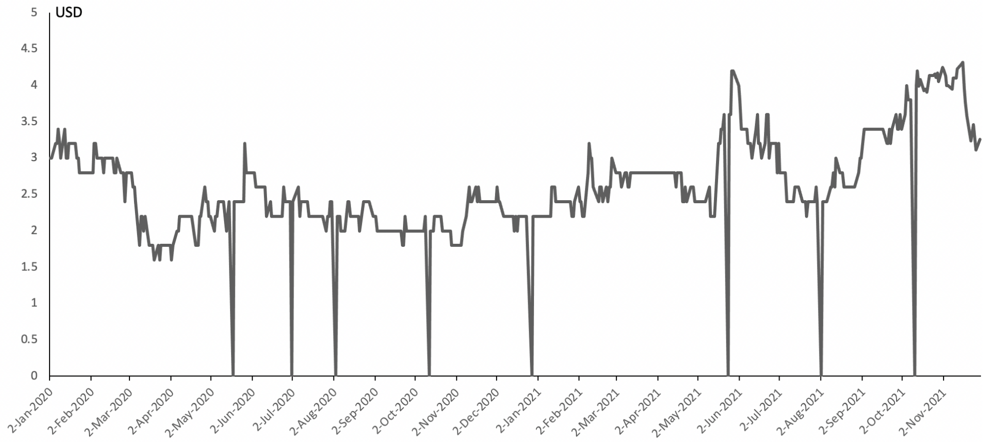
Note: This table presents the results of the VECM-MGARCH model defined in equations (1)â(2). Prices are measured in logrithimic format. a_{ij} and b_{ij} measure the direct impact of lagged volatility and shocks originated in market i affect the conditional volatility in market j, respectively. Standard deviation in parentheses. One, two and three asterisks represent statistical significance at 10, 5, and 1 percent, respectively.

Figure 1.1: Plots of natural gas, fertilizer, and corn prices (2011-2021)



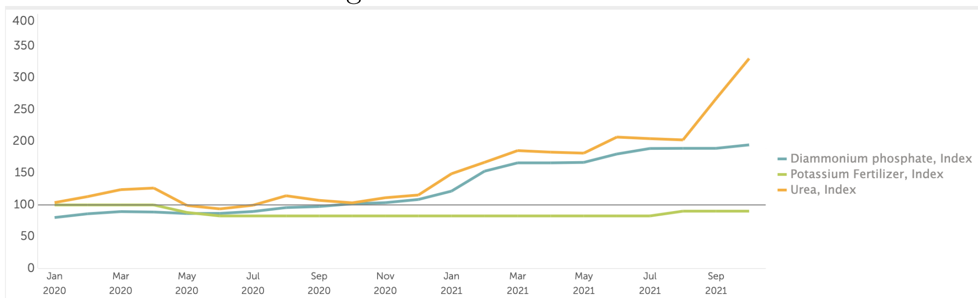
Source: Energy Information Administration and Yahoo Finance

Figure 1.2: Exchange Traded Funds (ETF) of Fertilizer



Source: Yahoo Fiance, from Jan 02 2020 to Nov 30 2021

Figure 1.3: Fertilizer Index



Source: International Monetary Fund Commodity Database

Figure 1.4: Coefficient of variation for natural gas, ammonia, and corn prices

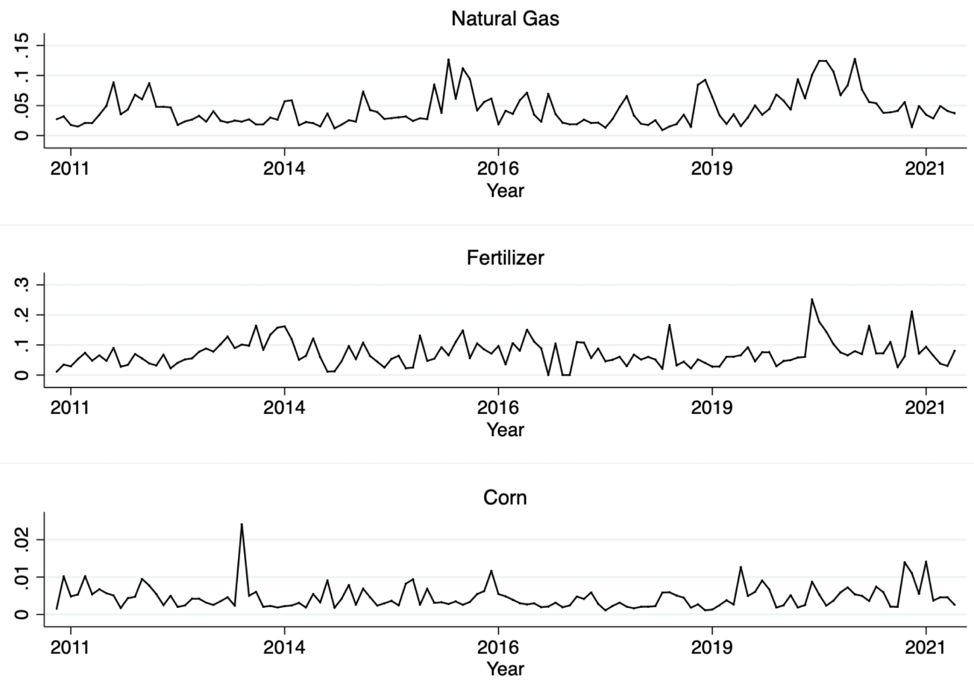
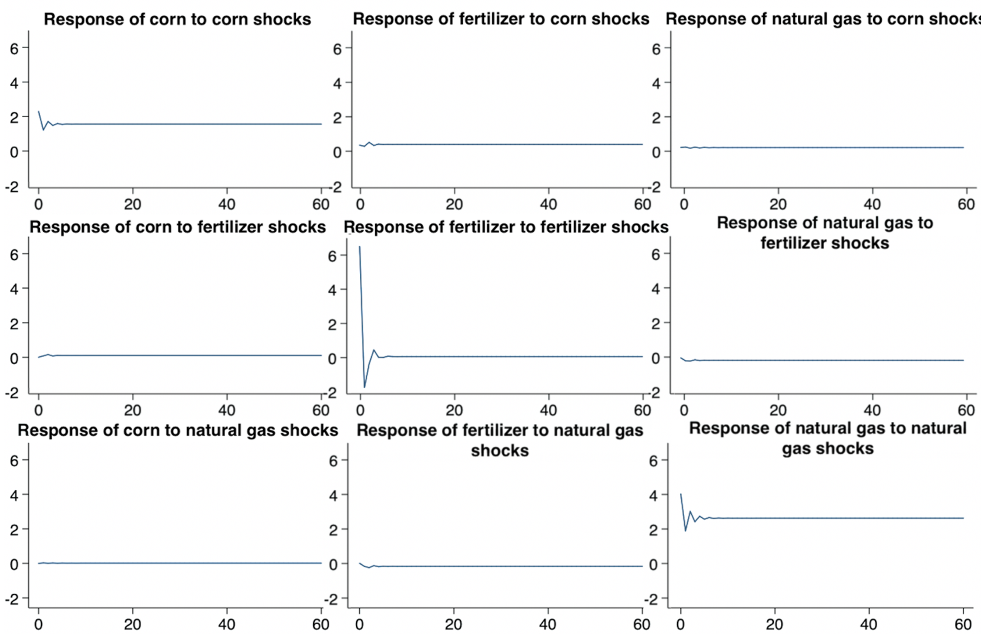


Figure 1.5: Plot of Impulse Response Function



References

Abbott, P. C., Hurt, C., and Tyner, W. E. (2008). What's driving food prices? Technical report.

Aderajew, T. S., Du, X., Pennings, J. M., and Trujillo-Barrera, A. (2020). Farm-level risk-balancing behavior and the role of latent heterogeneity. *Journal of Agricultural and Resource Economics*, 45(2):265–281.

Alcorn, C. (2021). Global food prices surge to their highest level in a decade. `https://www.cnn.com/2021/06/04/business/inflation-food-prices/index.html`.

AliResearch (2020). 2020 taobao live streaming economicv. `https://www.cbndata.com/report/2219/detail?isReading=report&page=1`.

Amblee, N. and Bui, T. (2008). Can brand reputation improve the odds of being reviewed on-line? *International Journal of Electronic Commerce*, 12(3):11–28.

Anderson, M. and Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, 122(563):957–989.

Aula, P. (2010). Social media, reputation risk and ambient publicity management. *Strategy & leadership*.

Baba, Y., Engle, R. F., Kraft, D. F., and Kroner, K. F. (1990). Multivariate simultaneous generalized arch. *Manuscript, University of California, San Diego, Department of Economics*.

Baek, J. and Koo, W. W. (2010). Analyzing factors affecting us food price inflation. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 58(3):303–320.

Bekkers, E., Brockmeier, M., Francois, J., and Yang, F. (2017). Local food prices and international price transmission. *World Development*, 96:216–230.

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS quarterly*, pages 351–370.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Brzustewicz, P. and Singh, A. (2021). Sustainable consumption in consumer behavior in the time of covid-19: Topic modeling on twitter data using lda. *Energies*, 14(18):5787.

Burton, S., Howlett, E., and Tangari, A. H. (2009). Food for thought: how will the nutrition labeling of quick service restaurant menu items influence consumersâ product evaluations, purchase intentions, and choices? *Journal of Retailing*, 85(3):258–273.

Cai, R. and Chi, C. G.-Q. (2021). Pictures vs. reality: roles of disconfirmation magnitude, disconfirmation sensitivity, and branding. *International Journal of Hospitality Management*, 98:103040.

Chen, C.-D., Zhao, Q., and Wang, J.-L. (2020). How livestreaming increases product sales: role of trust transfer and elaboration likelihood model. *Behaviour & Information Technology*, pages 1–16.

Chen, P.-Y., Chang, C.-L., Chen, C.-C., and McAleer, M. (2012). Modelling the effects of oil prices on global fertilizer prices and volatility. *Journal of Risk and Financial Management*, 5(1):78–114.

Chowdhury, M. A. F., Meo, M. S., Uddin, A., and Haque, M. M. (2021). Asymmetric effect of energy price on commodity price: New evidence from nardl and time frequency wavelet approaches. *Energy*, 231:120934.

Clement Addo, P., Fang, J., Asare, A. O., and Kulbo, N. B. (2021). Customer engagement and purchase intention in live-streaming digital marketing platforms. *The Service Industries Journal*, pages 1–20.

De Gorter, H. and Just, D. R. (2008). " water" in the us ethanol tax credit and mandate: Implications for rectangular deadweight costs and the corn-oil price relationship. *Review of Agricultural Economics*, 30(3):397–410.

Dillon, B. M. and Barrett, C. B. (2016). Global oil prices and local food prices: Evidence from east africa. *American Journal of Agricultural Economics*, 98(1):154–171.

Du, X., Khanna, M., Lu, L., Yang, X., and Zilberman, D. (2017). Contracting in the biofuel sector. In *Handbook of Bioenergy Economics and Policy: Volume II*, pages 401–425. Springer.

Duarte Alonso, A., O'neill, M., Liu, Y., and O'shea, M. (2013). Factors driving consumer restaurant choice: An exploratory study from the southeastern united states. *Journal of Hospitality Marketing & Management*, 22(5):547–567.

Dutta, A., Bouri, E., Saeed, T., and Vo, X. V. (2021). Crude oil volatility and the biodiesel feedstock market in malaysia during the 2014 oil price decline and the covid-19 outbreak. *Fuel*, 292:120221.

EIA (2021). Energy information administration (eia) forecasts u.s. winter natural gas bills will be 30% higher than last winter. `https://www.eia.gov/todayinenergy/detail.php?id=50076`.

Eissa, M. A. and Al Refai, H. (2019). Modelling the symmetric and asymmetric relationships between oil prices and those of corn, barley, and rapeseed oil. *Resources Policy*, 64:101511.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric theory*, 11(1):122–150.

Etienne, X. L., Trujillo-Barrera, A., and Wiggins, S. (2016). Price and volatility transmissions between natural gas, fertilizer, and corn markets. *Agricultural Finance Review*.

Fafchamps, M., Hill, R. V., and Minten, B. (2008). Quality control in nonstaple food markets: evidence from india. *Agricultural Economics*, 38(3):251–266.

Fernandez-Perez, A., Frijns, B., and Tourani-Rad, A. (2016). Contemporaneous interactions among fuel, biofuel and agricultural commodities. *Energy Economics*, 58:1–10.

Ferreira, P., Almeida, D., Dionísio, A., Bouri, E., and Quintino, D. (2022). Energy markets–who are the influencers? *Energy*, 239:121962.

Fowowe, B. (2016). Do oil prices drive agricultural commodity prices? evidence from south africa. *Energy*, 104:149–157.

Frick, B. and Simmons, R. (2013). The impact of individual and collective reputation on wine prices: empirical evidence from the mosel valley. *Journal of Business Economics*, 83(2):101–119.

Gan, Q., Ferns, B. H., Yu, Y., and Jin, L. (2017). A text mining and multidimensional sentiment analysis of online restaurant reviews. *Journal of Quality Assurance in Hospitality & Tourism*, 18(4):465–492.

Gardebroek, C. and Hernandez, M. A. (2013). Do energy prices stimulate food price volatility? examining volatility transmission between us oil, ethanol and corn markets. *Energy economics*, 40:119–129.

Gerten, J. and Topolinski, S. (2019). Shades of surprise: Assessing surprise as a function of degree of deviance and expectation constraints. *Cognition*, 192:103986.

Gilbert, C. L. (2010). How to understand high food prices. *Journal of agricultural economics*, 61(2):398–425.

Grossman, D. A. and Frieder, O. (2004). *Information retrieval: Algorithms and heuristics*, volume 15. Springer Science & Business Media.

Grunert, K. G., Bredahl, L., and Brunsø, K. (2004). Consumer perception of meat quality and implications for product development in the meat sectorâa review. *Meat science*, 66(2):259–272.

Ha, J., Park, K., and Park, J. (2016). Which restaurant should i choose? herd behavior in the restaurant industry. *Journal of foodservice business research*, 19(4):396–412.

Hatzenbuehler, P. L., Du, X., and Painter, K. (2021). Price transmission with sparse market information: The case of united states chickpeas. *Agribusiness*, 37(3):665–682.

Heng, Y., Gao, Z., Jiang, Y., and Chen, X. (2018). Exploring hidden factors behind online food shopping from amazon reviews: A topic mining approach. *Journal of Retailing and Consumer Services*, 42:161–168.

Hlee, S., Lee, J., Yang, S.-B., and Koo, C. (2019). The moderating effect of restaurant type on hedonic versus utilitarian review evaluations. *International Journal of Hospitality Management*, 77:195–206.

Hu, L., Sun, A., and Liu, Y. (2014). Your neighbors affect your ratings: on geographical neighborhood influence to rating prediction. In *Proceedings of the 37th*

international ACM SIGIR conference on Research & development in information retrieval, pages 345–354.

Hu, M. and Chaudhry, S. S. (2020). Enhancing consumer engagement in e-commerce live streaming via relational bonds. *Internet Research*.

Huang, J., Rogers, S., and Joo, E. (2014). Improving restaurants by extracting subtopics from yelp reviews. *iConference 2014 (Social Media Expo)*.

Ibrahim, M. H. (2015). Oil and food prices in malaysia: a nonlinear ardl analysis. *Agricultural and Food Economics*, 3(1):1–14.

IMF (2021). International monetary fund primary commodity prices. `https://data.imf.org/?sk=471DDDF8-D8A7-499A-81BA-5B332C01F8B9&sId=1547558078595`.

Ismail, A., Ihsan, H., Khan, S. A., and Jabeen, M. (2017). Price volatility of food and agricultural commodities: A case study of pakistan. *Journal of Economic Cooperation & Development*, 38(3):77–120.

Jacks, D. S., O'rourke, K. H., and Williamson, J. G. (2011). Commodity price volatility and world market integration since 1700. *Review of Economics and Statistics*, 93(3):800–813.

Ji, Q., Bouri, E., Roubaud, D., and Shahzad, S. J. H. (2018). Risk spillover between energy and agricultural commodity markets: A dependence-switching covar-copula model. *Energy Economics*, 75:14–27.

Jia, S., Zhang, X., Wang, X., and Liu, Y. (2018). Fake reviews detection based on lda. In *2018 4th International Conference on Information Management (ICIM)*, pages 280–283. IEEE.

Jia, S. S. (2021). Analyzing restaurant customersâ evolution of dining patterns

and satisfaction during covid-19 for sustainable business insights. *Sustainability*, 13(9):4981.

Jurafsky, D., Chahuneau, V., Routledge, B. R., and Smith, N. A. (2014). Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*.

Kalkuhl, M., von Braun, J., and Torero, M. (2016). Volatile and extreme food prices, food security, and policy: an overview. *Food price volatility and its implications for food security and policy*, pages 3–31.

Kang, K., Lu, J., Guo, L., and Li, W. (2021). The dynamic effect of interactivity on customer engagement behavior through tie strength: Evidence from live streaming commerce platforms. *International Journal of Information Management*, 56:102251.

Klostermann, J., Plumeyer, A., Böger, D., and Decker, R. (2018). Extracting brand information from social networks: Integrating image, text, and social tagging data. *International Journal of Research in Marketing*, 35(4):538–556.

Koirala, K. H., Mishra, A. K., D'Antoni, J. M., and Mehlhorn, J. E. (2015). Energy prices and agricultural commodity prices: Testing correlation using copulas method. *Energy*, 81:430–436.

Li, Y. and Peng, Y. (2021). What drives gift-giving intention in live streaming? the perspectives of emotional attachment and flow experience. *International Journal of Human–Computer Interaction*, pages 1–13.

Lian, J., Zhang, F., Xie, X., and Sun, G. (2017). Restaurant survival analysis with heterogeneous information. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 993–1002.

Lin, L., Gang, R., Hong, T., and Yang, S. B. (2019). Exploring simultaneous presentation in online restaurant reviews: An analysis of textual and visual content. *Asia Pacific Journal of Information Systems*, 29(2):181–202.

Lin, Z., Li, D., Janamanchi, B., and Huang, W. (2006). Reputation distribution and consumer-to-consumer online auction market structure: an exploratory study. *Decision Support Systems*, 41(2):435–448.

Liu, H., Feng, S., and Hu, X. S. (2022). Process vs. outcome: Effects of food photo types in online restaurant reviews on consumersâ purchase intention. *International journal of hospitality management*, 102:103179.

Lu, L., Nguyen, R., Rahman, M. M., and Winfree, J. (2021). Demand shocks and supply chain resilience: An agent based modelling approach and application to the potato supply chain. Technical report, National Bureau of Economic Research.

Lu, L., Tian, G., and Hatzenbuehler, P. (2022). How agricultural economists are using big data: a review. *China Agricultural Economic Review*.

Lu, L. and Winfree, J. (2021). Demand shocks and supply chain flexibility. In *NBER Conference âRisks in Agricultural Supply Chainsâ*. Springer Berlin/Heidelberg.

Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp. com. *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper*, (12-016).

Lucotte, Y. (2016). Co-movements between crude oil and food prices: A post-commodity boom perspective. *Economics Letters*, 147:142–147.

Luo, Y. and Xu, X. (2019). Predicting the helpfulness of online restaurant reviews using different machine learning algorithms: A case study of yelp. *Sustainability*, 11(19):5254.

Meyer, D. F., Sanusi, K. A., and Hassan, A. (2018). Analysis of the asymmetric impacts of oil prices on food prices in oil-exporting, developing countries. *Journal of International Studies*, 11(3).

Miller, R. K. (2021). The 2013 restaurant, food beverage market research handbook.

Netzer, O., Lemaire, A., and Herzenstein, M. (2019). When words sweat: Identifying signals for loan default in the text of loan applications. *Journal of Marketing Research*, 56(6):960–980.

Neubig, G. (2017). Neural machine translation and sequence-to-sequence models: A tutorial. *arXiv preprint arXiv:1703.01619*.

Newman, D., Lau, J. H., Grieser, K., and Baldwin, T. (2010a). Automatic evaluation of topic coherence. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pages 100–108.

Newman, D., Noh, Y., Talley, E., Karimi, S., and Baldwin, T. (2010b). Evaluating topic models for digital libraries. In *Proceedings of the 10th annual joint conference on Digital libraries*, pages 215–224.

Nosko, C. and Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment. Technical report, National Bureau of Economic Research.

Nwoko, I. C., Aye, G. C., and Asogwa, B. C. (2016). Oil price and food price volatility dynamics: The case of nigeria. *Cogent Food & Agriculture*, 2(1).

Nyga-Łukaszewska, H. and Aruga, K. (2020). Energy prices and covid-immunity: The case of crude oil and natural gas prices in the us and japan. *Energies*, 13(23):6300.

Oliveira, B. and Casais, B. (2018). The importance of user-generated photos in restaurant selection. *Journal of Hospitality and Tourism Technology*.

Ostrom, E. (2010). Analyzing collective action. *Agricultural economics*, 41:155–166.

Peng, C.-H., Yin, D., Wei, C.-P., and Zhang, H. (2014). How and when review length and emotional intensity influence review helpfulness: Empirical evidence from epinions. com.

Phua, J. and Ahn, S. J. (2016). Explicating the âlikeâon facebook brand pages: The effect of intensity of facebook use, number of overall âlikesâ, and number of friends'âlikesâ on consumers' brand outcomes. *Journal of Marketing Communications*, 22(5):544–559.

Pinto Gurdiel, L., Morales Mediano, J., and Cifuentes Quintero, J. A. (2021). A comparison study between coherence and perplexity for determining the number of topics in practitioners interviews analysis.

Piqueras-Fiszman, B. and Spence, C. (2014). *The perfect meal: the multisensory science of food and dining*. John Wiley & Sons.

Pizam, A. and Milman, A. (1993). Predicting satisfaction among first time visitors to a destination by using the expectancy disconfirmation theory. *International Journal of Hospitality Management*, 12(2):197–209.

Poushneh, A. and Rajabi, R. (2022). Can reviews predict reviewersâ numerical ratings? the underlying mechanisms of customersâ decisions to rate products using latent dirichlet allocation (lda). *Journal of Consumer Marketing*.

Radmehr, R. and Rastegari Henneberry, S. (2020). Energy price policies and food prices: Empirical evidence from iran. *Energies*, 13(15):4031.

Ramage, D., Dumais, S., and Liebling, D. (2010). Characterizing microblogs with topic models. In *Fourth international AAAI conference on weblogs and social media*.

Reardon, T., Belton, B., Liverpool-Tasie, L. S. O., Lu, L., Nuthalapati, C. S., Tasie, O., and Zilberman, D. (2021). E-commerce's fast-tracking diffusion and adaptation in developing countries. *Applied Economic Perspectives and Policy*, 43(4):1243–1259.

Ren, M., Vu, H. Q., Li, G., and Law, R. (2021). Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: implications for hospitality marketers. *Journal of Hospitality Marketing & Management*, 30(1):96–119.

Resano, H., Perez-Cueto, F. J., de Barcellos, M. D., Veflen-Olsen, N., Grunert, K. G., and Verbeke, W. (2011). Consumer satisfaction with pork meat and derived products in five european countries. *Appetite*, 56(1):167–170.

Resano, H., Sanjuán, A. I., and Albisu, L. M. (2007). Consumersâ acceptability of cured ham in spain and the influence of information. *Food Quality and Preference*, 18(8):1064–1076.

Rezitis, A. N. (2015). The relationship between agricultural commodity prices, crude oil prices and us dollar exchange rates: A panel var approach and causality analysis. *International Review of Applied Economics*, 29(3):403–434.

Roman, M., Górecka, A., and Domagała, J. (2020). The linkages between crude oil and food prices. *Energies*, 13(24):6545.

Saghaian, S., Nemati, M., Walters, C., and Chen, B. (2018). Asymmetric price volatility transmission between us biofuel, corn, and oil markets. *Journal of Agricultural and Resource Economics*, 43(1):46.

Sanyal, P., Malczynski, L. A., and Kaplan, P. (2015). Impact of energy price variability on global fertilizer price: application of alternative volatility models. *Sustainable Agriculture Research*, 4(526-2016-37964).

Seo, B. (2007). Asymptotic distribution of the cointegrating vector estimator in error correction models with conditional heteroskedasticity. *Journal of Econometrics*, 137(1):68–111.

Seo, Y., Kim, J., Choi, Y. K., and Li, X. (2019). In âlikesâ we trust: likes, disclosures and firm-serving motives on social media. *European Journal of Marketing*.

Serra, T., Zilberman, D., and Gil, J. (2011). Price volatility in ethanol markets. *European review of agricultural economics*, 38(2):259–280.

Shahnoushi, N., Sayed, S., Hezareh, R., and Tirgari Seraji, M. (2017). Investigation of relationship between world food prices and energy price: A panel sur approach. Technical report.

Siami-Namini, S. et al. (2019). Volatility transmission among oil price, exchange rate and agricultural commodities prices. *Applied Economics and Finance*, 6(4):41–61.

Su, C. W., Wang, X.-Q., Tao, R., and Oana-Ramona, L. (2019). Do oil prices drive agricultural commodity prices? further evidence in a global bio-energy context. *Energy*, 172:691–701.

Taghizadeh-Hesary, F., Rasoulinezhad, E., and Yoshino, N. (2018). Volatility linkages between energy and food prices: Case of selected asian countries.

Taghizadeh-Hesary, F., Rasoulinezhad, E., and Yoshino, N. (2019). Energy and food security: Linkages through price volatility. *Energy Policy*, 128:796–806.

Taobao (2013). Tmall seller's naming rules. `https://rule.tmall.com/tdetail-37.htm?spm=a2177.7731969.0.0.1168c32fOVuWXN&tag=self`.

Taobao (2019). Tmall 2020 annual software service fee, discounts and settlement standards. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Taobao (2020a). 2020 tmall sellers evaluation standards. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Taobao (2020b). Tmall seller's deposit rules. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Tian, G., Lu, L., and McIntosh, C. (2021). What factors affect consumersâ dining sentiments and their ratings: Evidence from restaurant online review data. *Food Quality and Preference*, 88:104060.

Tirole, J. (1993). A theory of collective reputations. *Research Papers in Economics University of Stockholm*, 9.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies*, 63(1):1–22.

Tirunillai, S. and Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of marketing research*, 51(4):463–479.

Tzimiropoulos, G. and Pantic, M. (2017). Fast algorithms for fitting active appearance models to unconstrained images. *International journal of computer vision*, 122(1):17–33.

Verbeke, W. (2005). Agriculture and the food industry in the information age. *European review of agricultural economics*, 32(3):347–368.

Wang, F., Chen, L., and Pan, W. (2016). Where to place your next restaurant? optimal restaurant placement via leveraging user-generated reviews. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 2371–2376.

Winfree, J. A. and McCluskey, J. J. (2005). Collective reputation and quality. *american Journal of agricultural Economics*, 87(1):206–213.

Wongkitrungrueng, A. and Assarut, N. (2018). The role of live streaming in building consumer trust and engagement with social commerce sellers. *Journal of Business Research*.

Wongkitrungrueng, A., Dehouche, N., and Assarut, N. (2020). Live streaming commerce from the sellersâ perspective: implications for online relationship marketing. *Journal of Marketing Management*, 36(5-6):488–518.

Wongpiyabovorn, O. (2021). Changes in the relationship between nitrogen fertilizer and natural gas prices in the us.

WorldBank (2020). Fertilizer prices to rise moderately in 2021. `https://blogs.worldbank.org/opendata/fertilizer-prices-rise-moderately-2021`.

Xiao, L., Fu, B., and Liu, W. (2018). Understanding consumer repurchase intention on o2o platforms: an integrated model of network externalities and trust transfer theory. *Service Business*, 12(4):731–756.

Xu, X., Wu, J.-H., and Li, Q. (2020). What drives consumer shopping behavior in live streaming commerce? *Journal of Electronic Commerce Research*, 21(3):144–167.

Yang, S.-B., Hlee, S., Lee, J., and Koo, C. (2017a). An empirical examination of online restaurant reviews on yelp. com: A dual coding theory perspective. *International Journal of Contemporary Hospitality Management*.

Yang, S.-B., Shin, S.-H., Joun, Y., and Koo, C. (2017b). Exploring the comparative importance of online hotel reviewsâ heuristic attributes in review helpfulness: a conjoint analysis approach. *Journal of Travel & Tourism Marketing*, 34(7):963–985.

Zhang, M. and Luo, L. (2022). Can consumer-posted photos serve as a leading indicator of restaurant survival? evidence from yelp. *Management Science*.

Zhou, S., Kan, P., Huang, Q., and Silbernagel, J. (2021). A guided latent dirichlet allocation approach to investigate real-time latent topics of twitter data during hurricane laura. *Journal of Information Science*, page 01655515211007724.

Chapter 2

Collective Reputation of Online Platforms in Agri-food Live Streaming E-commerce: Evidence from Taobao Live

2.1 Introduction

Live streaming e-commerce has emerged as a viable and rapidly growing marketing channel for agri-food vendors in recent years. By demonstrating quality (e.g. tasting apples and showing its pulp) and answering audience questions dynamically in real-time video, live streaming e-commerce helps vendors sell perishable food products and solve the quality concern caused by uncertainty in online shopping. The global live streaming e-commerce market keeps thriving and is projected to reach $247 billion by 2027[1]; it is expected to continue to grow steadily, accounting for $35 billion in sales and 3.3% of all U.S. e-commerce by 2024[2]. This promising market attracts numerous sellers and online platforms around the world jumping on the bandwagon. We are witnessing American entrepreneurs selling yogurt on Amazon Live, Singaporeans offering fresh fish through Facebook Live, and Chinese farmers selling fresh fruits and vegetables on Taobao Live[3].

With so many online platforms and products flooding into live streaming e-commerce, one determinant for consumers to make the decision is the collective reputation. It refers to the average quality of the goods produced by a group of firms to which an individual firm belongs(Tirole, 1993; Frick and Simmons, 2013). Except for product collective reputation, it includes online platforms collective reputation in the e-commerce context. For example, products from the same region are often viewed as

---

[1] $https://www.prnewswire.com/news-releases/global-live-streaming-market-is-projected-to-reach-247-billion-by-2027-301290473.html$

[2] $https://www.retailtouchpoints.com/topics/digital-commerce/livestreaming-update-coresight-debunks-5-common-myths$

[3] a live streaming platform owned by Alibaba

a group: some food products, like Washington apples or Idaho potatoes, that come from specific areas shares geographic labels to signal quality(Winfree and McCluskey, 2005); online platforms such as Whole Foods Market, Walmart, and Alibaba's Tmall, characterized as selling organic food, groceries, and brand-name goods, respectively. Both types of collective reputation are important because they serve as a tool for consumers to expect quality, select sellers, and decide the frequency of consumption for a online platform in live streaming for experience goods like agri-food products, for which consumers may not know quality prior to purchasing.

Despite the important role of two types of collective reputation, it is unknown how collective reputation of online platforms influences agri-food live streaming revenue, which is the question that this paper seeks to answer. On the one hand, live streaming assists consumers in perceiving the quality of agri-food products, allowing them to select vendor from another online platform, thus substituting the role of the collective reputation of online platform. For example, consumers may choose to buy bread from a relatively unknown seller on Facebook Live rather than spending time and effort searching on Amazon because Facebook Live allows them to build the quality expectation that "both breads have similar quality" by showing the bread in detail and answering any questions that consumer have. Live streaming, on the other hand, may act as a complement to the collective reputation of online platform. While live streaming video shows product detail and provides information that consumers want through interactions, consumers still perceive risk (e.g. concerns about shipping time) so they adopt the collective reputation of the online platform as one of the additional methods of selecting sellers or products.

In this paper, we explore how the collective reputation of online platform affect agri-food products live streaming revenue. We start by identifying the collective reputation of online platforms with platform rules differences between the Tmall and Taobao Marketplace sellers on Taobao Live. The collective reputation of online platforms comes from some platform rules and administration processes are mandatory only for Tmall sellers while none of these are compulsory for Taobao Marketplace sellers, which provides an ideal place to make identification. First, Tmall is a group of quality and brand-name product coalition that shares uniform quality standards: the prerequisite for entering Tmall is to sell brand-name or officially authorized brand products. Second, Tmall sellers must provides quality and money-back guarantee for products. Third, Tmall will operate as a third-party regulator of product quality, reviewing quality and screening out sellers that do not meet the quality standard to ensure the overall reputation. Besides, different entering barriers further filter out companies that do not meet the qualification: Tmall sellers must pay at least $10,000 of the security deposit plus annual technology and service fee[4]. As a result, consumers generally believe products from Tmall sellers have better quality and less risk than Marketplace sellers.

Then we collect data and construct a comprehensive dataset, which includes the daily real-time transaction records for 10 various types of fresh fruits (product price, sales volume, the number of Thumbs up, live streaming duration, etc). We also manually search each product's corresponding web page and document the detailed information of sellers and products (e.g., which online platform the seller belongs to, products origin, etc.) to avoid potential omitted variable biased estimates. After that, we use the fixed-effect model to control product type and origins, which can alleviate the time-invariant effect from geographic or product type to confirm the effect of the collective reputation of the online platform on live streaming revenue.

The result shows that collective reputation of online platform has significantly positive impact on live streaming revenue: Tmall sellers will earn an average more than 5,000 yuan (almost $900) per product per day than the Taobao Marketplace sellers in live streaming, which is nearly 4 times the average revenue per product per

[4] https://about.tmall.com/tmall/fee_schedule?spm=3.6635917.0.0.1dbb8f18XqZhw1#place

day of Taobao Marketplace sellers. Intuitively, it may because collective reputation of online platform, combined with live streaming, enables consumers to know more and perceive less risk about product: the collective reputation of online platform further guarantees the quality of the product and lower the risk perceived by consumers. This implies that the collective reputation of online platform plays an important role in live streaming, which can not be substituted by the live streaming.

Some unobservable variables, however, may lead to endogeneity problems. For example, low product reputation may cause consumers to lower their overall evaluation of the platform and discourage them from placing orders, which negatively impacts live streaming revenue. Therefore, we use the security deposit (one-time fee, which is used as collateral in the case of any damages incurred by Tmall.com or any customer) as an instrumental variable to address the endogeneity issues[5]. Our instrumental variable is valid because (i) security deposit is mandatory and pre-requirement for sellers who want to enter the Tmall; (ii) the amount of security deposit is associated only with a brand censoring status conducted by the Bureau of Industry and Commerce, which means it wonât be affected by any other variable in our model.

The results from IV approaches show the significant greater impact of the collective reputation of online platforms on live streaming revenue, compared to the results from the fixed-effect model: Tmall sellers will earn an average of 7,000 yuan (roughly \$ 1,100) per product per day more than sellers from Taobao Marketplace, which is nearly 6 times the average revenue per product per day from Taobao Marketplace. The practical implication of our research for agri-food retailers is that they should choose a platform with the higher collective reputation to lunch the live streaming by just simply switching to a higher collective reputation online platform(Tmall).

Our contribution to the research of live streaming and collective reputation is in the following three areas. First, we fill the research gap of existing live streaming

[5] $https://about.tmall.com/tmall/fee_schedule?spm=3.6635913.0.0.5bd4206ccn31vr\#place$

e-commerce studies that lacking agri-food products. Second, we extend our understanding of the collective reputation of online platforms to a new area: live streaming e-commerce, and show that it has a direct positive impact on live streaming revenue. Third, we expand the indirect measurement of online platform collective reputation by differentiating the platform's rules.

2.2 Literature Review

2.2.1 Live Streaming e-commerce: Current Research

Previous research on live streaming e-commerce can be classified into two categories: (i) research from the consumer side focus on factors influencing consumer engagement (e.g., sent text comments) and purchasing behaviors in live streaming. They include a social bond (such as interpersonal interactions, friendships) or structural bonds[6](Hu and Chaudhry, 2020), interactivity(Kang et al., 2021), and perceived value by consumers in live streaming(Wongkitrungrueng and Assarut, 2018). Some research also shows that the unique feature of live streaming, real-time interactions between consumers and retailers, can affect consumer's purchase intentions so that affect consumers' decision-making (Hu and Chaudhry, 2020; Xu et al., 2020; Clement Addo et al., 2021) (ii) literature from seller's perspective discuss strategies to increase live streaming sales(see Chen et al. (2020); Wongkitrungrueng et al. (2020)). For example, Wongkitrungrueng et al. (2020) proved that the combination of different live streaming strategies and approaches, such as offering limited quantity or price, may boost sales in Facebook Live.

The literature on the customer or seller side lacks real-time sales information and agri-food products so we don't know how exactly these strategies affect live

[6]defined as the value-adding services that are not available everywhere and are expensive or difficult for consumers to obtain if they terminate the relationship

streaming sales. Furthermore, although consumers can learn detailed information about the product in live streaming, they may not know some features of agri-food products like an apple's weight or odor through the video and interaction. As a result, the collective reputation of the online platform may be one of the factors that help consumers make purchase decisions because it makes consumers feel less risky. But we do not know how it impacts consumers' behaviors thus revenue in live streaming.

Our work is based on the perspective of sellers but distinguished from previous works by focusing on live streaming revenue, especially for agri-food products. We show that the collective reputation of online platforms is another factor that influences consumers' choice behaviors thus revenue in live streaming e-commerce. Moreover, our dataset includes complete and detailed transaction price and sales volume and live streaming variables (e.g., live streaming duration, times of thumbs up), which is the more accurate and direct measurement of sales.

2.2.2 Online platforms reputation in e-commerce

With the advent of e-commerce platforms such as Amazon or eBay, many online shopping platforms introduce seller or product score ratings (ranges from 0 to 5) to represent their reputation to reduce asymmetric information between sellers and consumers. There are many studies discussing the important role of reputation in e-commerce, which may also be divided into two parts: (i) research about the reputation mechanism within the online platform to explore the correlation between product or seller's reputation and online sales. The evidence from Lin et al. (2006) indicates that sellers with high reputation scores have a higher sales growth rate than sellers with lower reputation scores. Supported Lin's idea, Amblee and Bui (2008) conclude that brand reputation can attract more users to view products and increase sales and profits. (ii) studies regarding the effect of the reputation of online platforms on consumer purchase behaviors. For instance, the result from Xiao et al. (2018)

shows trust in online-to-offline platforms has positive impacts on consumer repurchase intentions. Since the reputation of the online platform is difficult to measure, some of the earlier literature used indirect measurement based on consumers' behaviors in the online platform. Nosko and Tadelis (2015) uses "whether a buyer returns to eBay to purchase again" to measure the quality of eBay as a whole.

The research of online platform collective reputation is still in the early stage compared to the reputation system within online platforms and the role of collective reputation of the online platform in the context of live streaming e-commerce remains unknown. Our study differs from the literature on live streaming and collective reputation in e-commerce context we use indirect measurement to measure the collective reputation of the online platform, but we utilize rules discrepancy for two similar platforms to identify the collective reputation of online platforms and we extend the role of collective reputation of online platforms to the context of live streaming e-commerce.

2.3 Background

2.3.1 Taobao Live and its two types of sellers

Founded by Alibaba in 2016, Taobao Live has the lionâs share of live-streaming e-commerce at roughly 80% in China[7]. Transaction values on Taobao Live reached more than 100 billion yuan ($15.3 billion) in 2018, with a year-on-year growth rate of nearly 400%[8]. In 2019, agri-food products sales on Taobao Live reached 6 billion yuan ($1 billion)[9]. From 2019 to 2020, over 1.4 million agri-food products live streaming sales conducted by over 60,000 farmers/sellers on the Taobao Live platform (AliResearch,

[7] https://www.forbes.com/sites/michellegreenwald/2020/12/10/live-streaming-e-commerce-is-the-rage-in-china-is-the-us-next/?sh=69d8e06f6535

[8] http://www.aliresearch.com/ch/information/informationdetails?articleCode=132760647053217792&type

[9] http://www.aliresearch.com/ch/information/informationdetails?articleCode=132760647053217792&type

2020). Taobao Live has two primary platforms(sellers): Taobao Marketplace and Tmall, which occupy 75% of the online agri-food products market share across China[10] and have many commonalities including product type, price, etc.

Two features are worth highlighting to differentiate Tmall from Taobao Marketplace. The first difference is in the entrance criteria, which are reflected in brand requirements, security deposit, and platform operation fee. (1) Brand requirement. Tmall requires that all Tmall sellers either own a brand or obtain brand authorization while Marketplace sellers are not required to do so. As a result, Tmall sellers need to develop their brands or become authorized sellers for other brands. Establishing a brand is an insufficient condition for selling on Tmall, however, since Tmall sellers must pass a rigorous brand evaluation implemented by Tmall. The brand evaluation, which is a completely independent process[11], involves methods for brand authorization (e.g., verifying whether brand authorization contracts are lawful) and evaluation of brand influence[12]. Even if some companies have good business performance (e.g., have a high-profit margin), they may not be eligible to enter Tmall due to insufficient evidence of brand influence. Tmall will also take the initiative to invite some companies with brand influence to enter.

(2) Security deposit and operation fee. Security deposit is the one-time fee, which is used as collateral in the case of any damages incurred by Tmall or any customer, unfrozen upon the termination of the Tmall Service Agreement[13]. Tmall and Taobao Marketplace require sellers to pay different amounts of security deposits: sellers on Tmall are mandatory to pay a security deposit based on their brand registration status before entering while sellers on Taobao Marketplace are not mandatory but

[10] http://industry.people.com.cn/n1/2019/0423/c413883-31045244.html

[11] Tmallâs entry requirements include company business profiles such as year established and management structure. The independent review process here means that it is not affected by parties or interests outside of Tmall.

[12] The definition and evaluation process of brand influence is confidential, but we have learned from public channels that one of the evaluation indicators is number of times of brand in search engines(like Google or Baidu) or the number of related news reports.

[13] $https://about.tmall.com/tmall/fee_schedule?spm = 3.7128306.0.0.529342f9dEOJFJ\#place$

encouraged to pay a minimum security deposit of 1,000 yuan ($154). For Tmall sellers, if their brand is under censoring by the Chinese Bureau of Industry and Commerce, then they need to pay a security deposit of at least 100,000 yuan ($15,384), while the security deposit for sellers holding registered brand is at least 50,000 yuan ($7,692)(Taobao, 2020b). For the platform operation fee, any individuals or companies can register as Taobao Marketplace sellers without charging but Tmall sellers must pay substantial software service fees to launch the Tmall platform (Taobao, 2019).

The second difference relates to seller operating rules within the platform. Tmall conducts an annual business performance evaluation once sellers have successfully entered Tmall, which includes an assessment of sales (e.g., monthly sales must match Tmall's sales objectives) and customer service ratings, etc. A Tmall seller cannot continue to sell on the Tmall platform if it fails an evaluation (Taobao, 2020a). However, Taobao Marketplace sellers are not subject to such evaluations. In addition, Tmall sellers must sign a quality and money-back guarantee agreement for which sellers will receive a fine of four times the transaction price if they sell the fake products, and are also subject to delivery time agreements (Tmall sellers must ship goods within 48 hours of purchase). These agreements reduce the risk of consumers buying low-quality products and shipping delays, respectively. Taobao Marketplace sellers do not have such compulsory agreements. The other operation rule difference between Tmall sellers and Marketplace sellers is the store naming rule and the specified number of products(for agri-food products, Tmall sellers need to sell at least 10 types of products), while marketplace sellers do not have such a requirement[14].

The distinct participation and operating rules between Taobao Marketplace and Tmall sellers provide an ideal setting to explore the role of the collective reputation of the online platform in agri-food live streaming. Consumers will typically believe

[14] https://service.tmall.com/support/tmall/knowledge-4781237.htm?spm=a223k.8006297.

that Tmall products are of higher quality and better after-sale service than those sold on Taobao Marketplace. Despite the differences in entry barriers and operations rules between Taobao Marketplace and Tmall, users can access any live streaming room for free. Additionally, consumers do not know whether it is a Tmall or Taobao Marketplace seller when selecting a live streaming room/portal until clicking on the âproduct detailsâ icon in live streaming video.

2.3.2 Agri-food Live Streaming: An Example From Taobao Live

Consumers follow four steps to purchase agri-food goods in live streaming: (1) search for products, (2) select and enter the live streaming room, (3) watch the live streaming videos and interact with sellers in real-time, and (4) place orders. To illustrate the agri-food live streaming shopping experience, we provide an example of purchasing apples in Taobao Live. Figure 1 describes the general process that leads from product searching through to purchasing within Taobao Live. Figures 2 to Figures 5 provide a graphic illustration of each step.

[Figure 1: Customer Decision Making Process in Live Streaming]

Step 2: Selecting and enter a live streaming room/portal. Customers will see a dashboard that offers a variety of fresh fruit seller live streaming room "business cards" after launching the app on their phones and searching for a product or category such as "fresh fruit" as shown in Figure 2[15]. A live streaming room âbusiness cardâ includes three types of basic information: (1) the number of customers who are currently watching the live streaming, (2) room/portal name, and (3) product pictures with a brief description (e.g., delicious, imported Vietnamese pineapple). Customers

[15] Customers can watch the live streaming on the computer as well, but the mobile app is officially recommended by Taobao as the viewing experience is better

can either choose one of the live streaming rooms or pull down to refresh the current page to get a new set of live streaming room recommendations, which they can do until they find a satisfactory room/portal. While watching the live streaming, if the product in the current live streaming room does not meet their demands, they can exit the room at any moment and return to the search dashboard to choose another room.

[Figure 2: Relevant Agri-food Products Live Streaming Rooms]

Step 3: Interacting with the streamer while watching live streaming. Figure 3 illustrates the experience of watching the live streaming after entering the seller room. Consumers will find several types of information on the live streaming page, including the product price, discounts and comments proposed by other consumers who are now watching the live streaming. Consumers can communicate with the streamer by texting in the âchat boxâ, and the seller can provide general or personalized answers in response. For example, some consumers may want to know the delivery time, while others are concerned with product taste. To express their concerns, they can type the text of their thoughts, which appear in the chatbox, and all other customers in the live streaming room will see. Meanwhile, the streamer answers the customer's request or queries orally in the video or by text. Figure 3 shows an example in which one consumer inquired about the apple's flavor (consumer A) while another inquired about the number of apples supplied per order (consumer B). The streamer responded to consumer A's question by taking a bite of an apple and describing its taste, and to consumer B's question by showing the packaging case with the contents displayed on it live on the video.

[Figure 3: Watching the live streaming videos and interacting with the streamers]

Step 4: Placing the order. Once consumers have learned enough information about apple and decide to place the order directly then they can click "buy now" on the live streaming room page as Figure 4 shown. On the other hand, consumers who want to learn more about apple than what is shown in the live streaming room can click the "apple picture" (Figure 3), which takes the consumer to a new transaction page with information including the product's past customer reviews, the seller's information (such as service rating), and other information (shown in Figure 5). When viewing the "product details" web page, the live streaming video is available through the mini window. Finally, like shopping on Amazon or eBay at the payment page, the consumer only needs to fill in the zip code, address, and contact information, and then pay with Alipay[16] to complete the transaction(Not shown here).

[Figure 4: Transaction page in the live streaming room]

[Figure 5: Transaction page on "product details" web page]

[16] A third-party mobile and online payment app developed by Alibaba

2.4 Data

We collected daily live streaming sales data from a professional third-party database (which contains millions of live streaming transaction records) and sellers and products information from product web pages (including Taobao Marketplace and Tmall sellers) to conduct the empirical analysis that determines the impact of the collective reputation of the online platform on revenue. We then merged the two datasets by matching the seller's name (or store name) to obtain the dataset used in the regression analysis. The first dataset contains live streaming information of 10 different kinds of fresh fruits with the highest total live streaming sales from June 1, 2020, to February 28, 2021. The second dataset consists of the live streaming seller and product information within the same time period.

We focus on fresh fruits live streaming for three reasons. First, fresh fruit sales may benefit more from live streaming to reduce costs than processed product sales. Fresh fruits, such as pears, have a significantly shorter shelf life than dry, processed agri-foods, such as canned pears, and hence the costs of inventory management, including the risk of quality degradation over time, are higher for fresh than for processed products. As a result, fresh fruit sellers are more likely to adopt live streaming to sell fruits quickly and save costs. Second, fresh fruit producers (farmers) participate in the sales without processing so they can directly communicate with consumers and reduce their uncertainty. Compared to relying on live streaming agents(or intermediary sellers) to sell products, farmers can directly participate in the sales process and answer consumer questions in more detail. Third, we found missing values in other agri-food product categories (e.g., fresh vegetables, meat) across the platforms. For example, we found pork or aquatic products records from Tmall sellers but no such records from Taobao Marketplace sellers. The fresh fruits data is relatively complete, hence we focus on fresh fruits.

2.4.1 Live Streaming Data

Live streaming revenue in our dataset refers to the sum of the volume of all products sold in live streaming multiplied by their transaction price (yuan) from the beginning to the end of the observation period. We focus on the live streaming revenue, though sellers make sales via multiple channels (e.g., via the website outside of when the live streaming occurs) within the observation period. The database shows the live streaming sales volume directly, so we sorted the fresh fruit live streaming sales volume from June 1, 2020, to February 28, 2021, and selected the top 10 different kinds of fresh fruits to obtain the dataset. Sales volumes that were outliers were dropped from the sample[17].

The first independent variable that was determined to likely influence live streaming revenue was "live streaming duration", a measure of the time length (in hours) of live streaming on a given day. This is because the longer the live streaming lasts, the more consumers are likely to come and watch the streaming and shop thus increasing revenue. Live streaming duration is the sum of the time length of the product in live streaming. For example, if the same apple product is offered in three different live streaming rooms, the product's live streaming duration is the total of the time the product was actively sold in those three rooms. Records with 0 live streaming hours throughout the observation period were removed since the seller was not selling via live streaming on those days.

The second independent variable was the number of âThumbs upâ. It refers to the number of people that hit the "Thumbs up" button from the beginning to the end of the live streaming session (notice that repeated "like" button clicks for the same account are not permitted). This is comparable to the number of "likes" a Twitter post receives. During live streaming, we employ "thumbs up" to assess the engagement between the audience and the streamer(Seo et al., 2019). It's also the

[17]Note that such exclusions did not impact our empirical results

evaluation of how the audience feels about the product, the seller, or the live streaming content(Phua and Ahn, 2016).

2.4.2 Sellers and products information

Although the first dataset provides us with comprehensive live streaming sales information, it lacks seller and product information, which are also important for explaining revenue. For example, agri-food productsâ origins (like the Florida oranges) are essential information regarding product collective reputation, which can affect consumer choices (Tirole, 1996; Winfree and McCluskey, 2005; Ostrom, 2010). Therefore, we manually collected seller and products information, including security deposits, sellersâ service ratings, and product origin, to construct the second dataset.

We first manually collected the origin of each product on the product detail page to control for the effect from the product collective reputation. The origin information includes province and county for products produced within China. Imported fruits (such as mangoes imported from Vietnam) are designated by country.

The second independent variable was the dynamic service rating (DSR) score, which is the service rating on the product detail page. The DSR is a dynamic rating that measures how satisfied customers are with a seller's customer service, incorporating factors like response time to customer requests and professionalism in addressing concerns. The DSR is calculated by both Tmall and Taobao Marketplace platform, which has the same calculation methods. We chose the rating on February 28, 2021, as the typical DSR for each seller because we were unable to get historical data on the DSR throughout time.

Thirdly, we add a variable to differentiate the Tmall and Taobao Marketplace platforms. To find the seller's platform attributions, we studied the naming rule of sellers across the platform. Tmall has unique and compulsory naming rules for its sellers, whose names must be suffixed with âFlagship Store[18]â, âSpecialty Store[19]â or "Authorized Store[20]" while the Taobao Marketplace sellers are not permitted to use such keywords as part of their store name(Taobao, 2013). By filtering the store name suffix, we can identify the seller's type. There is no naming rule requirement for live streaming room names for both Tmall sellers and Taobao Marketplace sellers, but Taobao Marketplace sellers are still not permitted to use suffixed key words that are used by Tmall sellers as their live streaming room name.

Finally, we complete the second dataset by collecting security deposit information. Since the amount of the security deposit is based on the brand(trademark) registration status (e.g., whether it is a registered trademark or under censoring) and platform rule, we obtain brand registration status for Tmall sellers via searching through an authoritative third-party website[21] to determine the amount of the security deposit. For Taobao Marketplace sellers, we check the seller information page displayed on the corresponding web page to find their security deposit amount.

2.4.3 Summary Statistics

Table 1 presents summary statistics after merging the two datasets by platform. The full sample contains 9,505 live streaming records (1,538 from Taobao Marketplace sellers and 7,967 from Tmall sellers, respectively). We can observe the following facts.

First, Tmall sellers' the average live streaming revenue is much higher than Taobao Marketplace sellers. This indicates that when consumers do not know whether they are Tmall or Taobao Marketplace sellers in live streaming, consumers are more in-

[18]Merchant is the brand (trademark) owner, or possesses exclusive authorization from the brand owner when opening their storefront on Tmall.com

[19]Merchant sells products of two or more brands that belong to the same product category

[20]Merchant holds authorization documents from the brand owner to open a store on Tmall.com

[21]https://www.tianyancha.com/. Search on this website is similar to Google, where the company or brand name is submitted, and the website provides the registration status of trademark: canceled or under censoring. From this, we can determine the amount of deposit required.

clined to place orders for products sold by Tmall sellers. This implies that consumers appear to make their consumption choice on quality considerations associated with different platforms. We can also observe a substantial gap between Tmall and Taobao Marketplace sellers regarding average revenue (9,188 versus 1,279, respectively). This disparity illustrates the important impact of reputation-related factors on live streaming revenue.

Second, live streaming measurements between two platforms are similar to each other. This is consistent with expectations because consumers are unaware of the platform type of the seller when selecting a live streaming room. Therefore, customers tend to focus on live streaming content before looking for information about the seller's type on the "product details" web page. The Taobao Marketplace sellers and Tmall sellers had almost the same average number of audience members (6.7 versus 6.5) and live streaming duration (11 versus 15). These two results show that the consumers in our sample may not have selection bias. That is, the relatively greater revenue for Tmall sellers may not be due to consumers intentionally choosing to watch the live streaming from Tmall sellers. Furthermore, Taobao Marketplace sellers have twice the number of âthumbs upâ of Tmall sellers (13.1 versus 6.5), which shows that Taobao Marketplace sellers may have better interactions than Tmall sellers. Despite this, it appears that more interactions may not narrow the revenue gap between two online platforms.

Third, the average security deposit of Tmall sellers is nearly 10 times that of Taobao sellers (83,507 versus 8,388). This is expected due to Tmallâs higher and mandatory deposit requirements.

Fourth, the most popular products sold by Tmall and Taobao Marketplace sellers were apples (23.7%), pears (13.8%), and dragon fruit(12.3%) regarding fruit type. This shows that consumers are concentrated in the live streaming rooms of specific types of fruit.

Lastly, the average service rating of Taobao Marketplace sellers was 0.1 points higher than that of Tmall sellers (4.69 versus 4.59), which shows that there is not a substantial difference between the platforms regarding average consumer satisfaction. These live streaming and seller's information variables between Tmall and Taobao Marketplace are similar to each other, but we can see that revenue in different online platform has a significant gap, which may be explained by the different online platform.

2.5 Empirical strategy and results

2.5.1 OLS fixed-effects model

The first model used to explain variation in live streaming revenue is OLS fixed-effects model that controls for fruit type and origin fixed effects. The fruit type and origin fixed effects enable us to alleviate bias and endogeneity problems caused by unobservable time-invariant fruit and origin characteristics that are difficult to account for given available data. For example, the taste of apples will not change in a short period and even if the brands are different there will be no obvious differences in the taste of the same type of apples. Furthermore, agri-food products that have the same origin may share benefits from collective reputation (Winfree and McCluskey, 2005) while it will not change in a short time. By using a fixed-effects model, we can control the effect from such factors even though these variables are not included in the dataset. The following equation (1) illustrates the baseline model:

$$y_{ijt} = \beta_0 + \beta_1 * Tmall_i + \beta * C'_{ijt} + \mu_j + \epsilon_{ijt} \tag{2.1}$$

where y_{ijt} is the live streaming revenue of the seller i's fruit type j at time t; $Tmall_i$ is the platform indicator, which equals to 1 if it is the seller from Tmall, and 0 for sellers from Taobao Marketplace .

C'_{ijt} denote the live streaming control variable vector, which includes $Price_{ijt}$, $ThumbsUp_{ijt}$ and $LiveDuration_{ijt}$. These variables are the average transaction prices, recorded "thumbs ups", and live streaming duration and for seller $i's$ live streaming product j at time t. We choose price as a control variable because it is the main product value information. The rest of variables are used to control for differences in live streaming across firms such live streaming time length ($LiveDuration_{ijt}$) and quality of interactions ($ThumbsUp_{ijt}$)(see Li and Peng, 2021). $ServiceRating_i$ is used for controlling time-invariant seller's service effects (e.g., customer service may not change in the selected period but still affects the purchase intention). μ_j is the fruit type and origin fixed effect (here we construct the interactions to capture this effect); β_0 is a constant term, and ϵ_{ijt} is a random error.

Table 2 presents the regression results for the baseline OLS model. We estimated four types of regression equations. All equations contain the fixed effect of fruit type and origin interaction and each of the four regressions had one additional control variable added from the previous one. We showed the regression results with an increasing number of control variables such that the column number (1 to 4) represents the number of control variables in each estimated regression model. Overall, we observed that Tmall is significantly ($p < 0.01$) and positively (+) associated with greater live streaming revenue across all models. The results imply that sellers on Tmall will, on average, earn from 4,783 yuan ($735) to 5,611 yuan ($865) per product per day more than sellers on Taobao Marketplace while holding other variables constant.

We can draw some preliminary conclusions from Table 2. First, the results from the model with all controls (column 4) show that Tmall sellers have much higher mean revenues per product per day than Taobao Marketplace sellers (5,611 yuan versus 1,279 yuan). Although the estimates of the Tmall effect continue to change across models, the changes are within a range of one unit of standard error, which shows that online platform and revenue have a generally stable relationship. Second, the R^2 improves greatly (by over 10%) from the model that only includes price as the control (column 1) to the model that also adds live streaming duration (column 2). This indicates that live streaming related variables are important for explaining liver streaming revenue variation. Third, the R^2 increases from 0.05 (column (1)) to 0.20 (column (4)), which suggests that model explanatory efficiency improves after accounting for more control variables.

2.5.2 IV estimation

Some unobservable variables may correlate with $Tmall_i$ and revenue thus relationship between the collective reputation of platform and revenue may have endogeneity issues. For example, Tmall rules require sellers to sell at least 10 types of products, but no such requirement for Taobao Marketplace sellers. Therefore, the number of products sold correlates with the seller's type, and under the assumption that a greater number of products sold is associated with higher revenue, an endogeneity issue emerges. To solve for endogeneity problem, we adopt the security deposit as an instrumental variable. Here we focus on discussing the validity of the security deposit as an instrumental variable.

A good instrument will be strongly correlated with the endogenous regressor, in our case $Tmall_i$, and uncorrelated with the error term ϵ_{ijt}. First, the security deposit must correlate with the endogenous variable ($Tmall_i$). As the Tmall rule indicates, sellers who wish to enter Tmall must pay the security deposit, which is a mandatory prerequisite imposed by Tmall. So security deposit correlates with Tmall by platform rules definition thus meets the association condition. Second, the security deposit needs to meet the exclusion restriction. According to the official Tmall platform rules, the amount of security deposit is only associated with the brand registration status and does not depend on any other variables such as brand influence or number of required sold products in Tmall. So the security deposit cannot affect the revenue through unobservable variables, and thus satisfies exclusion restriction.

The following two-stage equations were designed to account for the instrumental variable effects in the first stage, and then re-estimate the fixed effects regression with the estimated values for Tmall from the first stage regression included along with the same controls and fixed effects in the second stage. Equation (2) is to illustrate our first-stage estimation, where $Tmall_i$ is instrumented by $Deposit_i$. We then include the estimation results from the first stage $T\hat{mall}_i$ into the second stage with the same live streaming controls as in the OLS model.

$$Tmall_i = \alpha_0 + \alpha_1 * Deposit_i + \alpha * C'_{ijt} + \xi_j + \eta_{ijt} \tag{2.2}$$

$$y_{ijt} = \beta_0 + \beta_1 * T\hat{mall}_i + \beta * C'_{ijt} + \mu_j + \epsilon_{ijt} \tag{2.3}$$

Table 3 shows IV regression results, where panel A and panel B show the second stage and first stage results, respectively. Panel A suggests that $Tmall$ has a statistically significant ($p < 0.01$) and positive (+) relationship with revenue. The estimated effect is much greater than the estimate from the OLS model. This may be because some unonservable variables like the seller reputation that makes up the collective reputation of online platform correlates with security deposit: seller with higher reputation is more likely to enter Tmall and thus need to pay a large amount of security deposit; Taobao Marketplace seller with high reputation also tends to pay the security deposit to ensure their reputation. As a result, the OLS results in Table 2 are downward biased (from 7,059 to 5,611).

The results displayed in column (2) to column (4) show that sellers on Tmall would earn an average of over 5,000 yuan ($769) more than sellers from Taobao Marketplace, which allowed Tmall sellers to grow their revenue to 5 times the average revenue per product per day of Taobao Marketplace sellers.

Panel B indicates the regression results of the first stage. β_1 in the first stage is statistically significant ($p < 0.01$) with a standard error close to 0. The model that accounts for all control variables in column (4) has an R^2 value of 0.8, which is the highest among all models. These results confirm the association between the *Deposit* and *Tmall*. This is strong evidence that deposit can be used as an instrumental variable for *Tmall*.

2.6 Discussion & Limitation

2.6.1 Discussion

We attribute the substantial live streaming revenue gap to the collective reputation differences of online platforms. Live streaming may substitute the role of collective reputation for products, but the collective reputation of online platforms appears to predominate in consumers' choice in agri-food products because it not only offers quality and after-sale warranty that consumers care about but also guaranteed delivery time. In our case, Tmall provides consumers with such services that Taobao Marketplace sellers are not required. So consumers would believe that products from Tmall are more reliable in terms of quality and after-sales support, hence are less risky to purchase. Although customers have access to as much product information as possible in live streaming, such knowledge does not alleviate their concerns about product quality and other aspects. As a result, even Taobao Marketplace sellers have a longer live streaming duration and more interactions than Tmall sellers as shown in Table 1, consumers are still hesitant to purchase in Taobao Marketplace store, which generates a revenue gap. Our results confirm the reputation that enables economic agents to credibly offer warranties are important for explaining consumer purchase behavior(Fafchamps et al., 2008).

Second, positive externality of online platforms further widens the live streaming revenue gap between sellers on two platforms. This may be because the collective reputation of the platform has a positive impact on all products, which makes consumers more willing to buy them. Previous research proves that the collective reputation of the online platform has a positive externality on consumer's purchase decision(Nosko and Tadelis, 2015). In our case, the collective reputation of Tmall about the products quality and after-sales service makes consumers less concerned about all types of fruits in Tmall and reduces their search cost so they are more willing to purchase. Even it is a less well-known brand on the Tmall platform, as long as the product comes from Tmall, like apples from Washington State or red wine from Bordeaux, then consumers may consider it to be of higher quality and better after-sales service. So that consumers are willing to buy all products from Tmall rather than Taobao Marketplace for the same type of fruits then increase the sales of the entire category of fruits for Tmall sellers.

Third, our results mean that consumers value the service provided by the platform based on the fact that consumers are willing to consume agri-food products on a certain platform with a higher reputation. It has practical implications for platform designers and retailers. On the one hand, online platform rule designers can protect consumers' welfare while increasing seller's revenue by developing mandatory quality and after-sales service-related regulations. In our case, although Taobao Marketplace seller has the promise of quality or after-sales service as well, it is not mandatory like Tmall. In other words, it is the mandatory enforcement of the platform rules that protects the consumer's welfare and makes consumers less risky to purchase. Besides, building an entrance barrier not only filter company with bad operation performance but also helps to preserve the platform collective reputation, which may have a positive influence on customer purchasing behavior. On the other hand, for small and medium-sized enterprises who can not afford to be a member of higher reputation platform, signing product quality or after-sales agreements with the platform or marketplace even if they are not mandatory, and displaying them during live streaming may help consumers trust more in their products and increase their revenue.

Lastly, our results help agri-food vendors to measure whether the benefits are worth the costs of entering into higher collective reputation platforms. In our case, sellers on Tmall need to pay a large amount of security deposit and their operation fee is high. Even so, the live streaming revenue from the high collective reputation of the platform may offset this part of the cost. Our research provides decision-making references for sellers who intend to adopt live streaming to sell goods to measure the cost and benefit of entering the higher collective reputation platform.

2.6.2 Limitations

There are several limitations of this study. First, our dataset is limited to live to stream, which means that we cannot compare the effects on sales from before and after the adoption of live streaming. This issue relates to the business strategy among small or medium-sized firms regarding whether to adopt live streaming. Answering this question is left to future research.

Second, we cannot distinguish the effect from brand reputation or streamer's reputation on revenue. Tmall is a gathering platform for various brands(like a department store, but online). Consumers may buy products on Tmall because of their preference for the specific brand. Similarly, consumers may buy some products because they prefer a certain streamer (e.g., they are the fans of a celebrity). Future scholars can explore the ways to separate different sources of reputation-related effects.

Third, our research is limited to e-commerce platforms like Taobao Live. We don't know what role of collective reputation plays in social media live streaming like Facebook Live, and the difference between these platforms. Social media expands the spectrum of reputation risks and boosts risk dynamics(Aula, 2010), while the professional e-commerce live streaming platforms may not have the same effect thus generating different effects on seller's revenue. Future research can compare the role of the collective reputation of online platforms in a different types of live streaming platforms to find out a way to increase sales.

2.7 Conclusion

Live streaming is an emerging sales channel for agri-food products and we still do not know the mechanisms that increase agri-food live streaming revenue. We use a series of the platform rules of two platforms on Taobao Live to identify the collective reputation of online platforms and study its role in explaining the live streaming revenue gap. We use the fixed-effect model that controls for fruit type and origin to estimate the impact of the collective reputation of the online platform on revenue. Specifically, we use the security deposit as an instrumental variable to address the endogeneity problem and the unique feature of the security deposit (which is pre-requirement of entering Tmall and no correlations with other variables) proves its validity. Agri-food vendors on a higher collective reputation had an average of 7,059 yuan ($1,086) more live streaming revenue per product per day than those selling on the lower collective reputation of platform, based on results from the second stage regression.

Our results have valuable and practical implications for agri-food retailers. Particularly, the online platform's collective reputation is critical for increasing revenue in agri-food live streaming. As a result, small and medium-sized agri-food businesses who want to increase revenue by live streaming should consider becoming a seller on a well-known collective reputation platform.

Our contribution to food marketing and consumer choice research is that we provide empirical evidence of the collective reputation of the online platform that can positively influence consumers' purchasing behavior and its important role in live streaming e-commerce. Admittedly, our analysis was limited to 10 types of fresh fruits, and we only control the live streaming duration and thumbs up. However, we used the fixed effects model as well as an instrumental variables approach to obtain accurate estimation and achieve our primary goal, which was to answer the role of the collective reputation of the online platform in live streaming. Since agri-food products live streaming research is still in the initial stage, many other valuable topics can be explored in future research. For example, the live streaming revenue may vary for different types of agri-food products. Future researchers can collect data from a wider range of types of agri-food products to explore which type of agri-food products revenue will increase most in live streaming and expand on the findings of our study.

References

Abbott, P. C., Hurt, C., and Tyner, W. E. (2008). What's driving food prices? Technical report.

Aderajew, T. S., Du, X., Pennings, J. M., and Trujillo-Barrera, A. (2020). Farm-level risk-balancing behavior and the role of latent heterogeneity. *Journal of Agricultural and Resource Economics*, 45(2):265–281.

Alcorn, C. (2021). Global food prices surge to their highest level in a decade. `https://www.cnn.com/2021/06/04/business/inflation-food-prices/index.html`.

AliResearch (2020). 2020 taobao live streaming economicv. `https://www.cbndata.com/report/2219/detail?isReading=report&page=1`.

Amblee, N. and Bui, T. (2008). Can brand reputation improve the odds of being reviewed on-line? *International Journal of Electronic Commerce*, 12(3):11–28.

Anderson, M. and Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, 122(563):957–989.

Aula, P. (2010). Social media, reputation risk and ambient publicity management. *Strategy & leadership*.

Baba, Y., Engle, R. F., Kraft, D. F., and Kroner, K. F. (1990). Multivariate simultaneous generalized arch. *Manuscript, University of California, San Diego, Department of Economics*.

Baek, J. and Koo, W. W. (2010). Analyzing factors affecting us food price inflation. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 58(3):303–320.

Bekkers, E., Brockmeier, M., Francois, J., and Yang, F. (2017). Local food prices and international price transmission. *World Development*, 96:216–230.

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS quarterly*, pages 351–370.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Brzustewicz, P. and Singh, A. (2021). Sustainable consumption in consumer behavior in the time of covid-19: Topic modeling on twitter data using lda. *Energies*, 14(18):5787.

Burton, S., Howlett, E., and Tangari, A. H. (2009). Food for thought: how will the nutrition labeling of quick service restaurant menu items influence consumersâ product evaluations, purchase intentions, and choices? *Journal of Retailing*, 85(3):258–273.

Cai, R. and Chi, C. G.-Q. (2021). Pictures vs. reality: roles of disconfirmation magnitude, disconfirmation sensitivity, and branding. *International Journal of Hospitality Management*, 98:103040.

Chen, C.-D., Zhao, Q., and Wang, J.-L. (2020). How livestreaming increases product sales: role of trust transfer and elaboration likelihood model. *Behaviour & Information Technology*, pages 1–16.

Chen, P.-Y., Chang, C.-L., Chen, C.-C., and McAleer, M. (2012). Modelling the effects of oil prices on global fertilizer prices and volatility. *Journal of Risk and Financial Management*, 5(1):78–114.

Chowdhury, M. A. F., Meo, M. S., Uddin, A., and Haque, M. M. (2021). Asymmetric effect of energy price on commodity price: New evidence from nardl and time frequency wavelet approaches. *Energy*, 231:120934.

Clement Addo, P., Fang, J., Asare, A. O., and Kulbo, N. B. (2021). Customer engagement and purchase intention in live-streaming digital marketing platforms. *The Service Industries Journal*, pages 1–20.

De Gorter, H. and Just, D. R. (2008). " water" in the us ethanol tax credit and mandate: Implications for rectangular deadweight costs and the corn-oil price relationship. *Review of Agricultural Economics*, 30(3):397–410.

Dillon, B. M. and Barrett, C. B. (2016). Global oil prices and local food prices: Evidence from east africa. *American Journal of Agricultural Economics*, 98(1):154–171.

Du, X., Khanna, M., Lu, L., Yang, X., and Zilberman, D. (2017). Contracting in the biofuel sector. In *Handbook of Bioenergy Economics and Policy: Volume II*, pages 401–425. Springer.

Duarte Alonso, A., O'neill, M., Liu, Y., and O'shea, M. (2013). Factors driving consumer restaurant choice: An exploratory study from the southeastern united states. *Journal of Hospitality Marketing & Management*, 22(5):547–567.

Dutta, A., Bouri, E., Saeed, T., and Vo, X. V. (2021). Crude oil volatility and the biodiesel feedstock market in malaysia during the 2014 oil price decline and the covid-19 outbreak. *Fuel*, 292:120221.

EIA (2021). Energy information administration (eia) forecasts u.s. winter natural gas bills will be 30% higher than last winter. `https://www.eia.gov/todayinenergy/detail.php?id=50076`.

Eissa, M. A. and Al Refai, H. (2019). Modelling the symmetric and asymmetric relationships between oil prices and those of corn, barley, and rapeseed oil. *Resources Policy*, 64:101511.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric theory*, 11(1):122–150.

Etienne, X. L., Trujillo-Barrera, A., and Wiggins, S. (2016). Price and volatility transmissions between natural gas, fertilizer, and corn markets. *Agricultural Finance Review*.

Fafchamps, M., Hill, R. V., and Minten, B. (2008). Quality control in nonstaple food markets: evidence from india. *Agricultural Economics*, 38(3):251–266.

Fernandez-Perez, A., Frijns, B., and Tourani-Rad, A. (2016). Contemporaneous interactions among fuel, biofuel and agricultural commodities. *Energy Economics*, 58:1–10.

Ferreira, P., Almeida, D., Dionísio, A., Bouri, E., and Quintino, D. (2022). Energy markets–who are the influencers? *Energy*, 239:121962.

Fowowe, B. (2016). Do oil prices drive agricultural commodity prices? evidence from south africa. *Energy*, 104:149–157.

Frick, B. and Simmons, R. (2013). The impact of individual and collective reputation on wine prices: empirical evidence from the mosel valley. *Journal of Business Economics*, 83(2):101–119.

Gan, Q., Ferns, B. H., Yu, Y., and Jin, L. (2017). A text mining and multidimensional sentiment analysis of online restaurant reviews. *Journal of Quality Assurance in Hospitality & Tourism*, 18(4):465–492.

Gardebroek, C. and Hernandez, M. A. (2013). Do energy prices stimulate food price volatility? examining volatility transmission between us oil, ethanol and corn markets. *Energy economics*, 40:119–129.

Gerten, J. and Topolinski, S. (2019). Shades of surprise: Assessing surprise as a function of degree of deviance and expectation constraints. *Cognition*, 192:103986.

Gilbert, C. L. (2010). How to understand high food prices. *Journal of agricultural economics*, 61(2):398–425.

Grossman, D. A. and Frieder, O. (2004). *Information retrieval: Algorithms and heuristics*, volume 15. Springer Science & Business Media.

Grunert, K. G., Bredahl, L., and Brunsø, K. (2004). Consumer perception of meat quality and implications for product development in the meat sectorâa review. *Meat science*, 66(2):259–272.

Ha, J., Park, K., and Park, J. (2016). Which restaurant should i choose? herd behavior in the restaurant industry. *Journal of foodservice business research*, 19(4):396–412.

Hatzenbuehler, P. L., Du, X., and Painter, K. (2021). Price transmission with sparse market information: The case of united states chickpeas. *Agribusiness*, 37(3):665–682.

Heng, Y., Gao, Z., Jiang, Y., and Chen, X. (2018). Exploring hidden factors behind online food shopping from amazon reviews: A topic mining approach. *Journal of Retailing and Consumer Services*, 42:161–168.

Hlee, S., Lee, J., Yang, S.-B., and Koo, C. (2019). The moderating effect of restaurant type on hedonic versus utilitarian review evaluations. *International Journal of Hospitality Management*, 77:195–206.

Hu, L., Sun, A., and Liu, Y. (2014). Your neighbors affect your ratings: on geographical neighborhood influence to rating prediction. In *Proceedings of the 37th*

international ACM SIGIR conference on Research & development in information retrieval, pages 345–354.

Hu, M. and Chaudhry, S. S. (2020). Enhancing consumer engagement in e-commerce live streaming via relational bonds. *Internet Research*.

Huang, J., Rogers, S., and Joo, E. (2014). Improving restaurants by extracting subtopics from yelp reviews. *iConference 2014 (Social Media Expo)*.

Ibrahim, M. H. (2015). Oil and food prices in malaysia: a nonlinear ardl analysis. *Agricultural and Food Economics*, 3(1):1–14.

IMF (2021). International monetary fund primary commodity prices. `https://data.imf.org/?sk=471DDDF8-D8A7-499A-81BA-5B332C01F8B9&sId=1547558078595`.

Ismail, A., Ihsan, H., Khan, S. A., and Jabeen, M. (2017). Price volatility of food and agricultural commodities: A case study of pakistan. *Journal of Economic Cooperation & Development*, 38(3):77–120.

Jacks, D. S., O'rourke, K. H., and Williamson, J. G. (2011). Commodity price volatility and world market integration since 1700. *Review of Economics and Statistics*, 93(3):800–813.

Ji, Q., Bouri, E., Roubaud, D., and Shahzad, S. J. H. (2018). Risk spillover between energy and agricultural commodity markets: A dependence-switching covar-copula model. *Energy Economics*, 75:14–27.

Jia, S., Zhang, X., Wang, X., and Liu, Y. (2018). Fake reviews detection based on lda. In *2018 4th International Conference on Information Management (ICIM)*, pages 280–283. IEEE.

Jia, S. S. (2021). Analyzing restaurant customersâ evolution of dining patterns

and satisfaction during covid-19 for sustainable business insights. *Sustainability*, 13(9):4981.

Jurafsky, D., Chahuneau, V., Routledge, B. R., and Smith, N. A. (2014). Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*.

Kalkuhl, M., von Braun, J., and Torero, M. (2016). Volatile and extreme food prices, food security, and policy: an overview. *Food price volatility and its implications for food security and policy*, pages 3–31.

Kang, K., Lu, J., Guo, L., and Li, W. (2021). The dynamic effect of interactivity on customer engagement behavior through tie strength: Evidence from live streaming commerce platforms. *International Journal of Information Management*, 56:102251.

Klostermann, J., Plumeyer, A., Böger, D., and Decker, R. (2018). Extracting brand information from social networks: Integrating image, text, and social tagging data. *International Journal of Research in Marketing*, 35(4):538–556.

Koirala, K. H., Mishra, A. K., D'Antoni, J. M., and Mehlhorn, J. E. (2015). Energy prices and agricultural commodity prices: Testing correlation using copulas method. *Energy*, 81:430–436.

Li, Y. and Peng, Y. (2021). What drives gift-giving intention in live streaming? the perspectives of emotional attachment and flow experience. *International Journal of Human–Computer Interaction*, pages 1–13.

Lian, J., Zhang, F., Xie, X., and Sun, G. (2017). Restaurant survival analysis with heterogeneous information. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 993–1002.

Lin, L., Gang, R., Hong, T., and Yang, S. B. (2019). Exploring simultaneous presentation in online restaurant reviews: An analysis of textual and visual content. *Asia Pacific Journal of Information Systems*, 29(2):181–202.

Lin, Z., Li, D., Janamanchi, B., and Huang, W. (2006). Reputation distribution and consumer-to-consumer online auction market structure: an exploratory study. *Decision Support Systems*, 41(2):435–448.

Liu, H., Feng, S., and Hu, X. S. (2022). Process vs. outcome: Effects of food photo types in online restaurant reviews on consumersâ purchase intention. *International journal of hospitality management*, 102:103179.

Lu, L., Nguyen, R., Rahman, M. M., and Winfree, J. (2021). Demand shocks and supply chain resilience: An agent based modelling approach and application to the potato supply chain. Technical report, National Bureau of Economic Research.

Lu, L., Tian, G., and Hatzenbuehler, P. (2022). How agricultural economists are using big data: a review. *China Agricultural Economic Review*.

Lu, L. and Winfree, J. (2021). Demand shocks and supply chain flexibility. In *NBER Conference âRisks in Agricultural Supply Chainsâ*. Springer Berlin/Heidelberg.

Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp. com. *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper*, (12-016).

Lucotte, Y. (2016). Co-movements between crude oil and food prices: A post-commodity boom perspective. *Economics Letters*, 147:142–147.

Luo, Y. and Xu, X. (2019). Predicting the helpfulness of online restaurant reviews using different machine learning algorithms: A case study of yelp. *Sustainability*, 11(19):5254.

Meyer, D. F., Sanusi, K. A., and Hassan, A. (2018). Analysis of the asymmetric impacts of oil prices on food prices in oil-exporting, developing countries. *Journal of International Studies*, 11(3).

Miller, R. K. (2021). The 2013 restaurant, food beverage market research handbook.

Netzer, O., Lemaire, A., and Herzenstein, M. (2019). When words sweat: Identifying signals for loan default in the text of loan applications. *Journal of Marketing Research*, 56(6):960–980.

Neubig, G. (2017). Neural machine translation and sequence-to-sequence models: A tutorial. *arXiv preprint arXiv:1703.01619*.

Newman, D., Lau, J. H., Grieser, K., and Baldwin, T. (2010a). Automatic evaluation of topic coherence. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pages 100–108.

Newman, D., Noh, Y., Talley, E., Karimi, S., and Baldwin, T. (2010b). Evaluating topic models for digital libraries. In *Proceedings of the 10th annual joint conference on Digital libraries*, pages 215–224.

Nosko, C. and Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment. Technical report, National Bureau of Economic Research.

Nwoko, I. C., Aye, G. C., and Asogwa, B. C. (2016). Oil price and food price volatility dynamics: The case of nigeria. *Cogent Food & Agriculture*, 2(1).

Nyga-Łukaszewska, H. and Aruga, K. (2020). Energy prices and covid-immunity: The case of crude oil and natural gas prices in the us and japan. *Energies*, 13(23):6300.

Oliveira, B. and Casais, B. (2018). The importance of user-generated photos in restaurant selection. *Journal of Hospitality and Tourism Technology*.

Ostrom, E. (2010). Analyzing collective action. *Agricultural economics*, 41:155–166.

Peng, C.-H., Yin, D., Wei, C.-P., and Zhang, H. (2014). How and when review length and emotional intensity influence review helpfulness: Empirical evidence from epinions. com.

Phua, J. and Ahn, S. J. (2016). Explicating the âlikeâon facebook brand pages: The effect of intensity of facebook use, number of overall âlikesâ, and number of friends'âlikesâ on consumers' brand outcomes. *Journal of Marketing Communications*, 22(5):544–559.

Pinto Gurdiel, L., Morales Mediano, J., and Cifuentes Quintero, J. A. (2021). A comparison study between coherence and perplexity for determining the number of topics in practitioners interviews analysis.

Piqueras-Fiszman, B. and Spence, C. (2014). *The perfect meal: the multisensory science of food and dining*. John Wiley & Sons.

Pizam, A. and Milman, A. (1993). Predicting satisfaction among first time visitors to a destination by using the expectancy disconfirmation theory. *International Journal of Hospitality Management*, 12(2):197–209.

Poushneh, A. and Rajabi, R. (2022). Can reviews predict reviewersâ numerical ratings? the underlying mechanisms of customersâ decisions to rate products using latent dirichlet allocation (lda). *Journal of Consumer Marketing*.

Radmehr, R. and Rastegari Henneberry, S. (2020). Energy price policies and food prices: Empirical evidence from iran. *Energies*, 13(15):4031.

Ramage, D., Dumais, S., and Liebling, D. (2010). Characterizing microblogs with topic models. In *Fourth international AAAI conference on weblogs and social media*.

Reardon, T., Belton, B., Liverpool-Tasie, L. S. O., Lu, L., Nuthalapati, C. S., Tasie, O., and Zilberman, D. (2021). E-commerce's fast-tracking diffusion and adaptation in developing countries. *Applied Economic Perspectives and Policy*, 43(4):1243–1259.

Ren, M., Vu, H. Q., Li, G., and Law, R. (2021). Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: implications for hospitality marketers. *Journal of Hospitality Marketing & Management*, 30(1):96–119.

Resano, H., Perez-Cueto, F. J., de Barcellos, M. D., Veflen-Olsen, N., Grunert, K. G., and Verbeke, W. (2011). Consumer satisfaction with pork meat and derived products in five european countries. *Appetite*, 56(1):167–170.

Resano, H., Sanjuán, A. I., and Albisu, L. M. (2007). Consumersâ acceptability of cured ham in spain and the influence of information. *Food Quality and Preference*, 18(8):1064–1076.

Rezitis, A. N. (2015). The relationship between agricultural commodity prices, crude oil prices and us dollar exchange rates: A panel var approach and causality analysis. *International Review of Applied Economics*, 29(3):403–434.

Roman, M., Górecka, A., and Domagała, J. (2020). The linkages between crude oil and food prices. *Energies*, 13(24):6545.

Saghaian, S., Nemati, M., Walters, C., and Chen, B. (2018). Asymmetric price volatility transmission between us biofuel, corn, and oil markets. *Journal of Agricultural and Resource Economics*, 43(1):46.

Sanyal, P., Malczynski, L. A., and Kaplan, P. (2015). Impact of energy price variability on global fertilizer price: application of alternative volatility models. *Sustainable Agriculture Research*, 4(526-2016-37964).

Seo, B. (2007). Asymptotic distribution of the cointegrating vector estimator in error correction models with conditional heteroskedasticity. *Journal of Econometrics*, 137(1):68–111.

Seo, Y., Kim, J., Choi, Y. K., and Li, X. (2019). In âlikesâ we trust: likes, disclosures and firm-serving motives on social media. *European Journal of Marketing*.

Serra, T., Zilberman, D., and Gil, J. (2011). Price volatility in ethanol markets. *European review of agricultural economics*, 38(2):259–280.

Shahnoushi, N., Sayed, S., Hezareh, R., and Tirgari Seraji, M. (2017). Investigation of relationship between world food prices and energy price: A panel sur approach. Technical report.

Siami-Namini, S. et al. (2019). Volatility transmission among oil price, exchange rate and agricultural commodities prices. *Applied Economics and Finance*, 6(4):41–61.

Su, C. W., Wang, X.-Q., Tao, R., and Oana-Ramona, L. (2019). Do oil prices drive agricultural commodity prices? further evidence in a global bio-energy context. *Energy*, 172:691–701.

Taghizadeh-Hesary, F., Rasoulinezhad, E., and Yoshino, N. (2018). Volatility linkages between energy and food prices: Case of selected asian countries.

Taghizadeh-Hesary, F., Rasoulinezhad, E., and Yoshino, N. (2019). Energy and food security: Linkages through price volatility. *Energy Policy*, 128:796–806.

Taobao (2013). Tmall seller's naming rules. `https://rule.tmall.com/tdetail-37.htm?spm=a2177.7731969.0.0.1168c32fOVuWXN&tag=self`.

Taobao (2019). Tmall 2020 annual software service fee, discounts and settlement standards. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Taobao (2020a). 2020 tmall sellers evaluation standards. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Taobao (2020b). Tmall seller's deposit rules. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Tian, G., Lu, L., and McIntosh, C. (2021). What factors affect consumersâ dining sentiments and their ratings: Evidence from restaurant online review data. *Food Quality and Preference*, 88:104060.

Tirole, J. (1993). A theory of collective reputations. *Research Papers in Economics University of Stockholm*, 9.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies*, 63(1):1–22.

Tirunillai, S. and Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of marketing research*, 51(4):463–479.

Tzimiropoulos, G. and Pantic, M. (2017). Fast algorithms for fitting active appearance models to unconstrained images. *International journal of computer vision*, 122(1):17–33.

Verbeke, W. (2005). Agriculture and the food industry in the information age. *European review of agricultural economics*, 32(3):347–368.

Wang, F., Chen, L., and Pan, W. (2016). Where to place your next restaurant? optimal restaurant placement via leveraging user-generated reviews. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 2371–2376.

Winfree, J. A. and McCluskey, J. J. (2005). Collective reputation and quality. *american Journal of agricultural Economics*, 87(1):206–213.

Wongkitrungrueng, A. and Assarut, N. (2018). The role of live streaming in building consumer trust and engagement with social commerce sellers. *Journal of Business Research*.

Wongkitrungrueng, A., Dehouche, N., and Assarut, N. (2020). Live streaming commerce from the sellersâ perspective: implications for online relationship marketing. *Journal of Marketing Management*, 36(5-6):488–518.

Wongpiyabovorn, O. (2021). Changes in the relationship between nitrogen fertilizer and natural gas prices in the us.

WorldBank (2020). Fertilizer prices to rise moderately in 2021. `https://blogs.worldbank.org/opendata/fertilizer-prices-rise-moderately-2021`.

Xiao, L., Fu, B., and Liu, W. (2018). Understanding consumer repurchase intention on o2o platforms: an integrated model of network externalities and trust transfer theory. *Service Business*, 12(4):731–756.

Xu, X., Wu, J.-H., and Li, Q. (2020). What drives consumer shopping behavior in live streaming commerce? *Journal of Electronic Commerce Research*, 21(3):144–167.

Yang, S.-B., Hlee, S., Lee, J., and Koo, C. (2017a). An empirical examination of online restaurant reviews on yelp. com: A dual coding theory perspective. *International Journal of Contemporary Hospitality Management*.

Yang, S.-B., Shin, S.-H., Joun, Y., and Koo, C. (2017b). Exploring the comparative importance of online hotel reviewsâ heuristic attributes in review helpfulness: a conjoint analysis approach. *Journal of Travel & Tourism Marketing*, 34(7):963–985.

Zhang, M. and Luo, L. (2022). Can consumer-posted photos serve as a leading indicator of restaurant survival? evidence from yelp. *Management Science*.

Zhou, S., Kan, P., Huang, Q., and Silbernagel, J. (2021). A guided latent dirichlet allocation approach to investigate real-time latent topics of twitter data during hurricane laura. *Journal of Information Science*, page 01655515211007724.

Figure 2.1: Customer Decision Making Process in Live Streaming

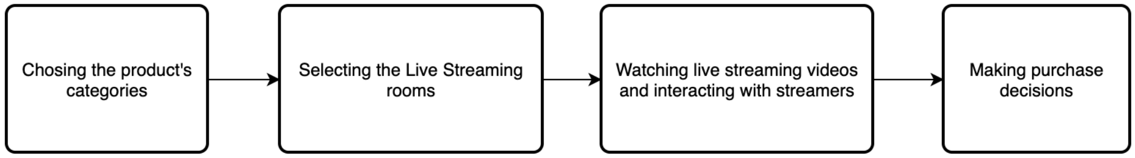


Table 2.1: Summary Statistics

Variables	Taobao Marketplace Sellers (N=1,538)		Tmall Sellers (N=7,976)		Total (N=9,505)	
	Mean	SD	Mean	SD	Mean	SD
Outcome variables						
Revenue (yuan)	1,279	14,381	9,188	55,858	7,908	51,547
Explanatory variables						
Live Streaming Duration (Hours)	11	12	15	18	14	17
Audience (10k viewers)	6.7	9.2	6.5	78.7	6.5	72.1
ThumbsUp (10k times)	13.1	28.4	6.5	62.2	7.5	58.2
Sellerâs Service Rating	4.7	0.1	4.6	0.1	4.6	0.1
Instrumental variables						
Deposit (yuan)	8,388	17,095	83,507	25,073	71,352	36,600
Fruit Type						
Apple	395 (25.7%)		1,861 (23.4%)		2,256 (23.7%)	
Coconut	153 (10.0%)		444 (5.6%)		597 (6.3%)	
Dragonfruit	212 (13.8%)		961 (12.1%)		1,173 (12.3%)	
Jujube	64 (4.2%)		271 (3.4%)		335 (3.5%)	
Kiwi	80 (5.2%)		766 (9.6%)		846 (8.9%)	
Lemon	158 (10.3%)		574 (7.2%)		732 (7.7%)	
MandarinOrange	145 (9.4%)		673 (8.4%)		818 (8.6%)	
Mango	105 (6.8%)		356 (4.5%)		461 (4.9%)	
Passionfruit	44 (2.9%)		936 (11.7%)		980 (10.3%)	
Pear	180 (11.7%)		1,127 (14.1%)		1,307 (13.8%)	

Figure 2.2: Agri-food Products Live Streaming Rooms

Figure 2.3: Watching the live streaming videos and interacting with the streamers

Figure 2.4: Transaction page within live streaming room

Figure 2.5: Transaction page within product detailed page

Table 2.2: OLS Fixed Effect

Variables	(1) Revenue	(2) Revenue	(3) Revenue	(4) Revenue
Tmall	4,783**	6,775***	4,990***	5,611***
	(1,967)	(1,819)	(1,809)	(1,860)
Price	155.0***	94.52***	69.50**	60.11*
	(32.65)	(30.22)	(30.02)	(30.73)
ThumbsUp		332.2***	326.6***	326.0***
		(8.284)	(8.223)	(8.235)
Live Duration			393.0***	394.5***
			(30.48)	(30.49)
Service Rating				7,516
				(5,258)
Constant	-3,863	-12,985	-14,448	-48,707
	(29,153)	(26,949)	(26,717)	(35,889)
N	9,505	9,505	9,505	9,505
R-squared	0.049	0.187	0.201	0.202
Fruit Type*Origin Dummy	YES	YES	YES	YES

Table 2.3: IV Regression

Variables	(1) Revenue	(2) Revenue	(3) Revenue	(4) Revenue
Panel A: Two-Stage Least Squares				
Tmall	4,157	7,893***	4,968**	7,059***
	(2,590)	(2,396)	(2,387)	(2,424)
Price	157.3***	90.31***	69.58**	53.90*
	(33.14)	(30.67)	(30.45)	(31.33)
ThumbsUp		332.3***	326.6***	326.1***
		(8.254)	(8.194)	(8.204)
Live Duration			393.1***	392.9***
			(30.42)	(30.42)
Service Rating				8,472
				(5,336)
Constant	-3,938	-12,854	-14,451	-52,901
	(29,044)	(26,847)	(26,613)	(36,034)
Observations	9,505	9,505	9,505	9,505
R^2	0.049	0.187	0.201	0.202
Type*Origin Dummy	YES	YES	YES	YES
Panel B: First Stage for Tmall				
Variables	(1) Tmall	(2) Tmall	(3) Tmall	(4) Tmall
Security Deposit	6.96e-06***	6.96e-06***	6.95e-06***	8.05e-06***
	(6.19e-08)	(6.19e-08)	(6.22e-08)	(7.00e-08)
Price	0.00317***	0.00317***	0.00316***	0.00248***
	(0.000109)	(0.000109)	(0.000109)	(0.000107)
ThumbsUp		9.13e-06	7.95e-06	-3.02e-05
		(3.07e-05)	(3.07e-05)	(2.94e-05)
Live Duration			8.15e-05	8.10e-05
			(0.000114)	(0.000109)
Service Rating				0.641***
				(0.0215)
Constant	-0.116	-0.116	-0.116	-3.043***
	(0.0998)	(0.0998)	(0.0998)	(0.137)
Observations	9,505	9,505	9,505	9,505
$F-stat$	475	469	462	512
R^2	0.782	0.782	0.782	0.801
Type*Origin Dummy	YES	YES	YES	YES

References

Abbott, P. C., Hurt, C., and Tyner, W. E. (2008). What's driving food prices? Technical report.

Aderajew, T. S., Du, X., Pennings, J. M., and Trujillo-Barrera, A. (2020). Farm-level risk-balancing behavior and the role of latent heterogeneity. *Journal of Agricultural and Resource Economics*, 45(2):265–281.

Alcorn, C. (2021). Global food prices surge to their highest level in a decade. `https://www.cnn.com/2021/06/04/business/inflation-food-prices/index.html`.

AliResearch (2020). 2020 taobao live streaming economicv. `https://www.cbndata.com/report/2219/detail?isReading=report&page=1`.

Amblee, N. and Bui, T. (2008). Can brand reputation improve the odds of being reviewed on-line? *International Journal of Electronic Commerce*, 12(3):11–28.

Anderson, M. and Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, 122(563):957–989.

Aula, P. (2010). Social media, reputation risk and ambient publicity management. *Strategy & leadership*.

Baba, Y., Engle, R. F., Kraft, D. F., and Kroner, K. F. (1990). Multivariate simultaneous generalized arch. *Manuscript, University of California, San Diego, Department of Economics*.

Baek, J. and Koo, W. W. (2010). Analyzing factors affecting us food price inflation. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 58(3):303–320.

Bekkers, E., Brockmeier, M., Francois, J., and Yang, F. (2017). Local food prices and international price transmission. *World Development*, 96:216–230.

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS quarterly*, pages 351–370.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Brzustewicz, P. and Singh, A. (2021). Sustainable consumption in consumer behavior in the time of covid-19: Topic modeling on twitter data using lda. *Energies*, 14(18):5787.

Burton, S., Howlett, E., and Tangari, A. H. (2009). Food for thought: how will the nutrition labeling of quick service restaurant menu items influence consumersâ product evaluations, purchase intentions, and choices? *Journal of Retailing*, 85(3):258–273.

Cai, R. and Chi, C. G.-Q. (2021). Pictures vs. reality: roles of disconfirmation magnitude, disconfirmation sensitivity, and branding. *International Journal of Hospitality Management*, 98:103040.

Chen, C.-D., Zhao, Q., and Wang, J.-L. (2020). How livestreaming increases product sales: role of trust transfer and elaboration likelihood model. *Behaviour & Information Technology*, pages 1–16.

Chen, P.-Y., Chang, C.-L., Chen, C.-C., and McAleer, M. (2012). Modelling the effects of oil prices on global fertilizer prices and volatility. *Journal of Risk and Financial Management*, 5(1):78–114.

Chowdhury, M. A. F., Meo, M. S., Uddin, A., and Haque, M. M. (2021). Asymmetric effect of energy price on commodity price: New evidence from nardl and time frequency wavelet approaches. *Energy*, 231:120934.

Clement Addo, P., Fang, J., Asare, A. O., and Kulbo, N. B. (2021). Customer engagement and purchase intention in live-streaming digital marketing platforms. *The Service Industries Journal*, pages 1–20.

De Gorter, H. and Just, D. R. (2008). " water" in the us ethanol tax credit and mandate: Implications for rectangular deadweight costs and the corn-oil price relationship. *Review of Agricultural Economics*, 30(3):397–410.

Dillon, B. M. and Barrett, C. B. (2016). Global oil prices and local food prices: Evidence from east africa. *American Journal of Agricultural Economics*, 98(1):154–171.

Du, X., Khanna, M., Lu, L., Yang, X., and Zilberman, D. (2017). Contracting in the biofuel sector. In *Handbook of Bioenergy Economics and Policy: Volume II*, pages 401–425. Springer.

Duarte Alonso, A., O'neill, M., Liu, Y., and O'shea, M. (2013). Factors driving consumer restaurant choice: An exploratory study from the southeastern united states. *Journal of Hospitality Marketing & Management*, 22(5):547–567.

Dutta, A., Bouri, E., Saeed, T., and Vo, X. V. (2021). Crude oil volatility and the biodiesel feedstock market in malaysia during the 2014 oil price decline and the covid-19 outbreak. *Fuel*, 292:120221.

EIA (2021). Energy information administration (eia) forecasts u.s. winter natural gas bills will be 30% higher than last winter. `https://www.eia.gov/todayinenergy/detail.php?id=50076`.

Eissa, M. A. and Al Refai, H. (2019). Modelling the symmetric and asymmetric relationships between oil prices and those of corn, barley, and rapeseed oil. *Resources Policy*, 64:101511.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric theory*, 11(1):122–150.

Etienne, X. L., Trujillo-Barrera, A., and Wiggins, S. (2016). Price and volatility transmissions between natural gas, fertilizer, and corn markets. *Agricultural Finance Review*.

Fafchamps, M., Hill, R. V., and Minten, B. (2008). Quality control in nonstaple food markets: evidence from india. *Agricultural Economics*, 38(3):251–266.

Fernandez-Perez, A., Frijns, B., and Tourani-Rad, A. (2016). Contemporaneous interactions among fuel, biofuel and agricultural commodities. *Energy Economics*, 58:1–10.

Ferreira, P., Almeida, D., Dionísio, A., Bouri, E., and Quintino, D. (2022). Energy markets–who are the influencers? *Energy*, 239:121962.

Fowowe, B. (2016). Do oil prices drive agricultural commodity prices? evidence from south africa. *Energy*, 104:149–157.

Frick, B. and Simmons, R. (2013). The impact of individual and collective reputation on wine prices: empirical evidence from the mosel valley. *Journal of Business Economics*, 83(2):101–119.

Gan, Q., Ferns, B. H., Yu, Y., and Jin, L. (2017). A text mining and multidimensional sentiment analysis of online restaurant reviews. *Journal of Quality Assurance in Hospitality & Tourism*, 18(4):465–492.

Gardebroek, C. and Hernandez, M. A. (2013). Do energy prices stimulate food price volatility? examining volatility transmission between us oil, ethanol and corn markets. *Energy economics*, 40:119–129.

Gerten, J. and Topolinski, S. (2019). Shades of surprise: Assessing surprise as a function of degree of deviance and expectation constraints. *Cognition*, 192:103986.

Gilbert, C. L. (2010). How to understand high food prices. *Journal of agricultural economics*, 61(2):398–425.

Grossman, D. A. and Frieder, O. (2004). *Information retrieval: Algorithms and heuristics*, volume 15. Springer Science & Business Media.

Grunert, K. G., Bredahl, L., and Brunsø, K. (2004). Consumer perception of meat quality and implications for product development in the meat sectorâa review. *Meat science*, 66(2):259–272.

Ha, J., Park, K., and Park, J. (2016). Which restaurant should i choose? herd behavior in the restaurant industry. *Journal of foodservice business research*, 19(4):396–412.

Hatzenbuehler, P. L., Du, X., and Painter, K. (2021). Price transmission with sparse market information: The case of united states chickpeas. *Agribusiness*, 37(3):665–682.

Heng, Y., Gao, Z., Jiang, Y., and Chen, X. (2018). Exploring hidden factors behind online food shopping from amazon reviews: A topic mining approach. *Journal of Retailing and Consumer Services*, 42:161–168.

Hlee, S., Lee, J., Yang, S.-B., and Koo, C. (2019). The moderating effect of restaurant type on hedonic versus utilitarian review evaluations. *International Journal of Hospitality Management*, 77:195–206.

Hu, L., Sun, A., and Liu, Y. (2014). Your neighbors affect your ratings: on geographical neighborhood influence to rating prediction. In *Proceedings of the 37th*

international ACM SIGIR conference on Research & development in information retrieval, pages 345–354.

Hu, M. and Chaudhry, S. S. (2020). Enhancing consumer engagement in e-commerce live streaming via relational bonds. *Internet Research*.

Huang, J., Rogers, S., and Joo, E. (2014). Improving restaurants by extracting subtopics from yelp reviews. *iConference 2014 (Social Media Expo)*.

Ibrahim, M. H. (2015). Oil and food prices in malaysia: a nonlinear ardl analysis. *Agricultural and Food Economics*, 3(1):1–14.

IMF (2021). International monetary fund primary commodity prices. `https://data.imf.org/?sk=471DDDF8-D8A7-499A-81BA-5B332C01F8B9&sId=1547558078595`.

Ismail, A., Ihsan, H., Khan, S. A., and Jabeen, M. (2017). Price volatility of food and agricultural commodities: A case study of pakistan. *Journal of Economic Cooperation & Development*, 38(3):77–120.

Jacks, D. S., O'rourke, K. H., and Williamson, J. G. (2011). Commodity price volatility and world market integration since 1700. *Review of Economics and Statistics*, 93(3):800–813.

Ji, Q., Bouri, E., Roubaud, D., and Shahzad, S. J. H. (2018). Risk spillover between energy and agricultural commodity markets: A dependence-switching covar-copula model. *Energy Economics*, 75:14–27.

Jia, S., Zhang, X., Wang, X., and Liu, Y. (2018). Fake reviews detection based on lda. In *2018 4th International Conference on Information Management (ICIM)*, pages 280–283. IEEE.

Jia, S. S. (2021). Analyzing restaurant customersâ evolution of dining patterns

and satisfaction during covid-19 for sustainable business insights. *Sustainability*, 13(9):4981.

Jurafsky, D., Chahuneau, V., Routledge, B. R., and Smith, N. A. (2014). Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*.

Kalkuhl, M., von Braun, J., and Torero, M. (2016). Volatile and extreme food prices, food security, and policy: an overview. *Food price volatility and its implications for food security and policy*, pages 3–31.

Kang, K., Lu, J., Guo, L., and Li, W. (2021). The dynamic effect of interactivity on customer engagement behavior through tie strength: Evidence from live streaming commerce platforms. *International Journal of Information Management*, 56:102251.

Klostermann, J., Plumeyer, A., Böger, D., and Decker, R. (2018). Extracting brand information from social networks: Integrating image, text, and social tagging data. *International Journal of Research in Marketing*, 35(4):538–556.

Koirala, K. H., Mishra, A. K., D'Antoni, J. M., and Mehlhorn, J. E. (2015). Energy prices and agricultural commodity prices: Testing correlation using copulas method. *Energy*, 81:430–436.

Li, Y. and Peng, Y. (2021). What drives gift-giving intention in live streaming? the perspectives of emotional attachment and flow experience. *International Journal of Human–Computer Interaction*, pages 1–13.

Lian, J., Zhang, F., Xie, X., and Sun, G. (2017). Restaurant survival analysis with heterogeneous information. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 993–1002.

Lin, L., Gang, R., Hong, T., and Yang, S. B. (2019). Exploring simultaneous presentation in online restaurant reviews: An analysis of textual and visual content. *Asia Pacific Journal of Information Systems*, 29(2):181–202.

Lin, Z., Li, D., Janamanchi, B., and Huang, W. (2006). Reputation distribution and consumer-to-consumer online auction market structure: an exploratory study. *Decision Support Systems*, 41(2):435–448.

Liu, H., Feng, S., and Hu, X. S. (2022). Process vs. outcome: Effects of food photo types in online restaurant reviews on consumersâ purchase intention. *International journal of hospitality management*, 102:103179.

Lu, L., Nguyen, R., Rahman, M. M., and Winfree, J. (2021). Demand shocks and supply chain resilience: An agent based modelling approach and application to the potato supply chain. Technical report, National Bureau of Economic Research.

Lu, L., Tian, G., and Hatzenbuehler, P. (2022). How agricultural economists are using big data: a review. *China Agricultural Economic Review*.

Lu, L. and Winfree, J. (2021). Demand shocks and supply chain flexibility. In *NBER Conference âRisks in Agricultural Supply Chainsâ*. Springer Berlin/Heidelberg.

Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp. com. *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper*, (12-016).

Lucotte, Y. (2016). Co-movements between crude oil and food prices: A post-commodity boom perspective. *Economics Letters*, 147:142–147.

Luo, Y. and Xu, X. (2019). Predicting the helpfulness of online restaurant reviews using different machine learning algorithms: A case study of yelp. *Sustainability*, 11(19):5254.

Meyer, D. F., Sanusi, K. A., and Hassan, A. (2018). Analysis of the asymmetric impacts of oil prices on food prices in oil-exporting, developing countries. *Journal of International Studies*, 11(3).

Miller, R. K. (2021). The 2013 restaurant, food beverage market research handbook.

Netzer, O., Lemaire, A., and Herzenstein, M. (2019). When words sweat: Identifying signals for loan default in the text of loan applications. *Journal of Marketing Research*, 56(6):960–980.

Neubig, G. (2017). Neural machine translation and sequence-to-sequence models: A tutorial. *arXiv preprint arXiv:1703.01619*.

Newman, D., Lau, J. H., Grieser, K., and Baldwin, T. (2010a). Automatic evaluation of topic coherence. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pages 100–108.

Newman, D., Noh, Y., Talley, E., Karimi, S., and Baldwin, T. (2010b). Evaluating topic models for digital libraries. In *Proceedings of the 10th annual joint conference on Digital libraries*, pages 215–224.

Nosko, C. and Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment. Technical report, National Bureau of Economic Research.

Nwoko, I. C., Aye, G. C., and Asogwa, B. C. (2016). Oil price and food price volatility dynamics: The case of nigeria. *Cogent Food & Agriculture*, 2(1).

Nyga-Łukaszewska, H. and Aruga, K. (2020). Energy prices and covid-immunity: The case of crude oil and natural gas prices in the us and japan. *Energies*, 13(23):6300.

Oliveira, B. and Casais, B. (2018). The importance of user-generated photos in restaurant selection. *Journal of Hospitality and Tourism Technology*.

Ostrom, E. (2010). Analyzing collective action. *Agricultural economics*, 41:155–166.

Peng, C.-H., Yin, D., Wei, C.-P., and Zhang, H. (2014). How and when review length and emotional intensity influence review helpfulness: Empirical evidence from epinions. com.

Phua, J. and Ahn, S. J. (2016). Explicating the âlikeâon facebook brand pages: The effect of intensity of facebook use, number of overall âlikesâ, and number of friends'âlikesâ on consumers' brand outcomes. *Journal of Marketing Communications*, 22(5):544–559.

Pinto Gurdiel, L., Morales Mediano, J., and Cifuentes Quintero, J. A. (2021). A comparison study between coherence and perplexity for determining the number of topics in practitioners interviews analysis.

Piqueras-Fiszman, B. and Spence, C. (2014). *The perfect meal: the multisensory science of food and dining*. John Wiley & Sons.

Pizam, A. and Milman, A. (1993). Predicting satisfaction among first time visitors to a destination by using the expectancy disconfirmation theory. *International Journal of Hospitality Management*, 12(2):197–209.

Poushneh, A. and Rajabi, R. (2022). Can reviews predict reviewersâ numerical ratings? the underlying mechanisms of customersâ decisions to rate products using latent dirichlet allocation (lda). *Journal of Consumer Marketing*.

Radmehr, R. and Rastegari Henneberry, S. (2020). Energy price policies and food prices: Empirical evidence from iran. *Energies*, 13(15):4031.

Ramage, D., Dumais, S., and Liebling, D. (2010). Characterizing microblogs with topic models. In *Fourth international AAAI conference on weblogs and social media*.

Reardon, T., Belton, B., Liverpool-Tasie, L. S. O., Lu, L., Nuthalapati, C. S., Tasie, O., and Zilberman, D. (2021). E-commerce's fast-tracking diffusion and adaptation in developing countries. *Applied Economic Perspectives and Policy*, 43(4):1243–1259.

Ren, M., Vu, H. Q., Li, G., and Law, R. (2021). Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: implications for hospitality marketers. *Journal of Hospitality Marketing & Management*, 30(1):96–119.

Resano, H., Perez-Cueto, F. J., de Barcellos, M. D., Veflen-Olsen, N., Grunert, K. G., and Verbeke, W. (2011). Consumer satisfaction with pork meat and derived products in five european countries. *Appetite*, 56(1):167–170.

Resano, H., Sanjuán, A. I., and Albisu, L. M. (2007). Consumersâ acceptability of cured ham in spain and the influence of information. *Food Quality and Preference*, 18(8):1064–1076.

Rezitis, A. N. (2015). The relationship between agricultural commodity prices, crude oil prices and us dollar exchange rates: A panel var approach and causality analysis. *International Review of Applied Economics*, 29(3):403–434.

Roman, M., Górecka, A., and Domagała, J. (2020). The linkages between crude oil and food prices. *Energies*, 13(24):6545.

Saghaian, S., Nemati, M., Walters, C., and Chen, B. (2018). Asymmetric price volatility transmission between us biofuel, corn, and oil markets. *Journal of Agricultural and Resource Economics*, 43(1):46.

Sanyal, P., Malczynski, L. A., and Kaplan, P. (2015). Impact of energy price variability on global fertilizer price: application of alternative volatility models. *Sustainable Agriculture Research*, 4(526-2016-37964).

Seo, B. (2007). Asymptotic distribution of the cointegrating vector estimator in error correction models with conditional heteroskedasticity. *Journal of Econometrics*, 137(1):68–111.

Seo, Y., Kim, J., Choi, Y. K., and Li, X. (2019). In âlikesâ we trust: likes, disclosures and firm-serving motives on social media. *European Journal of Marketing*.

Serra, T., Zilberman, D., and Gil, J. (2011). Price volatility in ethanol markets. *European review of agricultural economics*, 38(2):259–280.

Shahnoushi, N., Sayed, S., Hezareh, R., and Tirgari Seraji, M. (2017). Investigation of relationship between world food prices and energy price: A panel sur approach. Technical report.

Siami-Namini, S. et al. (2019). Volatility transmission among oil price, exchange rate and agricultural commodities prices. *Applied Economics and Finance*, 6(4):41–61.

Su, C. W., Wang, X.-Q., Tao, R., and Oana-Ramona, L. (2019). Do oil prices drive agricultural commodity prices? further evidence in a global bio-energy context. *Energy*, 172:691–701.

Taghizadeh-Hesary, F., Rasoulinezhad, E., and Yoshino, N. (2018). Volatility linkages between energy and food prices: Case of selected asian countries.

Taghizadeh-Hesary, F., Rasoulinezhad, E., and Yoshino, N. (2019). Energy and food security: Linkages through price volatility. *Energy Policy*, 128:796–806.

Taobao (2013). Tmall seller's naming rules. `https://rule.tmall.com/tdetail-37.htm?spm=a2177.7731969.0.0.1168c32fOVuWXN&tag=self`.

Taobao (2019). Tmall 2020 annual software service fee, discounts and settlement standards. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Taobao (2020a). 2020 tmall sellers evaluation standards. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Taobao (2020b). Tmall seller's deposit rules. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Tian, G., Lu, L., and McIntosh, C. (2021). What factors affect consumersâ dining sentiments and their ratings: Evidence from restaurant online review data. *Food Quality and Preference*, 88:104060.

Tirole, J. (1993). A theory of collective reputations. *Research Papers in Economics University of Stockholm*, 9.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies*, 63(1):1–22.

Tirunillai, S. and Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of marketing research*, 51(4):463–479.

Tzimiropoulos, G. and Pantic, M. (2017). Fast algorithms for fitting active appearance models to unconstrained images. *International journal of computer vision*, 122(1):17–33.

Verbeke, W. (2005). Agriculture and the food industry in the information age. *European review of agricultural economics*, 32(3):347–368.

Wang, F., Chen, L., and Pan, W. (2016). Where to place your next restaurant? optimal restaurant placement via leveraging user-generated reviews. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 2371–2376.

Winfree, J. A. and McCluskey, J. J. (2005). Collective reputation and quality. *american Journal of agricultural Economics*, 87(1):206–213.

Wongkitrungrueng, A. and Assarut, N. (2018). The role of live streaming in building consumer trust and engagement with social commerce sellers. *Journal of Business Research*.

Wongkitrungrueng, A., Dehouche, N., and Assarut, N. (2020). Live streaming commerce from the sellersâ perspective: implications for online relationship marketing. *Journal of Marketing Management*, 36(5-6):488–518.

Wongpiyabovorn, O. (2021). Changes in the relationship between nitrogen fertilizer and natural gas prices in the us.

WorldBank (2020). Fertilizer prices to rise moderately in 2021. `https://blogs.worldbank.org/opendata/fertilizer-prices-rise-moderately-2021`.

Xiao, L., Fu, B., and Liu, W. (2018). Understanding consumer repurchase intention on o2o platforms: an integrated model of network externalities and trust transfer theory. *Service Business*, 12(4):731–756.

Xu, X., Wu, J.-H., and Li, Q. (2020). What drives consumer shopping behavior in live streaming commerce? *Journal of Electronic Commerce Research*, 21(3):144–167.

Yang, S.-B., Hlee, S., Lee, J., and Koo, C. (2017a). An empirical examination of online restaurant reviews on yelp. com: A dual coding theory perspective. *International Journal of Contemporary Hospitality Management*.

Yang, S.-B., Shin, S.-H., Joun, Y., and Koo, C. (2017b). Exploring the comparative importance of online hotel reviewsâ heuristic attributes in review helpfulness: a conjoint analysis approach. *Journal of Travel & Tourism Marketing*, 34(7):963–985.

Zhang, M. and Luo, L. (2022). Can consumer-posted photos serve as a leading indicator of restaurant survival? evidence from yelp. *Management Science*.

Zhou, S., Kan, P., Huang, Q., and Silbernagel, J. (2021). A guided latent dirichlet allocation approach to investigate real-time latent topics of twitter data during hurricane laura. *Journal of Information Science*, page 01655515211007724.

Chapter 3

”Pictures are for reference only?” A large-scale photo content analysis in restaurant reviews

3.1 Introduction

With the development of consumer-generated restaurant review sites such as Yelp.com and OpenTable.com over recent decades, online restaurant rating (ORR) increasingly play a pivotal role for both business owners and customers. It usually based on a 5-point scale and directly influence restaurants in terms of revenues and reservations for business owners. For example, a half star increase will lead to 19% of reservation increasement(Anderson and Magruder, 2012) and one-star increase in reviews leads to a 5â9% increase in revenue for independent restaurants(Luca, 2016). Customers refer rating to help them expect quality and make dining decision. In fact, 80 % of customers tend to use a rating filter when searching for a restaurant[1]. The potential consumers may perceive a lower rating as a sign of poor restaurant quality or low popularity, and this would almost certainly impact their decision-making process(Ha et al., 2016).

While research on the factors influencing ORR has been highlighted in many areas, little research related to photo content analysis and ORR. A number of studies focus on text mining identified various elements or attributes that influence ORRs, such as food, service, context, price, and atmosphere(e.g., Gan et al. (2017)); text comment sentiments(e.g., Tian et al. (2021); Liu et al. (2022)), and geographical neighborhood(Hu et al., 2014). Compared to text review analysis, existing studies pay less attention to photo contents and have failed to explore the valuable information conveyed by business owner and customers’ photo content, which focus on

[1] https://www.reviewtrackers.com/blog/restaurant-star-rating/

the number of user-generated photos on review helpfulness and enjoyment instead of photo content(Hlee et al., 2019; Yang et al., 2017a).

Furthermore, the impact of photo content differences between customers and business owners on rating remains unknown. On the one hand, photos taken by business owners and other consumers in online platforms are important references for forming consumerâs quality expectation, which, in turn, affect customer satisfaction, trust, and behavioral intentions in the restaurant context(Cai and Chi, 2021). Two-thirds of avid restaurant customers say that they browse food pictures on social media and 75% of them pick a place to eat based on their evaluations of the pictures(Miller, 2021). On the other hand, how far the reality is apart from the expectation determines customer satisfaction and other post-consumption behaviors(Pizam and Milman, 1993; Burton et al., 2009), like their rating behaviors. Photos taken by customers reflect their actual perception of dining experience and preferences. Therefore, we still do not know how the different magnitudes of the expectancy disconfirmation between promotional photos by the business owners and actual consumption experience (reflected in the photos taken by customers) influence their rating behaviors and restaurant rating.

Our study seek to address the following important questions by digging deeply into photo content and understanding these questions provides business owners or managers with insights into customer's dining experience and their rating behaviors. (i) What are the major themes captured in customerâ photos? (ii)What are the differences between the projected photo in the promotional photos of restaurant managers and the perceived photo reflected in the photos of customer's? (iii) How does this photo content differences in different themes affect the online restaurant rating? and in which direction? (iv) Which themes differences contribute most to affect the restaurant rating?

To answer the questions above, we first utilized a web crawler to collect 343,337 photos (4,4316 for business owners, 338,095 for customers) from 303 restaurants in

Las Vegas between July 7, 2005, and March 14, 2022. In the first of part analysis, we adopt the Google Vision API to recognize the photo content of each photo content automatically. Then we identify the latent topic of each photo by building the Latent Dirichlet Allocation (LDA) modeling to answer the first question. In response to from the second to fourth questions, we calculate the photo content differences between business owners and customers for each restaurant based on framework by Ren et al. (2021). Finally, we merged the photo content different restaurant rating by unique business id, date, and rater's name. We conducted the OLS fixed-effect model to explore their relationship with restaurant ratings.

Results from the LDA model show that topics for all photos can be categorized into five interpretable topics: Ambiance, Dishes, Dessert & sweets, Protein Food, and Drinks & Drinkware. We find that photo content differences between business owners and customers in Dishes, Dessert & sweets, and Protein Food are positively associated with the restaurant rating. Differences in protein foods have the largest impact on restaurant ratings compared to other topics. This supports the conclusion by Zhang and Luo (2022), which shows that the proportion of food photos has the largest positive association with restaurant survival.

This paper makes two main contributions: first, our study fills the gap of lacking empirical evidence of the impact of photo content differences on rating in existing online restaurant review literature. To best of our knowledge,our paper is the first to adopt the visual recognition tool combining the topic modeling with an economic methodology to demonstrate photo content differences impact on restaurant rating; second, this paper has profound implications for business owners or managers that provide them with insights into consumerâs experience and preferences, which are useful for developing effective marketing strategies and online visual contents to attract future customers.

3.2 Background

3.2.1 Photo analysis in online restaurant reviews

Restaurant experiences are shared not only through related text descriptions but also through images of the food, drinks, or the restaurant environment that cannot be neglected(Yang et al., 2017b). With the easy access to the photos in online platforms, details of the product (e.g., food) are often provided in a picture, prompting customers to build a concrete expectation(Gerten and Topolinski, 2019). Previous studies regarding the image in online restaurant reviews focus on the number of photos instead of its content. For instance, Hlee et al. (2019) collected data from Yelp and found that the number of review photos positively affected review usefulness and enjoyment. This effect was stronger for casual restaurants than for luxury restaurants. The development of computer vision systems has provided increased flexibility and further automation options to researchers, thereby helping them find defects, sort products, and complete tasks faster and more efficiently than humans alone ever could (Tzimiropoulos and Pantic, 2017).

Recent years have seen a growing number of literature that employ image mining techniques in the context of online restaurant reviews to investigate the effects of visual content. For example, Yang et al. (2017a) combined review length and review readability with images of the physical environment, food, and beverages. They find that both aspects of imagery format (physical environment images and food and beverage images) positively relate to review enjoyment. In contrast, the review length and food and beverage images are the most critical factors affecting review usefulness and enjoyment. Lin et al. (2019) took one step further; they combined both text mining and image mining to investigate the similarity between textual-visual contents, which is the major factor in determining review usefulness but not review enjoyment.

Zhang and Luo (2022) expand their photos sample size from Yelp to 755,758 and further explore the relationship between photo attributions (e.g., brightness), content, and restaurant survival. Their results show that the informativeness of photos (e.g., the proportion of food photos) relates more to restaurant survival than photographic attributes (e.g., composition; brightness). In addition to the online platform like Yelp as the data source, Klostermann et al. (2018) focused on consumption experiences and collected 10,325 Instagram posts and images with the hashtag #mcdonalds, then used Google Could Vision API to identify 20 image clusters grouped into superordinate areas (e.g. fast food, illustrations, outdoor, selfies) that describe unique contents and contexts included in consumersâ created messages.

Yet research on the impact of specific photo content on rating in online reviews is rare. The purpose of our study is to fill this gap by exploring the effect of food photo type (process- vs. outcome-focused) in online restaurant reviews on customersâ purchase intention.

3.3 Data

We follow the road map (Figure 3.1) to answer the research question. As Figure 3.1 shows, we first develop a web crawler to get the original pixel of photos of restaurants in Las Vegas from Yelp.com and store them on a local disk by business id[2]. Next, photo content labels are obtained automatically by uploading all photos into Google Cloud Vision API. Google Cloud Vision API recognizes all photos and returns all photo labels as a format of the label; they a confidence score, and we store them as excel files by business ID. After that, we combine all excel files into a single file to obtain the raw dataset for photos. Then, we adopt the Latent Dirichlet Allocation(LDA) for identifying the latent topics of all photos after pre-processing the dataset. Finally, we merge the photo content dataset and its corresponding restaurant information by

[2]The business id is a unique id that represents the restaurant name.

matching their business ids, user names, and date. As a result, our dataset consists of two parts: the information about the restaurant in Las Vegas on Yelp.com, and each restaurant's online photo content (presented as labels, including both business owners and customers) recognized by Google Cloud Vision API.

3.3.1 Restaurant Information

The first dataset is from Tian et al. (2021), which is a panel dataset containing business attributes of 310 restaurants in Las Vegas from 2005 to 2017. This dataset contained business ID, restaurant name, historical rating categories (ranging from 0 to 5), and other 81 variables describing features such as whether the business has parking, pets allowed or not, or WiFi availability.

We selected Las Vegas as a target destination for our study rather than all ten cities in the Yelp dataset for two reasons. First, consumers are less likely to reside in such tourist city. Therefore, they rely more on online review platforms than on their past experiences or the experiences of other locals (Tian et al., 2021). As a result, text and photo reviews from tourists tend to be representative. Second, we do not have the historical restaurant information (e.g., rating) for cities other than Las Vegas.

3.3.2 Photo Collection

Although Yelp provides free access to the photo dataset, it contains a large part of the low pixel (lower than 30 Kb) of these photos, which creates a bias in recognizing photo content, and we can not identify whether a business owner or a customer took photos. Figure 2 (the original pixel of the picture on Yelp) and Figure 3 (low pixel photo from Yelp dataset) show an example of photo content recognition bias if we use low pixel photo from the Yelp public dataset. We can get more photo content labels and more precise recognition results that are not captured by low pixel photos from Yelp public dataset when uploading the original pixel photo into the vision recognition tool. The recognition result from Google Vision API clearly shows "White Rice" in the original pixel of the photo (pointed with a red circle in Figure 2), which is the same result as the human recognition. In contrast, the result from low pixel photos shows no such content. Besides, the photo of Las Vegas in a public dataset is less than 100, which accounts for less than 1% of the total number of photos.

We collect the original pixel of photos by developing a web crawler instead of using the public photo dataset provided by Yelp[3] to construct the photo content dataset. One of the advantages of using original pixel photos is that we can ensure accurate identification of the content of each photo and minimize recognition bias caused by low pixel photos. Second, we can know the photo taker by adding the code to identify the photo information. Yelp denotes every photo taken by business owners with the specific text description as "From the business owner" (as the Figure 3.4 shown)[4]. Otherwise, the photo without such a footnote is taken by customers. As a result, we can further explore the photo content differences between business owners and customers. Third, we can focus on our target city, Las Vegas, and photo analyze at a large scale. Based on the unique business ID, we collected 343,337 (4,431 for business owners, 338,095 for customers) photos from 303 restaurants between July 7, 2005 and March 14, 2022.

3.3.3 Photo Content Identification

Yelp has a massive amount of photos and classifies each of them into five categories, but there are limitations. Yelp has developed a deep learning model that yields a general photo classification as follows: food, drink, interior, outside, and menu[5]. This

[3]https://www.yelp.com/dataset

[4]When we inspect the Yelp web page, if the business owner took the photo, the web page element will show the "bizid"; if a customer took the photo, it will show its user id. Thus we can judge the photo takers by adding the corresponding code in Python

[5]$https://engineeringblog.yelp.com/2015/10/how-we-use-deep-learning-to-classify-business-photos-at-yelp.html$

general classification method, however, may not correctly categorize photo, capture the detailed content of a photo (e.g., subject feeling like smile or tableware) and hidden topics in photos. For instance, Yelp.com categorize the onion ring in Figure 3.5 as "Food"[6] rather than "Tableware or Drinkware". Moreover, comprehensive photo content analysis is challenging because online photos are usually available at large scales, and the captured contents are often diverse (Ren et al., 2021). Manual analysis is time-consuming and limited to a number of photos (Ma & Takagi, 2012).

Google Vision API[7] is used as the tool for automatic recognition of photo content at large scale in this study. Despite building deep learning models for visual recognition requiring considerable time and data with label annotations for training purposes, we have pre-trained tools like Google Vision API, Amazon Rekognition, or IBM Watson visual recognition. We choose Google Vision API because it has a pre-trained on a large-scale dataset of approximately 920 million photos, which can give us precise recognition results (Ren et al., 2021). In addition, We put 100 randomly selected photos from Yelp.com into the Google Vision API, Amazon Rekognition, and IBM Watson respectively, to compare the recognition precision among these three tools and our manual recognition. The results show that Google Vision API has the highest accuracy and most rich content labels for food photos and restaurant environments. This tool automatically identifies large-scale photos from consumers and business owners. Finally, we developed another Python script to upload all photos to Google Vision API, which automatically recognizes photo content and then returns labels for each photo at each restaurant.

Table 3.1 and Table 3.2 show examples of food and physical environmental photo content results after uploading Figure 6, Figure 7 (taken by the business owner) and Figure 8, Figure 9 (taken by the consumer) into the Google Vision API[8]. Then Google

[6] $https : //www.yelp.com/biz_photos/mercato - della - pescheria - las - vegas - las - vegas?select = CWrQX3o - -2u3GNJUTxBLGw$

[7] https://cloud.google.com/vision/

[8] For convenience, we only list three labels with the highest confidence scores and the three labels

Vision API will automatically return the recognition result in the form of labels (in text) with a confidence score. The label scores ranged from 0 (least confidence) to 1 (most confidence), indicating how confident the photo contains the specific labels. For instance, both photos have a confidence close to 1 to determine that they contain food. Treating a photo as having a particular entity is a common practice if its score exceeds 0.5 (Tan et al., 2019). To ensure the accuracy of photo content recognition, we only keep the labels with a confidence score over 0.6(Ren et al., 2021).

We get 2,671 unique labels after uploading all photos into Google Vision API. Then we merged two datasets by matching each dataset's business ID, rater's name, and date. The unit of observation was review by day. After data processing, the dataset contained 30,023 observations from 261 Las Vegas restaurants over 2,331 days (from 12 July 2005 to 10 December 2017). The panel dataset was unbalanced due to restaurants entering and exiting the market.

3.3.4 Photo Content Identification Results

We show the top 50 labels for customers (Figure 3.10) and business owners (Figure 3.11), respectively. The identified labels described specific customer and business owner features in their photos. The most frequent labels were Tableware, Ingredient, Food, Cuisine, and Recipe for customers and business owners, which suggests that both focus on specific restaurant or food characteristics. These outcomes were consistent with the fact that such labels were common restaurant features, likely to be captured in the photos taken inside the restaurant. Other labels describing the food types (such as Fast food, Finger Food), and detailed dining environment (like chair) were also identified.

---

with the lowest confidence scores.

3.4 Photo Content Analysis

3.4.1 LDA Modeling

The Latent Dirichlet Allocation (LDA), a generative probabilistic model for discovering latent semantic topics from a large text corpus, has become one of the most popular tools for mining big text data. LDA is used to analyze the online restaurant reviews for a variety of purposes, such as extracting and labeling the review dimensions (including food/taste, experience, location, and value) (Luo and Xu, 2019), finding the topics customers care about most(Huang et al., 2014), predicting the restaurant survival(Lian et al., 2017), detecting fake reviews(Jia et al., 2018), or selecting the best restaurant placement(Wang et al., 2016). Recent research by Jia (2021) collected 651,703 restaurant-user-generated reviews from Dianping (a Chinese version of Yelp.com) to analyze restaurant customersâ evolution of dining patterns and satisfaction during COVID-19 (from 2019-2020). By adopting the LDA, they identified twenty review topics, mostly food. The results suggest that customers have been paying fewer visits to restaurants since the outbreak, assigning lower ratings, and showing limited evidence of spending more.

We applied Latent Dirichlet Allocation (LDA) to identify unobserved topics within the data for two reasons. First, LDA discovers hidden topics in a pile of photo content labels and measures the probability of sorting each label into a topic. The photo recognition result from Google Vision API shows a couple of seemingly unrelated labels such as âfood,â âwaterâ, and âtablewareâ in a photo, which can be associated with a single group of âdiningâ when processed by a human. LDA can connect such labels and reveal topics containing similar linguistic cues (Poushneh and Rajabi, 2022). Second, LDA is one of the most popular topic modeling techniques (Blei et al., 2003) and a wide range of research fields adopt LDA as the main methodology and proves it has better performance than other topic modeling algorithms with proven probability(Ramage et al., 2010; Brzustewicz and Singh, 2021). We performed the topic modeling using the LDA implementation of the Gensim library(Rehurek and Sojka 2010).

The basic idea of LDA in our context is that photos are represented as random mixtures over latent topics, where each topic is characterized by a distribution over labels. Accordingly, LDA assumes that photos, represented as bags of label content, exhibit different topics in different proportions, as Figure 3.12 shows. Each D photo in the dataset is a mixture over K latent topics (e.g., the "Service", "Dining Environment"). Moreover, each topic is a multinomial distribution of W words vocabulary (corresponding to "labels" in photo). For example, photo 2 in Figure 3.12 is half about Topic 1 (50%) while photo 2 has a low probability about Topic 2 (1%). Each topic is represented as a probability distribution over controlled labels, usually all the labels appearing in the photo collection. In our case, Topic 1 has labels like âriceâ (8%), "pizza" (5%), which have high correlations with "Food" and Topic 3 has words like âtablesâ (8%), âdrinkwareâ (6%) with high probability with "Dining Environment". According to Blei et al. (2003), the probability of a corpus is obtained as follows:

$$p(D|\alpha, \beta) = \prod_{d=1}^{M} \int P(\theta_d|\alpha)(\prod_{n=1}^{N_d} \sum_{z_{d_n}} P(z_{d_n}|\theta_d) P(w_{d_n}|z_{d_n}, \beta)) d\theta_d \qquad (3.1)$$

where α and β are the Dirichlet parameter on topic distribution over the labels and the Dirichlet parameter on the label's distribution, respectively. θ_d is the topic distribution for photo d, z_{d_n} is the topic for the nth label in the photo d, and w_{d_n} is a specific label in photo d. N_d denotes the number of labels. The "Gensim" package in Python automatically sets α and β based on the label distribution in our bag of labels.

We pre-process the dataset after getting all photo labels. We first removed all numbers, punctuation, percentages, and brackets. Next, we removed low-frequency labels that appeared less than 1 % in all photo labels(e.g. the 'carnaroli', 'night', etc). The LDA model is a probability-based model that performs badly for those less frequently appearing patterns of photos because lower frequency means that these patterns contribute less to the target function and thus become less important to be considered by the model (Zhou et al., 2021). Removing rare labels reduces the risk that the results are also influenced by outlier labels(Netzer et al., 2019; Tirunillai and Tellis, 2014)

Determination of Optimal Number of Topics

Because LDA is an unsupervised technique, there is no prior information on the number of topics in our label set. Thus, we conducted perplexity and coherence tests to determine the optimal number of topics. On the one hand, perplexity is intuitively defined as âhow confused is the model about its decisionâ (Neubig, 2017). More accurately, it expresses the value âif we randomly picked words from the probability distribution calculated by the language model at each time step, on average, how many words would it have to pick to get the correct one?â (Neubig, 2017). It monotonically decreases the likelihood of the test data and measures how well a model describes a document according to a generative process based on the learned set of topics(Brzustewicz and Singh, 2021; Blei et al., 2003). Perplexity is inversely related to the model predictive likelihood, so a lower perplexity score indicates better generalization performance and is calculated as follows (Blei et al., 2003) :

$$perplexity(D) = exp\left\{-\frac{\sum_{d=1}^{M} logp(w_d)}{\sum_{d=1}^{M} N_d}\right\} \tag{3.2}$$

where D is the set of label sets, N_d is the number of labels in photo d, w_d is the specific label in photo d, and $p(w_d)$ is the generation of probability of w_d.

Although perplexity can help decide the optimal number of topics, sometimes it can be contrary to human judgements(Newman et al., 2010b). Moreover, models that only rely on perplexity are at risk of generating chained, unbalanced, intruded, or random topics, as non-related words may be included on the same topic(Pinto Gurdiel et al., 2021). So a metric based on the co-occurrence of words within the document is needed. A popular metric of semantic validation is the coherence score.

Coherence score is the other metric to select the optimal number of topics. It is based on the degree of semantic similarity between high-scoring words within the topic, thereby giving the human interpretable topics(Grossman and Frieder, 2004). A topic is said to be coherent when its most likely terms can be interpreted and associated with a single semantic concept(Newman et al., 2010a). For example, âcocktailâ, âwineâ, and âbeerâ are items that can be easily associated with the topic of âalcoholâ or 'drinks'. Newman et al. (2010a) utilize the Pointwise Mutual Information (PMI) to measure the probability of seeing two labels within the same topic compared to seeing them individually. The coherence score is calculated as follows:

$$PMI(w_i, w_j) = log(\frac{P(w_i, w_j)}{P(w_i)P(w_j)}) \tag{3.3}$$

Different numbers of topics will have different levels of perplexity and coherence scores. The "Gensim" library provides the perplexity and coherence score after we have set up all parameters. We choose a certain range of number of topics and decide the optimal number if the model has low perplexity and high consistency by jointly judging the results from these two metrics.

3.4.2 LDA Results

Before conducting LDA topic modeling, results from perplexity and coherence jointly help to decide the optimal number of topics. We select five as the optimal number of topics based on Figure 13 and Figure 14. Perplexity shows that we can select the number of topics from five to ten because of their lower perplexity (around 34) compared to other number of topics (over 36), which indicates models with a larger number of topics suffer more severe perplexity degradation. Perplexity exhibits a sharp decline when increasing from 4 to 5 topics, followed by a stable trend (from 5 to 10) and then increasing when selecting over 11 topics. It means a good fit of our model and the model's efficiency in predicting topics is greatly improved when selecting 5 to 10 topics.

A higher coherence score model indicates better interpretable topics and semantic coherence to humans. Figure 14 shows a higher coherence score (above 0.80) than the other number of topics (below 0.80) if we select five or seven as the number of topics. Moreover, the coherence score shows a sharp increase when changing four from five as the optimal number of topics, which means the interpretability of the model has improved substantially. However, in the LDA models with seven topics, a larger number of topics were judged to be less meaningful in discerning a unique and salient theme than the model with five topics. As a robuset check, we plot the labels of 7 topics (appendix), which shows that topic 2, topic 3, and topic 4 are difficult to interpret and some labels have overlaps.

We present the visualization of five topic distribution via the LDAvis package in Python. The LDAvis visualization result shows semantical distinct topics, which indicates the quality of topics (see Figure 3.15). The larger distance between bubbles (topics) shows more sarity. Figure 3.15 indicates no correlation between the five topics, implying that each topic is unique and informative. The bubble size shows the topic's prominence based on the frequencies of relevant words, which means Topic 3 and Topic 5 are the most prominent topics, followed by Topic 2. Table X presents the five topics and the most representative photo content labels for each topic based on the relevance score calculated using λ=0.6 (Sievert and Shirley 2014).

Next, we name each topic after setting five as the optimal number of topics. Topic 1 includes the labels relating to the restaurant's outside and inside environment, like "Room", "Chair", "Interior design", "Lighting". As a result, we name Topic 1 as "Ambience". We name Topic 2 as "Dish" because it contains various types of food such as "Comfort Food", "Staple Food", "Fast Food". Topic 3 consists of different types of desserts, such as "Sweetness", "Dessert", so we refer to Topic 3 as "Dessert and sweets". Results in Table x indicate that Topic 4 is more likely to consist of words describing the meat, such as "Steak", "Pork", "Red Meat". Consequently, we name Topic 4 as "Protein Food". Topic 5 contains labels describing drinks and drinkware, such as "Cocktail", "Alcoholic beverage", "Barware". Thus, we name Topic 5 as "Drinks & Drinkware".

3.4.3 Photo content difference

Once the topics were recognized from the LDA models for the entire dataset, we conducted quantitative analyses to explore the photo contents differences between customers and business owners, based on Ren et al. (2021). We first computed the overall fraction of each topic in the photo collection at each restaurant for business owners. Because photos posted by business owners are updated less frequently and the content is relatively fixed compared to consumers. We also computed the daily fraction of each topic for photos taken by customers at each restaurant. Finally, we get the photo content differences between them by taking the differences between those two fractions. For example, suppose there are 20 photos by distinct semantical topics, indicating restaurant i, respectively. The number of photos with a specific topic j for business owner's photos is four while the number of consumer's photos with the same topic is ten so the topic proportions are 0.2 and 0.5 and their differences is 0.3 at date t. The photo content differences can be calculated as following:

$$Diff_{ijt} = \frac{|l_{cijt}|}{|L_{ijt}|} - \frac{|l_{bij}|}{|L_{ij}|} \tag{3.4}$$

where $|L_{ij}|$ is a entire photo collection in which $L_{it} = \{l_1, l_2, ...l_m\}$ have m photos for Topic j at restaurant i; $\frac{|l_{bij}|}{|L_{ij}|}$ is defined as the proportion of number of photos posted by business owners of Topic j for restaurant i; $|L_{ijt}|$ is a list of all possible topics in the entire customer's photo collection for restaurant i of Topic j at date t; $\frac{|l_{cijt}|}{|L_{ijt}|}$ is the proportion captured in the photos of topic j by consumers for restaurant i at date t. We standardized the differences to make photo content differences comparable across topics and restaurants.

3.5 Econometric Analysis

3.5.1 Summary Statistics

Table 3.4 presents summary statistics for the panel dataset after merging photo content dataset and restaurant information by restaurant id, user name, and date. From the table we can observe (1) the average rating for restaurant is around 4.3 with one standard deviation towards 5 points, from which we can infer customers are satisfied with their dining experience. (2) the mean number of characters of review is close to 400, which indicates that the customers tend to leave text comments (around 3 or 4 paragraphs). (3) For photo content differences of five topics, the largest and smallest difference in mean value is Topic 1 (Ambiance) (0.151) and Topic 5 (Drink & Drinkware) (0.0038), respectively. Topic 2 (Dish) shows more variation (0.384) than the other topics due to its highest standard deviation. (4) the mean value of photo content differences for Topic 2, Topic 3 (Dessert & sweets), and Topic 4 (Protein Food) is positive, which shows that customers focus more on these topics.

3.5.2 OLS Fixed-effect Regression

We adopt a fixed-effect model to explore the association between differences of topics for five topics and restaurant ratings. The advantage of using the fixed-effect model is that it will mitigate bias and endogeneity issues caused by unobservable restaurant characteristics(Tian et al., 2021). For instance, factors like restaurant interior design would affect customer's perceptions of cleanliness, safety, and healthiness, which may affect the online rating and their willingness to take photos(Liu et al., 2022).

$$Rating_{it} = \beta_0 + \sum_{j=1}^{4} \beta_j * Diff_{ijt} + \alpha_5 * Length_{it} + \sigma_i + e_{ijt} \tag{3.5}$$

where $Rating_{it}$ is the rating for restaurant i at date t. $Diff_{ijt}$ is the standardized photo content differences indices between customers and business owners for restaurant i, topic j at date t[9]. $Length_{it}$ [10] is the word count of text comments for restaurant i at date t; σ_i is the the restaurant fixed effect for each restaurant i; e_{ijt} is the error term.

Table 3.5 presents four regression results for the OLS fixed-effect model. (1) The first column indicates that four topics have a statistically significant ($p < 0.01$) effect on rating, except for Topic 3 (Dessert & sweets). Topic 1 (Ambiance) Topic 2 (Dish), and Topic 4 (Protein Food) are positively correlate with rating. Moreover, the differences in Topic 2 (Dish) has greater impact on rating compared to other topics.

The second column adds the length of text comments and its quadratic term to control for customer's perceptions of the helpfulness text comments(Heng et al., 2018). We can observe that the coefficient of Topic 1 (Ambiance) turns insignificant and the coefficient of Topic 2 (Dish) and Topic 4 (Protein Food) decrease around

[9] We excluded âTopic 5 (Drink & Drinkware)â category to avoid multicollinearity because the total share of the five topics is equal to 1. Besides, it is the topic with the lowest difference between customers and business owners.

[10] It represents the amount of information and the elaborateness of the reviewer, which can positively affect review usefulness(Peng et al., 2014; Yang et al., 2017a). As a result, we add it to control for the usefulness of text comment(Heng et al., 2018)

10%. Besides, the length of review and its quadratic term are significantly correlated with rating and coefficient of Topic 3(Dessert & sweets) turns to significantly positive, which shows it is necessary to include the length as a control to the equation to reduce the omitted variable bias in estimation.

The third and fourth columns add the restaurant fixed effects. The photo content differences for Topic 2 remain statistically significant positive while Topic 3 turn from insignificant (in columns 1 and 2) to statistically significant positive ($p < 0.05$). Besides, the R^2 exhibit considerable increase from 0.003 (column 1) to 0.129 (column 4), which was the additional variation of the dependent variable explained by the restaurant fixed effects. Overall, the combination of the four topic indices explains about 12% of the variation in the restaurant ratings.

We can draw two preliminary conclusions from Table 3.5. First, the results from the fixed-effect model with all controls (column 4) show that all photo content differences for four topics are statistical significantly and positively affect restaurant rating. Topic 2 (Dish), Topic 3 (Dessert & sweets) and Topic 4 (Protein Food) are positively correlated with rating, from which we can expect that 1-standard-deviation increase in these topics will lead to a 0.359 unit, 0.134 unit, and 0.413 (roughly a half star) increase in rating, respectively. So for business owners, they can improve restaurant ratings by increasing the number of photos regarding these topics. This result is consistent with the result from the summary statistics that we found customers tend to post more photos about these topics.

Second, the coefficients both produce a significant decrease (over 10%) in column 3 and column 4, suggesting that our regression results in the first and second columns may be biased without controlling the information from text comments.

3.6 Discussion

Our results do not show the effect of the difference in photos of customers and business owners in Topic 1 (Ambiance), which generally refers to the dining environment, on restaurant ratings. This may be because the food provided (quality, taste) was central to customer's decision to favor one restaurant over another, though prior positive experience, a clean production/service environment and hospitable service are additional factors that most strongly influenced restaurant choice(Duarte Alonso et al., 2013). Therefore, customers care more about whether food instead of dining environment meet their expectation or not , thus influencing their ratings behaviors.

The regression result shows that the restaurant rating is positively correlated with the photo content differences in Topic 2 (Dish), which shows that business owners can boost rating by increasing number of photo (e.g., add more photos about the dish decoration) regarding Topic 2 (Dish) (1-standard-deviation increase in Topic 2 will lead to a 0.359 unit of increase in rating, which results in increasing restaurant rating toward 4.5). This finding supports results that the food photo's virtual content empirically influences the diner's expectations and their subsequent evaluations(Piqueras-Fiszman and Spence, 2014), which further affects restaurant rating behaviors. Moreover, customers continually adjust their expectations as they acquire fresh information about the product/service and use accumulated experience as a basis for subsequent evaluation(Bhattacherjee, 2001). It supports the conclusion that more number of photos on relevant topics, will provide more information to consumers thus influencing their rating behavior.

The regression result also implies that, given the number of customer's photos content in Topic 3(Dessert), business owners can improve restaurant ratings by increasing subjects in photos about Topic 3 (1-standard-deviation increase in Topic 3 will increase the rating by 0.13). This result supports that dessert can be a predictor of rating in text review analysis and indicates that it can influence rating in photo analysis(Jurafsky et al., 2014).

Finally, our results show that the photo content differences in protein food has the largest and positive impact on rating compared to other topics. It reflects that consumers are sensitive to protein foods, which can significantly affect their ratings of restaurants. The gap between consumers' quality expectations and perceptions of quality can be especially large when evaluating protein foods(Grunert et al., 2004). Protein foods (e.g., meat) is mostly generic or commodity-type (it is often sold unbranded and even unlabeled), and consumers do not have enough, and often conflicting, information to form their expectations(Verbeke, 2005). Furthermore, the main determinants of satisfaction may influence overall satisfaction to a different degree depending on the level of protein product processing(Resano et al., 2011). The processing of protein food affects both the taste, the quality of the food, and overall satisfaction(Resano et al., 2007), which may affect the restaurant rating. Hence, business owners or managers can consider adding more photo about protein food to provide more information and influencing customer's rating behaviors.

3.7 Conclusion and policy implication

Online platforms provide massive valuable information and restaurant rating is critical for both business owners and customers. Understanding how the latent information captured by photos provides business owners or managers with insights into customerâs perception and preference, which help them make targeted marketing strategies. However, previous studies have focused on numeric features (i.e., review length) and textural features of text review (i.e., readability) on restaurant rating, photo content potential has not been effectively utilized due to the barrier in manual analysis approach in traditional studies(Ren et al., 2021).

In this paper, we first construct the dataset by collecting the photo on Yelp.com and then we recognized the photo content automatically at large scale with Google Vision API; lastly, we build the LDA model to explore the latent topic of the photo and figure out photo content differences between business owners and its correlations with rating. We find that (1) the top three frequent photo content for business owners and customers are tableware, ingredient, food; (2) the latent topics for customer's photo are "Ambiance", "Dish", "Dessert & sweets ", "Protein Food", and "Drinks & Drinkware"; (3) "Dish" account for most proportion for both business owners and customers ; (4) the photo content differences regarding the topic of dish, dessert & sweets, and protein food were positively related to ratings (5) the discrepancy in protein food photos was found the largest and positive impact on rating.

For the managerial implications, our findings confirm the importance of pictorial-based marketing efforts in affecting customers rating behaviors by illustrating the positive effect of photo content differences between business owners and customers in different topics on rating. Our results indicate that when restaurant owners or managers promote their restaurant, they should provide more photos regarding the food than the environment. Specially, since customers are sensitive to protein food, business owners and managers should try to use pictures that realistically portray the food for promotion purposes to avoid raising customer's expectations too high.

Our study has two limitations and can be further extended in some areas. First, although we control the effect of text comment on consumers' rating behavior, we can not exclude the effect of other customers' photos on consumers' formation of expectations since customers tend to select restaurants by searching for food images by other customers(Oliveira and Casais, 2018), which may affect rating behaviors. And to control for such impact means we must know exactly which photos each consumer is referencing when viewing restaurant reviews. Given that our dataset is as of 2017, we cannot control for such an impact. Nevertheless, we used fixed-effect and the result from robust check to ensure the accuracy of our results. Since the Yelp.com have the vote indicates the helpfulness of the photo(similar to the helpfulness of text comments), future research may consider including the helpfulness of photo into their research.

Second, previous literature show that the photo attributions such as the proportion of food to the overall photo area, the dominant color of the food photo affect the perception of quality for customers(Zhang and Luo, 2022). Even so, our study is still have meaningful implications. The photo by business owner's photo may be overly retouched, but customers will refer other consumers' photos that reflect the real dining experience, thus mitigating the effect of photo attribution. Subsequent studies can get the photo attributions from Google Vision API and control them that influence restaurant ratings.

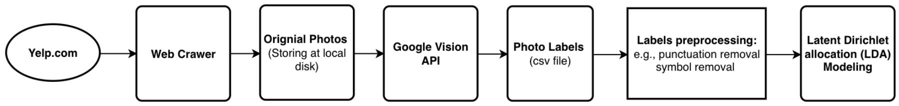


Figure 3.1: Road Map

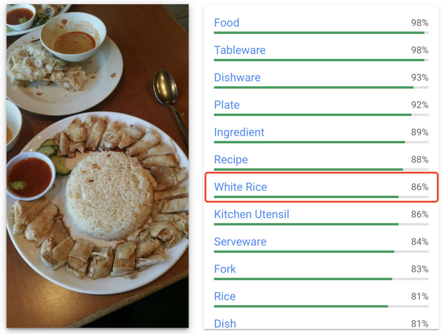


Figure 3.2: Recognition results from original photo

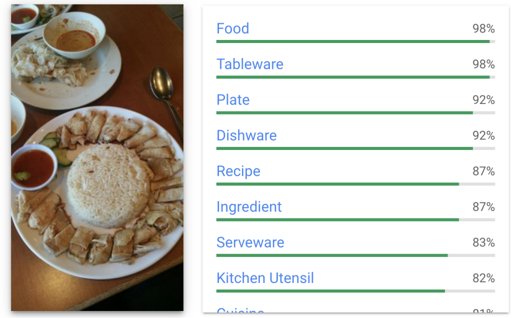


Figure 3.3: Recognition results from Yelp Dataset

Figure 3.4: Identification of Business Owners

Figure 3.5: Yelp Photo Classification

Figure 3.6: Food photo by business owner

Figure 3.7: Food photo by consumer

Figure 3.8: Environment Photo by business owner

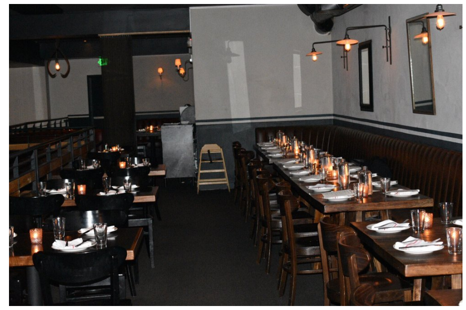


Figure 3.9: Environment photo by consumer

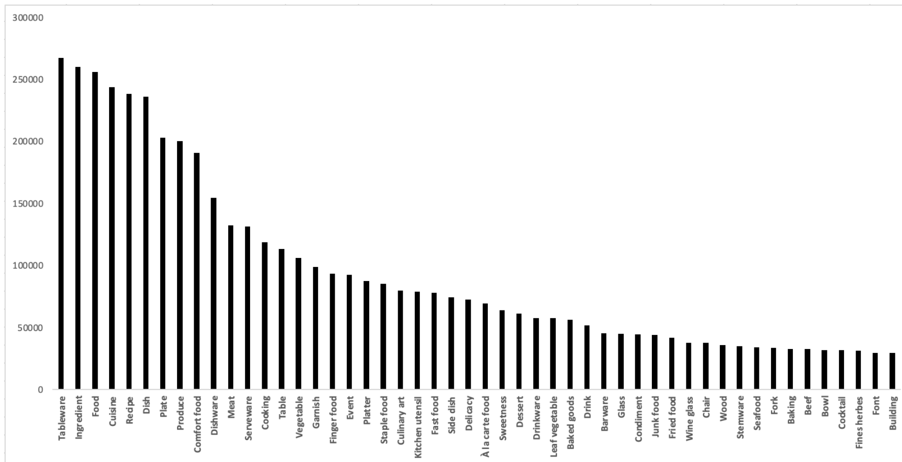


Figure 3.10: Popular labels captured from online customer photos

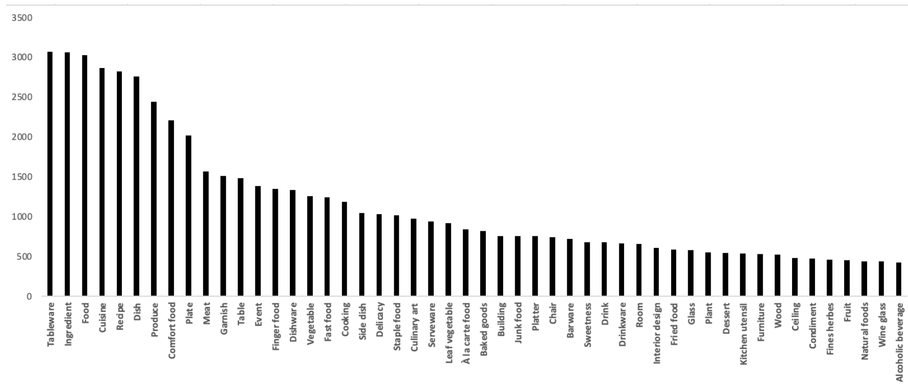


Figure 3.11: Popular labels captured from online business owners photos

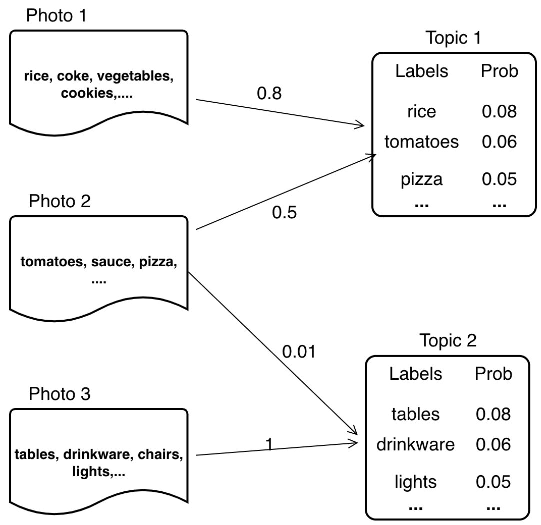


Figure 3.12: LDA examples

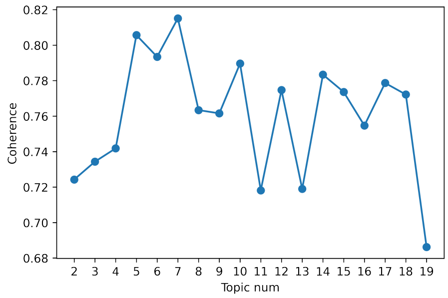


Figure 3.13: Coherence Score

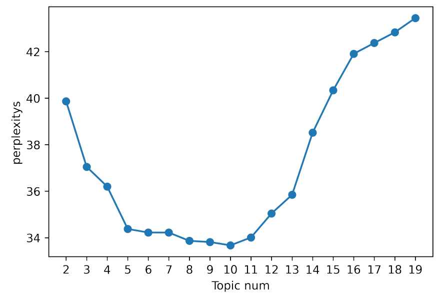


Figure 3.14: Perplexity

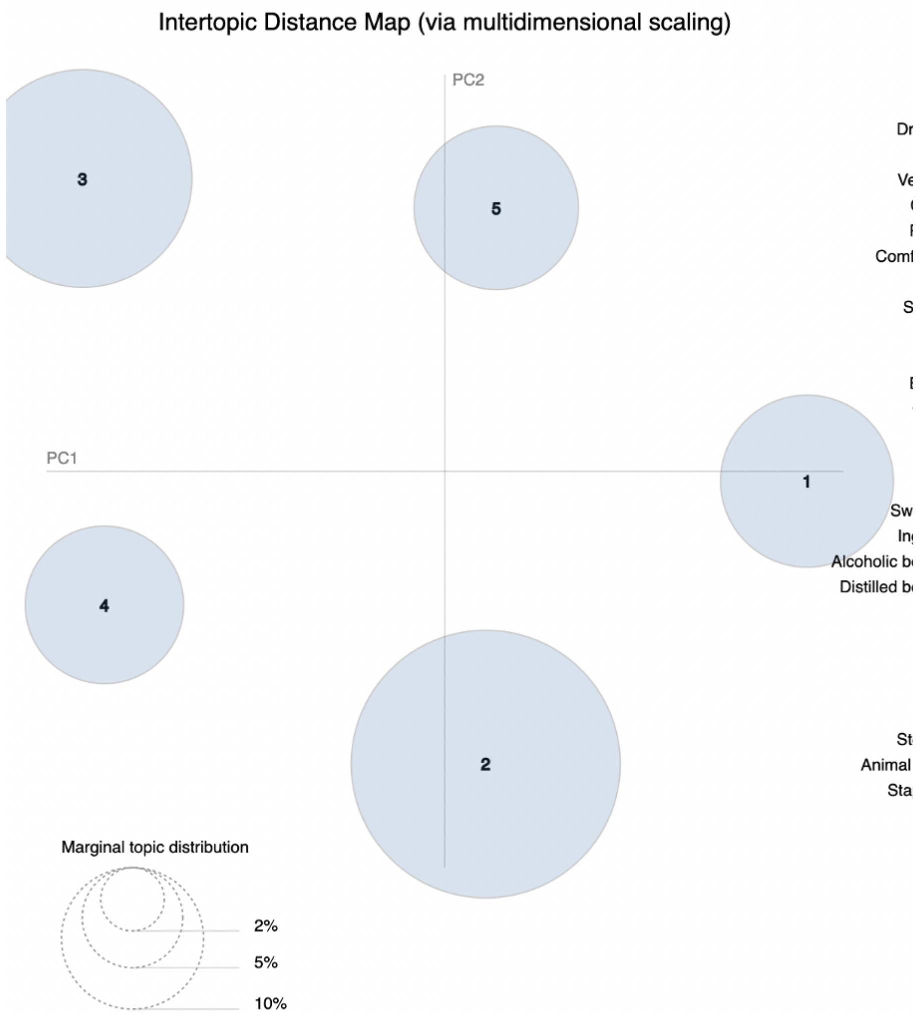


Figure 3.15: Visualization for topic distribution

Business Owner Photo	Score	Consumer Photo	Score
Food	0.974	Food	0.974
Tableware	0.971	Drinking water	0.794
Dishware	0.930	Ingredient	0.773
Still life photography	0.523	Chinese food	0.511
Tablecloth	0.517	Coquelicot	0.509
Fruit cake	0.517	Vegetarian food	0.500

Table 3.1: Food photo content recognition results

Business Owner Photo	Score	Consumer Photo	Score
Furniture	0.964	Table	0.974
Table	0.921	Furniture	0.954
Light	0.910	Property	0.943
Restaurant	0.533	Meal	0.531
Projection Screen	0.510	Plate	0.519
Chandelier	0.500	Diner	0.500

Table 3.2: Restaurant photo content recognition results

Table 3.3: Top 10 terms in each of five topics.

Topic 1: Ambiance	Topic 2: Dishes	Topic 3: Desserts and sweets	Topic 4: Protein Foods	Topic 5: Drinks & Drinkware
Event	Recipe	Sweetness	Beef	Drink
Building	Produce	Dessert	Animal product	Drinkware
Font	Food	Serveware	Steak	Cocktail
Room	Cuisine	Cuisine	Flesh	Alcoholic beverage
Chair	Ingredient	Dish	Pork	Distilled beverage
Interior design	Dish	Dishware	Meat	Barware
Wood	Comfort food	Food	Red meat	Stemware
Ceiling	Tableware	Ingredient	Venison	Glass
Flooring	Side dish	Cake	Cooking	Champagne stemware
City	Vegetable	Tableware	Veal	Non-alcoholic beverage

Table 3.4: Summary Statistics

Variables	Mean	Std. Dev	Min	Max	Observarion
		Yelp Review			
Rating	4.279	0.928	1	5	30,017
Length of Review (# of Characters)	398.2	871.1	1	6,653	30,017
		Photo Content Differences Index			
Topic 1 (Ambience)	-0.151	0.367	-1	1	30,017
Topic 2 (Dish)	0.0660	0.384	-1	1	30,017
Topic 3 (Desserts & sweets)	0.0435	0.289	-1	1	30,017
Topic 4 (Protein Foods)	0.0456	0.208	-1	0.667	30,017
Topic 5 (Drinks & Drinkware)	-0.00381	0.202	-1	0.667	30,017

Table 3.5: OLS Fixed-effect Regression Results

Variables	(1) Rating	(2) Rating	(3) Rating	(4) Rating
Std. Photo Content Differences (Topic 1)	0.160***	0.0847	0.0755	-0.0475
	(0.0581)	(0.0576)	(0.0673)	(0.0669)
Std. Photo Content Differences (Topic 2)	0.355***	0.304***	0.419***	0.359***
	(0.0547)	(0.0541)	(0.0632)	(0.0626)
Std. Photo Content Differences (Topic 3)	0.0298	-0.0337	0.168**	0.134**
	(0.0612)	(0.0607)	(0.0688)	(0.0682)
Std. Photo Content Differences (Topic 4)	0.187***	0.155***	0.433***	0.413***
	(0.0590)	(0.0584)	(0.0652)	(0.0646)
Review Length		-0.000540***		-0.000523***
		(2.12e-05)		(2.12e-05)
$ReviewLength^2$		1.29e-07***		1.25e-07***
		(5.41e-09)		(5.46e-09)
Constant	3.900***	4.112***	3.711***	3.931***
	(0.108)	(0.107)	(0.120)	(0.120)
Observations	30,017	30,017	30,017	30,017
R-squared	0.003	0.024	0.111	0.129
Restaurant FE	No	No	YES	YES

References

Abbott, P. C., Hurt, C., and Tyner, W. E. (2008). What's driving food prices? Technical report.

Aderajew, T. S., Du, X., Pennings, J. M., and Trujillo-Barrera, A. (2020). Farm-level risk-balancing behavior and the role of latent heterogeneity. *Journal of Agricultural and Resource Economics*, 45(2):265–281.

Alcorn, C. (2021). Global food prices surge to their highest level in a decade. `https://www.cnn.com/2021/06/04/business/inflation-food-prices/index.html`.

AliResearch (2020). 2020 taobao live streaming economicv. `https://www.cbndata.com/report/2219/detail?isReading=report&page=1`.

Amblee, N. and Bui, T. (2008). Can brand reputation improve the odds of being reviewed on-line? *International Journal of Electronic Commerce*, 12(3):11–28.

Anderson, M. and Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, 122(563):957–989.

Aula, P. (2010). Social media, reputation risk and ambient publicity management. *Strategy & leadership*.

Baba, Y., Engle, R. F., Kraft, D. F., and Kroner, K. F. (1990). Multivariate simultaneous generalized arch. *Manuscript, University of California, San Diego, Department of Economics*.

Baek, J. and Koo, W. W. (2010). Analyzing factors affecting us food price inflation. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 58(3):303–320.

Bekkers, E., Brockmeier, M., Francois, J., and Yang, F. (2017). Local food prices and international price transmission. *World Development*, 96:216–230.

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS quarterly*, pages 351–370.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Brzustewicz, P. and Singh, A. (2021). Sustainable consumption in consumer behavior in the time of covid-19: Topic modeling on twitter data using lda. *Energies*, 14(18):5787.

Burton, S., Howlett, E., and Tangari, A. H. (2009). Food for thought: how will the nutrition labeling of quick service restaurant menu items influence consumersâ product evaluations, purchase intentions, and choices? *Journal of Retailing*, 85(3):258–273.

Cai, R. and Chi, C. G.-Q. (2021). Pictures vs. reality: roles of disconfirmation magnitude, disconfirmation sensitivity, and branding. *International Journal of Hospitality Management*, 98:103040.

Chen, C.-D., Zhao, Q., and Wang, J.-L. (2020). How livestreaming increases product sales: role of trust transfer and elaboration likelihood model. *Behaviour & Information Technology*, pages 1–16.

Chen, P.-Y., Chang, C.-L., Chen, C.-C., and McAleer, M. (2012). Modelling the effects of oil prices on global fertilizer prices and volatility. *Journal of Risk and Financial Management*, 5(1):78–114.

Chowdhury, M. A. F., Meo, M. S., Uddin, A., and Haque, M. M. (2021). Asymmetric effect of energy price on commodity price: New evidence from nardl and time frequency wavelet approaches. *Energy*, 231:120934.

Clement Addo, P., Fang, J., Asare, A. O., and Kulbo, N. B. (2021). Customer engagement and purchase intention in live-streaming digital marketing platforms. *The Service Industries Journal*, pages 1–20.

De Gorter, H. and Just, D. R. (2008). " water" in the us ethanol tax credit and mandate: Implications for rectangular deadweight costs and the corn-oil price relationship. *Review of Agricultural Economics*, 30(3):397–410.

Dillon, B. M. and Barrett, C. B. (2016). Global oil prices and local food prices: Evidence from east africa. *American Journal of Agricultural Economics*, 98(1):154–171.

Du, X., Khanna, M., Lu, L., Yang, X., and Zilberman, D. (2017). Contracting in the biofuel sector. In *Handbook of Bioenergy Economics and Policy: Volume II*, pages 401–425. Springer.

Duarte Alonso, A., O'neill, M., Liu, Y., and O'shea, M. (2013). Factors driving consumer restaurant choice: An exploratory study from the southeastern united states. *Journal of Hospitality Marketing & Management*, 22(5):547–567.

Dutta, A., Bouri, E., Saeed, T., and Vo, X. V. (2021). Crude oil volatility and the biodiesel feedstock market in malaysia during the 2014 oil price decline and the covid-19 outbreak. *Fuel*, 292:120221.

EIA (2021). Energy information administration (eia) forecasts u.s. winter natural gas bills will be 30% higher than last winter. `https://www.eia.gov/todayinenergy/detail.php?id=50076`.

Eissa, M. A. and Al Refai, H. (2019). Modelling the symmetric and asymmetric relationships between oil prices and those of corn, barley, and rapeseed oil. *Resources Policy*, 64:101511.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric theory*, 11(1):122–150.

Etienne, X. L., Trujillo-Barrera, A., and Wiggins, S. (2016). Price and volatility transmissions between natural gas, fertilizer, and corn markets. *Agricultural Finance Review*.

Fafchamps, M., Hill, R. V., and Minten, B. (2008). Quality control in nonstaple food markets: evidence from india. *Agricultural Economics*, 38(3):251–266.

Fernandez-Perez, A., Frijns, B., and Tourani-Rad, A. (2016). Contemporaneous interactions among fuel, biofuel and agricultural commodities. *Energy Economics*, 58:1–10.

Ferreira, P., Almeida, D., Dionísio, A., Bouri, E., and Quintino, D. (2022). Energy markets–who are the influencers? *Energy*, 239:121962.

Fowowe, B. (2016). Do oil prices drive agricultural commodity prices? evidence from south africa. *Energy*, 104:149–157.

Frick, B. and Simmons, R. (2013). The impact of individual and collective reputation on wine prices: empirical evidence from the mosel valley. *Journal of Business Economics*, 83(2):101–119.

Gan, Q., Ferns, B. H., Yu, Y., and Jin, L. (2017). A text mining and multidimensional sentiment analysis of online restaurant reviews. *Journal of Quality Assurance in Hospitality & Tourism*, 18(4):465–492.

Gardebroek, C. and Hernandez, M. A. (2013). Do energy prices stimulate food price volatility? examining volatility transmission between us oil, ethanol and corn markets. *Energy economics*, 40:119–129.

Gerten, J. and Topolinski, S. (2019). Shades of surprise: Assessing surprise as a function of degree of deviance and expectation constraints. *Cognition*, 192:103986.

Gilbert, C. L. (2010). How to understand high food prices. *Journal of agricultural economics*, 61(2):398–425.

Grossman, D. A. and Frieder, O. (2004). *Information retrieval: Algorithms and heuristics*, volume 15. Springer Science & Business Media.

Grunert, K. G., Bredahl, L., and Brunsø, K. (2004). Consumer perception of meat quality and implications for product development in the meat sectorâa review. *Meat science*, 66(2):259–272.

Ha, J., Park, K., and Park, J. (2016). Which restaurant should i choose? herd behavior in the restaurant industry. *Journal of foodservice business research*, 19(4):396–412.

Hatzenbuehler, P. L., Du, X., and Painter, K. (2021). Price transmission with sparse market information: The case of united states chickpeas. *Agribusiness*, 37(3):665–682.

Heng, Y., Gao, Z., Jiang, Y., and Chen, X. (2018). Exploring hidden factors behind online food shopping from amazon reviews: A topic mining approach. *Journal of Retailing and Consumer Services*, 42:161–168.

Hlee, S., Lee, J., Yang, S.-B., and Koo, C. (2019). The moderating effect of restaurant type on hedonic versus utilitarian review evaluations. *International Journal of Hospitality Management*, 77:195–206.

Hu, L., Sun, A., and Liu, Y. (2014). Your neighbors affect your ratings: on geographical neighborhood influence to rating prediction. In *Proceedings of the 37th*

international ACM SIGIR conference on Research & development in information retrieval, pages 345–354.

Hu, M. and Chaudhry, S. S. (2020). Enhancing consumer engagement in e-commerce live streaming via relational bonds. *Internet Research*.

Huang, J., Rogers, S., and Joo, E. (2014). Improving restaurants by extracting subtopics from yelp reviews. *iConference 2014 (Social Media Expo)*.

Ibrahim, M. H. (2015). Oil and food prices in malaysia: a nonlinear ardl analysis. *Agricultural and Food Economics*, 3(1):1–14.

IMF (2021). International monetary fund primary commodity prices. `https://data.imf.org/?sk=471DDDF8-D8A7-499A-81BA-5B332C01F8B9&sId=1547558078595`.

Ismail, A., Ihsan, H., Khan, S. A., and Jabeen, M. (2017). Price volatility of food and agricultural commodities: A case study of pakistan. *Journal of Economic Cooperation & Development*, 38(3):77–120.

Jacks, D. S., O'rourke, K. H., and Williamson, J. G. (2011). Commodity price volatility and world market integration since 1700. *Review of Economics and Statistics*, 93(3):800–813.

Ji, Q., Bouri, E., Roubaud, D., and Shahzad, S. J. H. (2018). Risk spillover between energy and agricultural commodity markets: A dependence-switching covar-copula model. *Energy Economics*, 75:14–27.

Jia, S., Zhang, X., Wang, X., and Liu, Y. (2018). Fake reviews detection based on lda. In *2018 4th International Conference on Information Management (ICIM)*, pages 280–283. IEEE.

Jia, S. S. (2021). Analyzing restaurant customersâ evolution of dining patterns

and satisfaction during covid-19 for sustainable business insights. *Sustainability*, 13(9):4981.

Jurafsky, D., Chahuneau, V., Routledge, B. R., and Smith, N. A. (2014). Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*.

Kalkuhl, M., von Braun, J., and Torero, M. (2016). Volatile and extreme food prices, food security, and policy: an overview. *Food price volatility and its implications for food security and policy*, pages 3–31.

Kang, K., Lu, J., Guo, L., and Li, W. (2021). The dynamic effect of interactivity on customer engagement behavior through tie strength: Evidence from live streaming commerce platforms. *International Journal of Information Management*, 56:102251.

Klostermann, J., Plumeyer, A., Böger, D., and Decker, R. (2018). Extracting brand information from social networks: Integrating image, text, and social tagging data. *International Journal of Research in Marketing*, 35(4):538–556.

Koirala, K. H., Mishra, A. K., D'Antoni, J. M., and Mehlhorn, J. E. (2015). Energy prices and agricultural commodity prices: Testing correlation using copulas method. *Energy*, 81:430–436.

Li, Y. and Peng, Y. (2021). What drives gift-giving intention in live streaming? the perspectives of emotional attachment and flow experience. *International Journal of Human–Computer Interaction*, pages 1–13.

Lian, J., Zhang, F., Xie, X., and Sun, G. (2017). Restaurant survival analysis with heterogeneous information. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 993–1002.

Lin, L., Gang, R., Hong, T., and Yang, S. B. (2019). Exploring simultaneous presentation in online restaurant reviews: An analysis of textual and visual content. *Asia Pacific Journal of Information Systems*, 29(2):181–202.

Lin, Z., Li, D., Janamanchi, B., and Huang, W. (2006). Reputation distribution and consumer-to-consumer online auction market structure: an exploratory study. *Decision Support Systems*, 41(2):435–448.

Liu, H., Feng, S., and Hu, X. S. (2022). Process vs. outcome: Effects of food photo types in online restaurant reviews on consumersâ purchase intention. *International journal of hospitality management*, 102:103179.

Lu, L., Nguyen, R., Rahman, M. M., and Winfree, J. (2021). Demand shocks and supply chain resilience: An agent based modelling approach and application to the potato supply chain. Technical report, National Bureau of Economic Research.

Lu, L., Tian, G., and Hatzenbuehler, P. (2022). How agricultural economists are using big data: a review. *China Agricultural Economic Review*.

Lu, L. and Winfree, J. (2021). Demand shocks and supply chain flexibility. In *NBER Conference âRisks in Agricultural Supply Chainsâ*. Springer Berlin/Heidelberg.

Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp. com. *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper*, (12-016).

Lucotte, Y. (2016). Co-movements between crude oil and food prices: A post-commodity boom perspective. *Economics Letters*, 147:142–147.

Luo, Y. and Xu, X. (2019). Predicting the helpfulness of online restaurant reviews using different machine learning algorithms: A case study of yelp. *Sustainability*, 11(19):5254.

Meyer, D. F., Sanusi, K. A., and Hassan, A. (2018). Analysis of the asymmetric impacts of oil prices on food prices in oil-exporting, developing countries. *Journal of International Studies*, 11(3).

Miller, R. K. (2021). The 2013 restaurant, food beverage market research handbook.

Netzer, O., Lemaire, A., and Herzenstein, M. (2019). When words sweat: Identifying signals for loan default in the text of loan applications. *Journal of Marketing Research*, 56(6):960–980.

Neubig, G. (2017). Neural machine translation and sequence-to-sequence models: A tutorial. *arXiv preprint arXiv:1703.01619*.

Newman, D., Lau, J. H., Grieser, K., and Baldwin, T. (2010a). Automatic evaluation of topic coherence. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pages 100–108.

Newman, D., Noh, Y., Talley, E., Karimi, S., and Baldwin, T. (2010b). Evaluating topic models for digital libraries. In *Proceedings of the 10th annual joint conference on Digital libraries*, pages 215–224.

Nosko, C. and Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment. Technical report, National Bureau of Economic Research.

Nwoko, I. C., Aye, G. C., and Asogwa, B. C. (2016). Oil price and food price volatility dynamics: The case of nigeria. *Cogent Food & Agriculture*, 2(1).

Nyga-Łukaszewska, H. and Aruga, K. (2020). Energy prices and covid-immunity: The case of crude oil and natural gas prices in the us and japan. *Energies*, 13(23):6300.

Oliveira, B. and Casais, B. (2018). The importance of user-generated photos in restaurant selection. *Journal of Hospitality and Tourism Technology*.

Ostrom, E. (2010). Analyzing collective action. *Agricultural economics*, 41:155–166.

Peng, C.-H., Yin, D., Wei, C.-P., and Zhang, H. (2014). How and when review length and emotional intensity influence review helpfulness: Empirical evidence from epinions. com.

Phua, J. and Ahn, S. J. (2016). Explicating the âlikeâon facebook brand pages: The effect of intensity of facebook use, number of overall âlikesâ, and number of friends'âlikesâ on consumers' brand outcomes. *Journal of Marketing Communications*, 22(5):544–559.

Pinto Gurdiel, L., Morales Mediano, J., and Cifuentes Quintero, J. A. (2021). A comparison study between coherence and perplexity for determining the number of topics in practitioners interviews analysis.

Piqueras-Fiszman, B. and Spence, C. (2014). *The perfect meal: the multisensory science of food and dining*. John Wiley & Sons.

Pizam, A. and Milman, A. (1993). Predicting satisfaction among first time visitors to a destination by using the expectancy disconfirmation theory. *International Journal of Hospitality Management*, 12(2):197–209.

Poushneh, A. and Rajabi, R. (2022). Can reviews predict reviewersâ numerical ratings? the underlying mechanisms of customersâ decisions to rate products using latent dirichlet allocation (lda). *Journal of Consumer Marketing*.

Radmehr, R. and Rastegari Henneberry, S. (2020). Energy price policies and food prices: Empirical evidence from iran. *Energies*, 13(15):4031.

Ramage, D., Dumais, S., and Liebling, D. (2010). Characterizing microblogs with topic models. In *Fourth international AAAI conference on weblogs and social media*.

Reardon, T., Belton, B., Liverpool-Tasie, L. S. O., Lu, L., Nuthalapati, C. S., Tasie, O., and Zilberman, D. (2021). E-commerce's fast-tracking diffusion and adaptation in developing countries. *Applied Economic Perspectives and Policy*, 43(4):1243–1259.

Ren, M., Vu, H. Q., Li, G., and Law, R. (2021). Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: implications for hospitality marketers. *Journal of Hospitality Marketing & Management*, 30(1):96–119.

Resano, H., Perez-Cueto, F. J., de Barcellos, M. D., Veflen-Olsen, N., Grunert, K. G., and Verbeke, W. (2011). Consumer satisfaction with pork meat and derived products in five european countries. *Appetite*, 56(1):167–170.

Resano, H., Sanjuán, A. I., and Albisu, L. M. (2007). Consumersâ acceptability of cured ham in spain and the influence of information. *Food Quality and Preference*, 18(8):1064–1076.

Rezitis, A. N. (2015). The relationship between agricultural commodity prices, crude oil prices and us dollar exchange rates: A panel var approach and causality analysis. *International Review of Applied Economics*, 29(3):403–434.

Roman, M., Górecka, A., and Domagała, J. (2020). The linkages between crude oil and food prices. *Energies*, 13(24):6545.

Saghaian, S., Nemati, M., Walters, C., and Chen, B. (2018). Asymmetric price volatility transmission between us biofuel, corn, and oil markets. *Journal of Agricultural and Resource Economics*, 43(1):46.

Sanyal, P., Malczynski, L. A., and Kaplan, P. (2015). Impact of energy price variability on global fertilizer price: application of alternative volatility models. *Sustainable Agriculture Research*, 4(526-2016-37964).

Seo, B. (2007). Asymptotic distribution of the cointegrating vector estimator in error correction models with conditional heteroskedasticity. *Journal of Econometrics*, 137(1):68–111.

Seo, Y., Kim, J., Choi, Y. K., and Li, X. (2019). In âlikesâ we trust: likes, disclosures and firm-serving motives on social media. *European Journal of Marketing*.

Serra, T., Zilberman, D., and Gil, J. (2011). Price volatility in ethanol markets. *European review of agricultural economics*, 38(2):259–280.

Shahnoushi, N., Sayed, S., Hezareh, R., and Tirgari Seraji, M. (2017). Investigation of relationship between world food prices and energy price: A panel sur approach. Technical report.

Siami-Namini, S. et al. (2019). Volatility transmission among oil price, exchange rate and agricultural commodities prices. *Applied Economics and Finance*, 6(4):41–61.

Su, C. W., Wang, X.-Q., Tao, R., and Oana-Ramona, L. (2019). Do oil prices drive agricultural commodity prices? further evidence in a global bio-energy context. *Energy*, 172:691–701.

Taghizadeh-Hesary, F., Rasoulinezhad, E., and Yoshino, N. (2018). Volatility linkages between energy and food prices: Case of selected asian countries.

Taghizadeh-Hesary, F., Rasoulinezhad, E., and Yoshino, N. (2019). Energy and food security: Linkages through price volatility. *Energy Policy*, 128:796–806.

Taobao (2013). Tmall seller's naming rules. `https://rule.tmall.com/tdetail-37.htm?spm=a2177.7731969.0.0.1168c32fOVuWXN&tag=self`.

Taobao (2019). Tmall 2020 annual software service fee, discounts and settlement standards. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Taobao (2020a). 2020 tmall sellers evaluation standards. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Taobao (2020b). Tmall seller's deposit rules. `https://rule.tmall.com/tdetail-11001469.htm?spm=service_hall.20224364.0.0.10933f35Z9kr4I&tag=self`.

Tian, G., Lu, L., and McIntosh, C. (2021). What factors affect consumersâ dining sentiments and their ratings: Evidence from restaurant online review data. *Food Quality and Preference*, 88:104060.

Tirole, J. (1993). A theory of collective reputations. *Research Papers in Economics University of Stockholm*, 9.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies*, 63(1):1–22.

Tirunillai, S. and Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of marketing research*, 51(4):463–479.

Tzimiropoulos, G. and Pantic, M. (2017). Fast algorithms for fitting active appearance models to unconstrained images. *International journal of computer vision*, 122(1):17–33.

Verbeke, W. (2005). Agriculture and the food industry in the information age. *European review of agricultural economics*, 32(3):347–368.

Wang, F., Chen, L., and Pan, W. (2016). Where to place your next restaurant? optimal restaurant placement via leveraging user-generated reviews. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 2371–2376.

Winfree, J. A. and McCluskey, J. J. (2005). Collective reputation and quality. *american Journal of agricultural Economics*, 87(1):206–213.

Wongkitrungrueng, A. and Assarut, N. (2018). The role of live streaming in building consumer trust and engagement with social commerce sellers. *Journal of Business Research*.

Wongkitrungrueng, A., Dehouche, N., and Assarut, N. (2020). Live streaming commerce from the sellersâ perspective: implications for online relationship marketing. *Journal of Marketing Management*, 36(5-6):488–518.

Wongpiyabovorn, O. (2021). Changes in the relationship between nitrogen fertilizer and natural gas prices in the us.

WorldBank (2020). Fertilizer prices to rise moderately in 2021. `https://blogs.worldbank.org/opendata/fertilizer-prices-rise-moderately-2021`.

Xiao, L., Fu, B., and Liu, W. (2018). Understanding consumer repurchase intention on o2o platforms: an integrated model of network externalities and trust transfer theory. *Service Business*, 12(4):731–756.

Xu, X., Wu, J.-H., and Li, Q. (2020). What drives consumer shopping behavior in live streaming commerce? *Journal of Electronic Commerce Research*, 21(3):144–167.

Yang, S.-B., Hlee, S., Lee, J., and Koo, C. (2017a). An empirical examination of online restaurant reviews on yelp. com: A dual coding theory perspective. *International Journal of Contemporary Hospitality Management*.

Yang, S.-B., Shin, S.-H., Joun, Y., and Koo, C. (2017b). Exploring the comparative importance of online hotel reviewsâ heuristic attributes in review helpfulness: a conjoint analysis approach. *Journal of Travel & Tourism Marketing*, 34(7):963–985.

Zhang, M. and Luo, L. (2022). Can consumer-posted photos serve as a leading indicator of restaurant survival? evidence from yelp. *Management Science*.

Zhou, S., Kan, P., Huang, Q., and Silbernagel, J. (2021). A guided latent dirichlet allocation approach to investigate real-time latent topics of twitter data during hurricane laura. *Journal of Information Science*, page 01655515211007724.