Meal Kit Preferences during COVID-19 Pandemic:

Exploring User-Generated Content with Natural Language Processing Techniques

Danyi Qi[1]*, Yuetong Xu[2], Jane Chen[3], Ran Li[4]

1 Department of Agricultural Economics and Agribusiness and the LSU AgCenter, Louisiana State University

2 Department of Mathematics, University of California, Los Angeles

3 Homesite Insurance, Boston

4 Department of Agricultural, Environmental and Development Economics, Ohio State University

Abstract

We analyzed 51,497 customer reviews for nine meal kit companies in the United States in 2019 – 2020 using Natural Language Processing and Latent Dirichlet Allocation (LDA)-based topic model and derived four topics that customers discussed in the comments, including “Experience,” “Food Quality”, “Convenience” and “Service”. We compared the prevalence of the four identified topics in the customers comments before and after the pandemic between conventional and affordable meal kit companies using a difference-in-difference model. We found that conventional meal kit customers, who more likely had higher income than those subscribed to affordable meal kits, significantly valued the service provided by the meal kit companies, including delivery and customer service, more after the outbreak of the pandemic. As trade-off, those customers placed less emphasis on the quality of the products including the freshness and the type of food served and the food and family experience that they had from the meal kit. We discussed the implications of these results on the impacts of COVID-19 pandemic on U.S consumers’ food values.

Keyword: Meal kit, Machine learning, User-Generated Content, Natural Language Processing, Difference in Difference

*corresponding author: dqi@agcenter.lsu.edu

Acknowledgement:

Danyi Qi acknowledge funding support from 2020-2021 Louisiana Board of Regents Support Fund (AWD-002915)

Introduction

The recent Covid-19 pandemic has induced substantial disruption to the food supply chain and greatly shifted household food behaviors. Besides health care, the food sector is one of the most affected industries (Nicola et al., 2020). Before the pandemic, approximately 54% of food is consumed at eating establishments away from home (Grashuis et al., 2020), which mostly were temporarily shut down since the outbreak of COVID-19. According to a survey in July 2020, more than half of U.S. shoppers (55%) were eating at home more often since the pandemic began, while about a quarter of those people complaining about being exhausted by food preparation and cooking at home (Acosta, 2020). Corresponding to the abrupt decline of expenditure in food service, the sale for eGrocery skyrocketed. On the other hand, a growing number of consumers shifted to online grocery shopping due to the concerns of COVID-19 (Mercatus).

Meal kit service is a subscription food service that deliveries fresh pre-portioned food and ingredients directly to customers' homes so that specific meals can be more conveniently prepared at home (Troy & Acosta, 2018). The meal kit industry, starting in Europe, entered the U.S. market around 2012 with companies Blue Apron and Hello Fresh. Initially, the meal kit companies attracted customers through the convenience of obtaining and cooking food, offered recipes, and the claim of reducing food waste (Judkis, 2017 and SUBTA, 2020). The development of this novel service caused disruptions in the food retail industry as it offers healthy food options to households while saving time for families by reducing grocery trips to stores (Troy & Acosta, 2018). Though with many favorable innovations, meal kit industries were faced with great challenges from an unstable and low customer base before the pandemic. According to a report from NPD, an estimated only 4% of U.S. consumers had ever used meal kit services before 2019 (Durbin, 2019), which was in part a consequence of its high pricing and unsatisfied expectation of convenience. Responding to those complaints, new entrance, marketed as affordable meal kits and distinct from the conventional meal kit companies (e.g., Hello Fresh), rapidly attracted attention by delivering almost pre-cooked meal kits at much lower prices (e.g., Every Plate and Dinnerly at $4.99/person/meal). Unlike conventional meal kits that targeted busy and higher-income customers who value fresh and healthy eating diets, affordable meal kits targeted busy and lower-income customers who desire to obtain enough calories to maintain body function in a timely and financially efficient manner.

The outbreak of the COVID-19 pandemic featured with restaurant shut down and worldwide stay-at-home order provided unique opportunities to the meal kit industry. As fresh food became less accessible and dining at restaurants exposed risks, meal kit services that deliver somewhat prepared fresh groceries directly to home were identified as an alternative to grocery shopping and take-out from restaurants to lift the burden for grocery shopping and food preparation. Consequently, this industry witnessed incredible success in 2020. For example, monthly meal-kit sales in the U.S. doubled in the first month of the pandemic (Leon, 2020). Despite the rapid growth, little is known about consumers' preferences toward this novel product and its dynamic relationship with the COVID-19 pandemic progression. Chang and Meyerhoefer (2020) testified the shift to online food shopping as consumer responses to the spread of COVID-19 cases. Several recent studies in the U.S. investigated the changes in food shoppers' preferences during the pandemic based on online surveys (Ellison et al., 2020; Grashuis et al., 2020). None of these focused on the meal kits specifically.

We investigated the changes in consumer preferences to meal kits since the COVID-19 pandemic started by employing Natural Language Processing (NLP) based on customer review comments. To understand the changes in consumer focuses during the pandemic, we compared customer review ratings and the topics discussed in review comments among nine meal kit companies before and after the United States announced a national emergency on March 13th (i.e., 1/4/2019 – 12/22/2020). Among the nine meal kit companies, two were affordable meal kits while the rest were conventional meal kit companies. We scrapped 51,497 customer review ratings and comments from Trustpilot.com and processed the review comments using NLP techniques and topic modeling. As an important field in machine learning, NLP techniques exploit rich information in word documents by transforming text content to quantitative information (Gentzkow et al., 2019). We then trained a Latent Dirichlet Allocation (LDA)-based topic model (Blei et al. 2003; Heng et al. 2018; Yin and Chen 2020), using all customer review texts, which yields four key topics underlying each customer reviews. LDA is a commonly used NLP technique and is well suited to understanding our large sample of the customer reviews since it is able to analyze the topical content of a large number of lengthy documents in an objective and replicable way and relies only on very limited assumptions about the text information (Blei, 2012; Blei et al., 2003; Dyer et al., 2017). Compared with the traditional survey methods, LDA obtains quantitative details on customer preferences directly from customer review texts rather than deriving insights indirectly from customers' responses to hypothetical scenarios (e.g., discrete choice experience) (Jelodar et al., 2019). Therefore, LDA has advantages in deriving novel consumer insights that researchers failed to generate hypotheses on and in providing estimations that suffer less from hypothetical bias and/or social desirability bias.

The four topics generated from our LDA model, from the highest to lowest in prevalence, include food experience (e.g., with keywords like cook, recipe, family), food quality (e.g., meat, quality, option), convenience (e.g., easy, instruction, follow), and services (e.g., service, delivery, receive). To investigate the heterogeneous impacts of the COVID-19 pandemic on the low versus high-income consumers, we split our sample into treatment (e.g., Hello Fresh priced $7-$10/meal) and control groups (affordable meal kit e.g., Dinnerly priced at $4.99/meal) and evaluated the response differences between the two groups using the Difference-in-Difference (DID) method on customer attention focus and review ratings separately, where year and month fixed effects, the progression of COVID-19 (e.g., current and lagged confirmed cases), and customers' characteristics (e.g., if invited by the platform) are controlled.

Our preliminary results find "Convenience" as the most prevalent topic and its prevalence grew significantly during the pandemic when customers demand alleviation from in-person grocery shopping and home cooking (Acosta 2020). "Service", on the other side, was the least prevalent topic, and its occurrence was associated with lower ratings before the pandemic. During the pandemic, the discussion around this topic increased significantly among expensive meal kit subscribers who started to appreciate the contactless feature of the meal kits, and the occurrence of this topic is now strongly associated with higher ratings. Unfortunately, such changes were not observed among the comments in the control group, which may suggest a disadvantaged state of the lower-income customers (e.g., less alert and/or less capable of reducing contacts with people). "Food quality", including the discussion about the healthiness and diversity of the food, happened to have a similar prevalence between the two groups before the pandemic. As the expensive meal kits customers started to appreciate the logistics feature of the products, "Food quality" is less frequently mentioned, which reflects a tradeoff that customers make during the pandemic. Before the pandemic, the customers in the treatment group placed a higher emphasis on the "*Experience*" that meal kits introduced to their family, e.g., the cooking time with family and new food experience from novel recipes. A great "*Experience*" described in a comment was usually associated with higher ratings in the treatment group. Like "*Food quality*", the prevalence of the discussion related to "*Experience*" also decreased since the pandemic outbreak. Our results are robust regardless of how we define the starting point of the pandemic (as the start of either the first wave or the second wave) and whether we control for the progression or the severity of COVID-19.

This is the first study investigating the latent factors that affect consumer preferences for meal kits. By comparing the customer review comments between the conventional versus affordable meal kits before and during the pandemic, we study the dynamic relationship between the COVID-19 pandemic and customer preferences for meal kits. This study also provides initial evidence and implications for heterogeneous impacts of the COVID-19 pandemic on food behavior and food value among the low versus high-income consumers. Further, using an example of meal kits, we proposed a useful approach of conducting causal analysis by leveraging advanced machine learning algorithms in text mining to help researchers exploit online user-generated content and understand consumers' attitudes and preferences to food products.

The remainder of the article proceeds as follows: first, we introduced the data sources for this study. Then, we described the LDA modeling that we employed to derive quantitative information from customer comments. Later on, we explained the identification strategies used to estimate the impacts of the COVID-19 pandemic on consumers' preference changes. We then conclude and discuss limitations and future work.

Data

New York Times COVID-19 data

We used daily COVID-19 confirmed cases and death cases from the New York Times, an open data source collected from government and health department available through Github, to quantify the progression of the pandemic. The data starts from January 21st, 2020, the day after the first COVID-19 case in the United States. As of December 22nd, 2020, the last date in our study, we collected a total of 337 entries for cumulative counts of coronavirus cases and correlated deaths. Daily new cases are calculated by subtracting the cumulative case count of the previous date from the current date's cumulated case count, and confirmed cases and deaths from COVID-19 before the outbreak, January 21st, 2020, were coded as 0 (Figure 1)

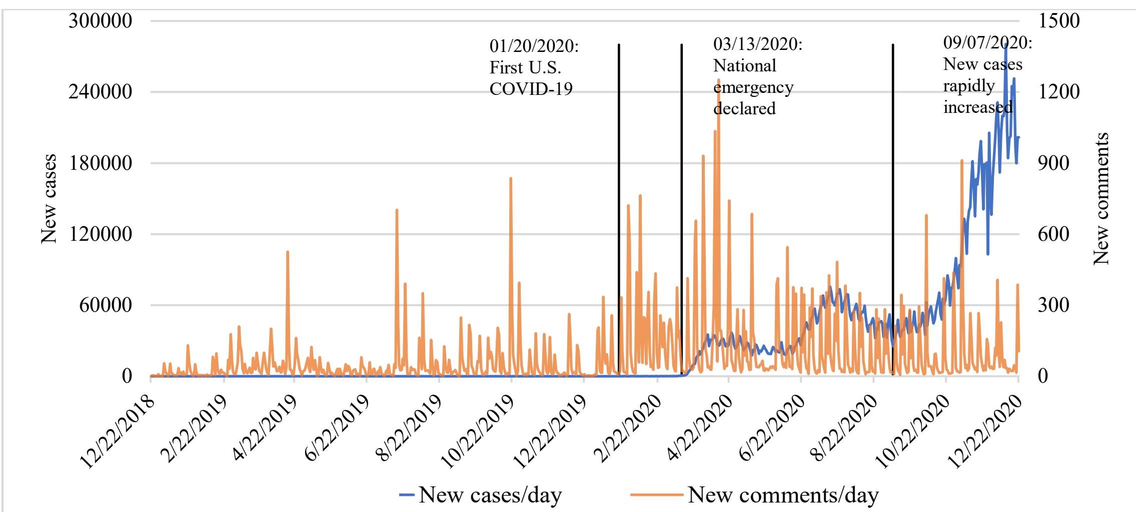

Figure 1. Count of meal kit comments on Trustpilot.com and daily COVID-19 new confirmed cases and

Customer review for meal kit companies

We collected 51, 497 customer reviews for meal kit companies from Trustpilot.com from 2019-2020. Nine companies had a fairly large number of customer reviews in our study time frame, including Home Chef, Freshly, Green Chef, Hello Fresh, Daily Harvest, Blue Apron, Gobble, Every Plate, and Dinnerly. Whenever reviews from multiple regions were provided, we selected the reviews from the U.S. company (e.g., we only collected data from U.S. Hello Fresh and omitted the reviews from U.K. Hello Fresh). Among the nine selected companies, two companies, Every Plate and Dinnerly, were advertised as affordable meal kits with a unit price per meal per person less than $5. All other companies had a unit price equal to or higher than $6.99. The reviews from the two affordable meal kits companies account for 36% of our total sample (Figure 2).

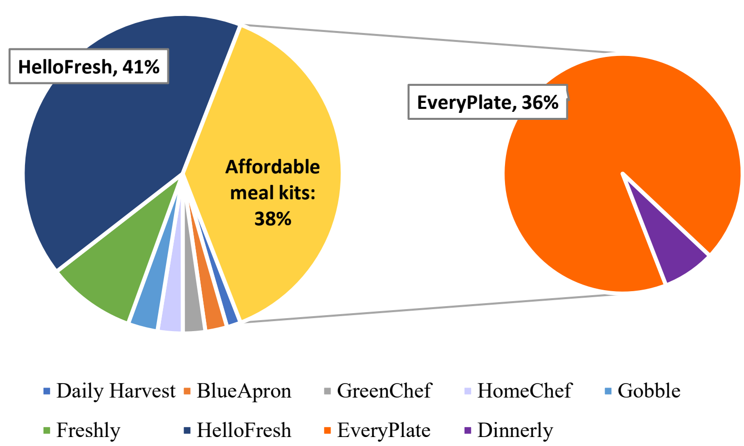

Figure 2. Customer review comment composition by company

On Trustpilot.com, each review is separately displayed in a small card as shown in Figure 3. We used data scrapping, a technique in which a computer program extracts human-readable data from a website, to collect the following information about reviews and about the customers who provided the reviews (Figure 3): *Consumer Name*, the username the reviewer is using to post the review; *Number of Reviews*, the total number of reviews that this username has posted on Trustpilot.com; *Overall Rating*, from 1 (bad) to 5 (excellent) how the reviewer evaluated the meal kit service that he/she received from the specific company; *Review Text*, the posted review content; *Invited*, whether the company invited the reviewer to provide the review; *Date*, the date the review was posted.

Figure 3. An example of customer review comments on Trustpilot.com

Next, based on the data collected from Trustpolit.com, we generated several variables to quantify the characteristics of the review. First, we split the *Date* information into variables including *Year*, *Month*, and *Day* to describe the common time patterns within a certain time range. Then, we counted the number of words from each comment using R as an indicator to quantify how much effort the customers put into his/her review.

Table 1. Summary statistics for customer review

Variable	**Obs**	**Mean**	**std**	**Min**	**Max**
Rating	51,497	4.36	1.02	1	5
Comment Length	51,497	40.97	44.92	0	922
Invited	51,497	0.95	0.22	0	1
# of comments	51,497	1.38	0.94	1	24

The review texts are key to understanding consumers' attitudes and individual preferences regarding the overall service provided by the meal kit companies. We used R to clean and process the review texts. We utilized a corpus consisting of all review text to remove irrelevant and infrequent terms. In the process, we removed numbers and special characters and transformed all terms to lowercase. Words with the same root and part of speech are combined into one term since they convey a similar meaning. For example, “delivery” and “deliveries” are combined into the term “delivery”. We applied a stop word list to remove irrelevant words. In particular, we removed basic and common terms that don’t provide insightful information, such as “have”, “give”, “will”. In addition, the names of meal kit companies are removed since the company name is separate. The word frequency of the final corpus is presented in Fig. 4 using Word Clouds. Fig. 4 shows that the most frequently mentioned words include “meal”, “good”, “recipe”, “fresh”, “easy”, and “service”, describing different aspects of a review.

Figure 4. Word cloud for meal kit customer review comments

Method

Latent Dirichlet Allocation

To better understand the contents of customer reviews for meal kit companies and quantify the underlying text information for further empirical analysis, we employed the Latent Dirichlet Allocation (LDA) method on 51, 497 customer reviews collected. LDA is a widely used natural language processing technique for both research and commercial purposes, such as financial document analysis and social media information retrieval (Dyer et al., 2017; Hong and Davison, 2010; Jelodar et al., 2019). As a topic modeling methodology first developed by Blei et al. (2003), LDA is a Bayesian computational linguistic technique that identifies the underlying set of topics in a collection of documents. These topics can best summarize the observed text in the documents in an objective and replicable matter with limited assumptions. Thus, the LDA model can reduce the complexity of the text corpus by capturing key information and statistical relationships from text and converting them into computable quantitative data.

According to Blei (2012) and Blei et al. (2003), LDA defines a *word* as the basic unit of discrete data in a *document* (in our case, a document is a customer review), a *corpus* as a set of *documents* (in our case, the corpus is our sample of customer reviews), and a latent *topic* (a topic is one underlying dimension generated from the text of all customer reviews using the LDA model) as characterized by a distribution over *words*. As a Bayesian generative model, LDA associates *words*, *documents,* the *corpus*, and the latent *topic* in the following way: each *document* is distributed over a set of latent *topics* and each *topic* is a multinomial distribution over a *word* vocabulary.

This procedure could be graphically illustrated in Figure 5. As seen in the figure, we define hyperparameters α and β as the Dirichlet parameter on the topic distribution over words and the Dirichlet parameter on the word distribution; M and N stand for documents and the repeated choice of topics and words within a document; $\boldsymbol{\theta}$ is the topic distribution and θ_m can be denoted as the topic distribution for a single document m; $\boldsymbol{z}$ is a set of K topics and $\boldsymbol{w}$ is a set of N words. LDA goes through each word w in a specific document m and for each topic k and operates in the following generative procedure: first, it chooses N (number of words) ~ *Poisson* (ξ) and θ (topic distribution) ~ *Dirichlet* (α); then for each of the N words w_n, LDA determines a topic z_n ~ *Multinomial* (θ) and further picks a word w_n from $p(w_n| z_n, \beta)$, which is a multinomial probability conditional on the topic z_n (Heng et al., 2018). At the end of the process, the LDA model generates the keywords associated with each of the topics and the probability that each document (i.e., customer review) is associated with a specific topic.

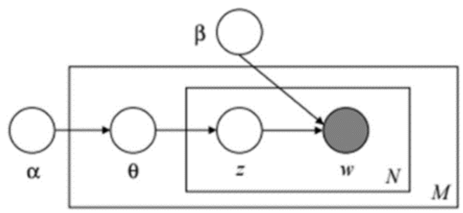

Figure 5. Graphical Illustration of LDA (See Blei et al., 2003)

In this study, we used software R and its package *ldatunning* which derives from *topicmodels* to analyze the customer reviews of meal kit services that we scrapped from Trustpilot.com (Hornik and Grün, 2011). When employing the LDA model, the most important procedure is to determine the optimized number of topics that characterize all text information. For this purpose, we used metrics developed by Cao et al. (2009) and Deveaud et al. (2014) to pin down the optimal topic number for the review text data. A good topic number should minimize Arun2010 and CaoJuan2009 metrics and maximize Deveaud2014 and Griffiths2004 metrics. As can be seen in Figure 6, we tested topic numbers ranging from 2 to 50 where Arun2010 and CaoJuan2009 metrics were relatively low while Deveaud2014 and Griffiths2004 metrics were relatively high. Based on the visualization of the four metrics, the optimal topic numbers range from 4 to 20. For simplicity and interpretability of the analysis, we chose 4 as our number of topics.

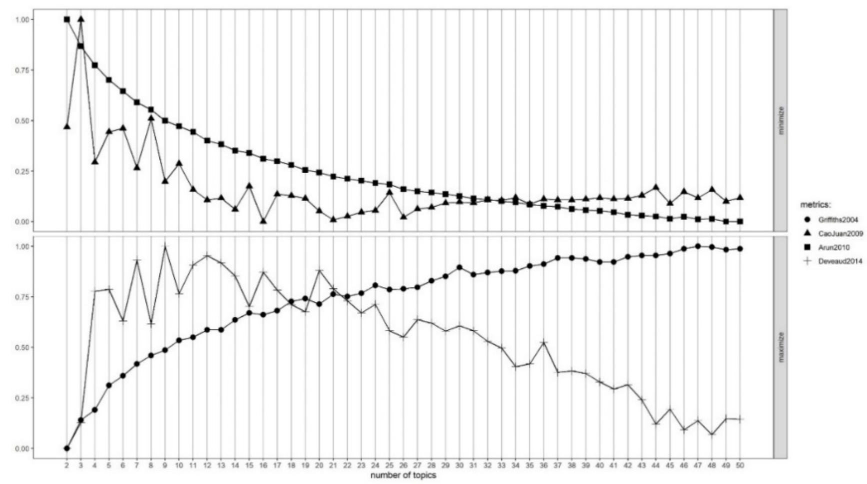

Figure 6. Selection of the number of topics

After determining the number of topics, the LDA model can exhibit each topic with a set of words from the training process. Table 2 presents the four topics and associated top keywords, as well as the corresponding percentage of each topic within customer reviews[1]. Since all text information was stemmed for generality before the LDA modeling, the key words output from LDA were in the words' the stem form. For easy interpretation, we present one example of the output with the stem of the keywords in Table 2 (e.g., delivery from the stem of "deliveri"), while the actual words used in the customer review text could be any forms from the stem (e.g., deliveries from the stem of "deliveri"). As displayed in Table 2, Topic 1 (*Experience*), featuring discussions on the customers' cooking or food experiences, accounts for 38% of the customer review texts. The most relevant keywords generated from this topic include "easy", "time", "ingredient", "prepare", and "family". Topic 2 (*Service*) focuses on the services provided by meal kit companies and a customer review with high prevalence on this topic is more likely to have keywords like "customer", "service", "delivery", "help", "cancel", and "refund". The average prevalence of Topic 2 among all review text is slightly less than 20%. Topic 3 (*Quality*) heavily emphasizes on the food quality that meal kits can provide and accounts for 23% of the review texts. Some keywords associated with this

[1] Prevalence or percentage of each topic within a customer review should add up to 1.

topic include “ingredient”, “fresh”, “meat”, “vegetable”, and “quality”. Lastly, Topic 4 (*Convenience*), which accounts for the rest 20% of the text, places an emphasis on the convenience offered by the meal kit service. The most critical keywords in this topic are “ingredient”, “follow”, “prepare” and “instruction” etc. To better demonstrate the identified topics, for each topic, we included one example of the original comments that had a prevalence greater than 95% for the topic discussed in that row.

Table 2. summary statistics for topics

	Prevalence (%)	Min (%)	Max (%)	Top key words	Example of review comments
Topic 1 (Experience)	37.85	0.06	99.42	easy, fresh, cook, recipe, delicious, time, ingredient, dinner, prepare, new, family, follow, service, grocery, use, home, us, work, quality, husband	The kids and I have been enjoying receiving the boxes! We Love to cook in the kitchen together and the meals are fantastic.
Topic 2 (Service)	19.07	0.06	99.80	service, custom, delivery, time, help, receive, fresh, cancel, credit, delivery, call, refund, arrive, quick, use, thank, account, email, ship, send	Overall company provides a good service. Recipes for meals are very good. Shipping is fast and kept updated with what is shipped and when it is to arrive.
Topic 3 (Quality)	22.97	0.05	99.55	recipe, ingredient, fresh, meat, time, package, produce, use, chicken, vegetable, quality, receive, little, item, option, veggie, potato, portion, pack, dish	Top end food products that taste delicious. I like the fresh garlic and ginger plus fresh vegetables that comes in each package. The meats are low fat and very fresh. Also, a plus for the meatless options. Love the pastas!
Topic 4 (Convenience)	20.10	0.04	98.28	easy, recipe, fresh, delict, service, ingredient, follow, quality, price, option, prepare, variety, portion, delivery, time, tasty, taste, use, choice, instruct	Nice selection of meal options. Great price! Easy to follow instructions

Empirical Strategy

To investigate the impacts of the COVID-19 pandemic on consumers’ preferences to meal kit services, we compared the changes of customer preferences before and after the pandemic for the meal kits, which were substantially influenced by the pandemic versus those that were minimally affected by COVID-19. We hypothesize that, during the pandemic, the customers, especially those with higher household income and could work from home during the COVID-19 lockdown (Angelucci et al., 2020), would demand better service from the meal kit companies which provided human contact free services, e.g., speedy fresh food delivery to the front door and excellent online customer services. On the other hand, due to the longer hours staying at home, those customers may value the time-saving feature from the meal kits less. Since higher-income individuals who worked at home during the lockdown were more likely to be the customers from conventional meal kits with higher prices, we expect the topics discussed in the comments from those companies may have been changed since the outbreak of COVID-19 (e.g., Hello Fresh). However, affordable meal kits (e.g., Every Plate), given their minimal costs (i.e., $4.99/person/meal), had a more constrained ability to improve their services during the pandemic further. Further, their target customers were more likely the ones with lower household income and those who were less likely to work from home during the pandemic according to the recent surveys. Therefore, customers of the affordable meal kit companies may value the contact-free features less than the customers who had subscriptions from conventional meal kit companies. In that case, the focuses of the comments from this group may remain similar before and after the pandemic. Therefore we categorize the affordable meal kit companies, including Every Plate and Dinnerly, as the control group and all other meal kit companies as treatment groups whose price per person per meal is equal to or above $6.99.

In Figure 7, we presented the Google trend index for Hello Fresh (blue line) and Ever Plate (orange line), the two most commented companies from the treatment and control groups, respectively, from January 2019 to December 2020. Comments from Hello Fresh account for 66% of the treatment group and 41% of the total sample, while Every Plate accounts for 95% of the control group and 36% of the total sample. Hence, we believe these two companies are representative enough for their group. The Google Trend confirmed our hypothesis that the COVID-19 pandemic less influenced the companies in the control group.

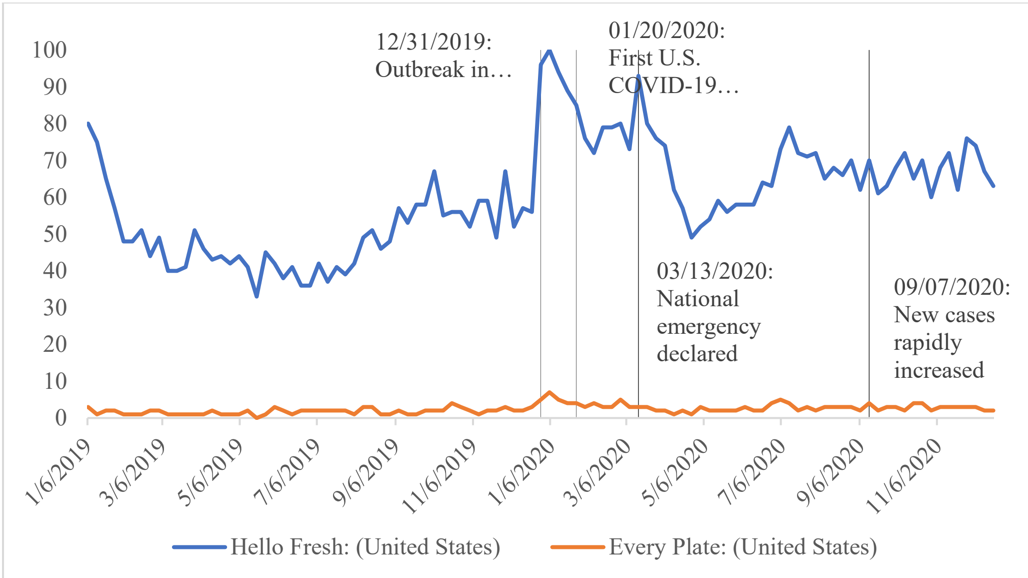

Figure 7 Google trends for Hello Fresh and Every Plate in 2019-2020

As shown in Figure 7, Hello Fresh is a more well-known company than Every Plate during the entire study period. The Google trends from the two companies, though different, were relatively parallel before the pandemic from January 2019 to January 2020. Since the outbreak of COVID-19 in late December 2019 in Wuhan in China and the first case confirmed in the U.S. on January 20th 2020, the discussion related to Hello Fresh online increased substantially and it enjoyed a higher level of attention in the entire year in 2020 compared to pre-pandemic periods. Similar attention increase was not observed from Every Plate, which further confirmed our hypothesis that *Affording meal kits* were less affected by the pandemics and can work as the control groups to capture the general time trend for the entire meal kit industry.

Difference in Difference

The basic empirical specification to estimate the impacts of COVID-19 pandemic on meal kit comment focus is as follows:

$$
\begin{aligned}
Log(Topic\%_{cit}+1) = a_0 &+ \beta_1 Treat_c \times Post_t \times Log(Covid_t+1) + \beta_2 Treat_c \times Log(Covid_t+1)_t \\
&+\beta_3 Post_t \times Log(Covid_t+1)_t + \beta_4 Treat_c \times Post_t \\
&+\beta_4 Treat_c + \beta_5 Post_t + \beta_6 Log(Covid_t+1) \\
&+\gamma' \mathbf{X}_{cit} + Month_t + Comp_c + \varepsilon_{cit}
\end{aligned} \tag{1}
$$

where *Topic* $\%_{cit}$ is the estimated prevalence of the interested topic in the comment from company *c*, reviewer *i*, on day *t*. $Treat_c$ is equal to 1 if the company is in the treatment group, otherwise 0 if the company in the control group (i.e., *Affordable meal kits*). $Post_t$ is an indicator of the outbreak of the pandemic. If the comment was posted after the outbreak, $Post_t$ is equal to 1 otherwise 0. $Covid_t$ measures the severity of the pandemic with metrics including new cases on the posting date or the lagged news cases. In this model, we included month by year fixed effects and company fixed effects to control the common time trends shared by all companies by month ($Month_t$) and also the time invariant variation ($Comp_c$). We also controlled for the characteristics of the comment and the reviewer with $\mathbf{X}_{cit}$ including whether the reviewer was invited to provide the review, how many comments the reviewer had posted on this website, and the length of the comment. The coefficient of interest is β_1 which measures the distinct changes in discussion topics identified from the review comments between the treatment and the control as the pandemic progressed.

The validity of the *Difference in Difference* (DID) specification depends on the parallel trend assumption, i.e., the control group works as a valid control in the sense that, before the pandemic, the topic focuses between the treatment and control groups followed a parallel trend and were similar enough in 2019. While this parallel trend assumption cannot be tested comprehensively, graphical analysis and partial tests are possible. Our findings show that the parallel trend assumption holds for both graphical (Figure 7 and 8) and regression analysis. According to Figure 7, the Google trends between the two groups were flat parallel in 2019, supporting the hypothesis that the overall discussion and attentions the two groups attracted online before the pandemic were parallel. In Figure 8, we investigated the parallel trends in detail by comparing the prevalence of each topic between the two groups before and after the pandemic. The blue lines represent the 60-day moving average of the topic prevalence among the comments for the treatment companies while the red lines showcase the topic prevalence among the comments for the control companies.

According to Figure 8, the prevalence of the four topics in the treatment and control groups shared parallel trends before the pandemic in 2019. We formally test this parallel trend assumption in Table 3, where we only used the sample in 2019 before the pandemic and the differences in the topic prevalence between treatment and control groups were compared by month to the difference in January 2019 which is the first month of our study period. For all the four topics, the differences between the treatment and the control groups were similar to the first month in the entire year, except for February and October 2019. The insignificant changes in the differences between the two groups before the pandemic suggest the identification strategy that we adopted is valid.

Table 3. Parallel test

	Topic 1	Topic 2	Topic 3	Topic 4
	Experience	Service	Quality	Convenience
VARIABLES	(1)	(2)	(3)	(4)
Omitted: Treat x Jan				
Treat x Feb	-0.191*	0.349***	-0.016	-0.142**
	(0.105)	(0.073)	(0.090)	(0.068)
Treat x March	0.086	-0.056	-0.066	0.036
	(0.120)	(0.084)	(0.102)	(0.077)
Treat x April	0.028	-0.008	-0.006	-0.014
	(0.106)	(0.074)	(0.090)	(0.068)
Treat x May	0.070	-0.026	-0.034	-0.009
	(0.104)	(0.073)	(0.089)	(0.067)
Treat x June	0.108	-0.034	-0.079	0.004
	(0.104)	(0.073)	(0.089)	(0.067)
Treat x July	0.087	-0.035	-0.049	-0.003
	(0.104)	(0.072)	(0.089)	(0.067)
Treat x August	0.086	0.026	-0.148	0.036
	(0.121)	(0.085)	(0.104)	(0.078)
Treat x September	0.108	-0.107	0.005	-0.005
	(0.123)	(0.086)	(0.105)	(0.080)
Treat x October	-0.215**	0.349***	-0.011	-0.123*
	(0.104)	(0.073)	(0.089)	(0.067)
Treat x November	0.069	-0.022	-0.002	-0.045
	(0.105)	(0.073)	(0.090)	(0.068)
Treat x December	0.117	0.019	-0.145	0.008
	(0.109)	(0.076)	(0.093)	(0.070)
Company fixed effects	Yes	Yes	Yes	Yes

Month fixed effects	Yes	Yes	Yes	Yes
Observations	12,142	12,142	12,142	12,142
R-squared	0.057	0.148	0.022	0.075

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

We note that the companies in the treatment and control groups were different on many metrics even before the pandemic. For example, companies in the treatment groups were more expensive, enjoyed higher customer ratings (Table 4), and attracted more attention on the market (Figure 6). The topics that the customers from the two groups of companies discussed were also slightly different, i.e., customers in the treatment groups paid more attention to the food experiment they enjoyed from the meal kits (keywords including family, time, home, and husband) while customers from the control groups discussed more services provided by meal kit companies (Figure 7 and Table 4). DID model allows for the unobservables to be correlated with treatment as long as the time-varying unobservables do not affect the topic discussion differently between the treatment and control groups (Jack and Suri, 2014). In that case, though the coefficient for the market liberalization could be biased, the variable *Treat* should capture the endogeneity and give an unbiased estimation for the coefficients for the interactions.

Table 4. Topics statistics by periods

	Pre-pandemic (before 1/20/2020)				During the pandemic (after			
	Expensive		Affordable		Expensive		Affordable	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Rating	4.443	0.946	4.236	1.050	4.467	0.956	4.177	1.144
Topic 1 (Experience)	0.428	0.343	0.358	0.329	0.358	0.345	0.355	0.335
Topic 2 (Service)	0.144	0.241	0.139	0.231	0.271	0.351	0.167	0.267
Topic 3 (Quality)	0.215	0.276	0.266	0.305	0.200	0.273	0.269	0.314
Topic 4 (Convenience)	0.215	0.276	0.266	0.305	0.200	0.273	0.269	0.314

Results

Several critical points emerged when described the progression of the pandemic. The first critical point was 12/31/2019, when the outbreak of the COVID-19 in Wuhan was released and induced a great discussion about the potential damage that this virus could cause and a global pandemic that it could end in without a good control (Lee, 2020). Due to the concerns of a global pandemic that this virus could cause, the discussion on the potential responses to the outbreak and meal kit delivery increased. Therefore, we observed a surge in Hello Fresh discussion online (Figure 7). However, meal kit customers in the U.S, at that point time, hadn't fully digested the news in China and didn't expect to make decision changes by themselves. Therefore, though a surge in discussion on meal kits was a witness in late December 2019 and early January 2020, the topics underlying among meal kit comments were not changed significantly until a first case was confirmed in the U.S on January 20th 2020. After the first COVID-19 case was confirmed in California, we witnessed an immediate increase in the discussion related to *Services* (keywords including delivery, customer service, and arrive) that meal kit companies can provide (Topic 2). Such increase was greater among the comments in the treatment group while smaller in the control group. The second critical point occurred around September 2020 when the United States witnessed a second wave of the case surge and the number of new confirmed cases and death each day increased scarily. As the pandemic progressed into the more terrifying second wave, more and more customers from the treatment group valued the service that meal kit companies provided and the prevalence of the second topic, *Service*, increased dramatically from 32% to nearly 70%. Compared to the increased discussion in service, as a trade-off, all other three topics suffered a decrease in their prevalence among the posted comments. The topics related to the quality of food and the quality of the experience suffered the most decrease. Starting September 2020, the discussion focused on Topic 1 (*Experience*) decreased more than 15 percentage points, from 31% to 14% and the prevalence for Topic 3 focusing on the quality and the types of the food provided by the meal kits reduced more than 10%. Customers in the treatment group paid fairly less attention to the topic related to *Convenience* since the outbreak of COVID-19 in the U.S. and the prevalence of the fourth topics among their comments gradually decreased from 20% in April 2020 to less than 10% at the end of our study period in December 2020. The changes in the control group, however, were not obvious in all four identified topis. The trends in 2020 after the pandemic were fairly similar to the ones in 2019 before the pandemic.

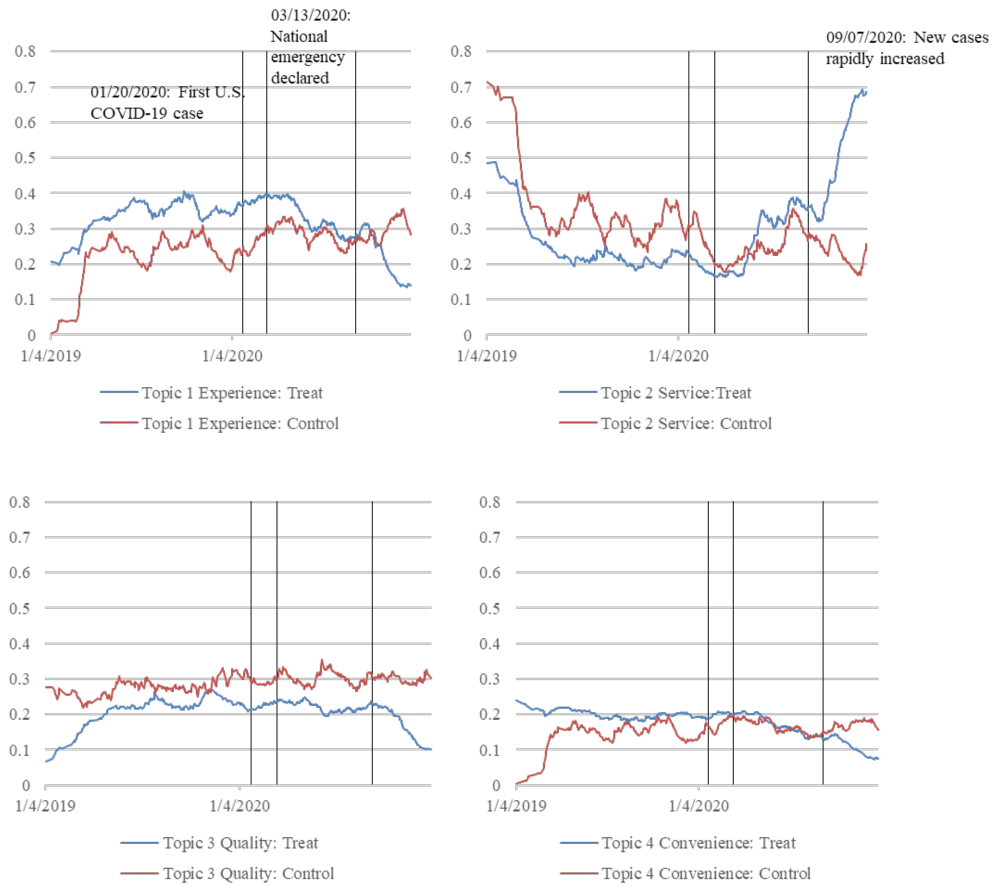

Figure 8. 60-day moving average of the prevalence of each topic

We estimated the impacts of the COVID-19 pandemic on customer's discussion focuses for meal kit services by comparing the changes of topic prevalence predicted from review comments between the treatment and control groups before and after the pandemic. Table 5 presents the estimations based on equation (1) where panel A defined the outbreak of the COVID-19 pandemic based on the national emergency declaration on March 13th, 2020 (the start of the first wave in the U.S) while panel B defined the outbreak based on the start of the second wave, September 7th, 2020 when the number of new cases per day started increasing rapidly.

In column (1) – (4) and column (9) – (12), we defined the pandemic as whether or not the pandemic outbroke (or entered the second wave) using a dummy variable *Post*. The coefficient for *Treat x Post* compared the different changes in the outcome variables between the treatment and control groups after the national emergency declaration. Compared to the control group, comments from the treatment groups had a significantly greater increase in the discussion related to the services provided by meal kit companies (1.8%) while less increase in the discussion on the topic related to food experience (-2.2%). Such changes were greater when defining the *Pre* and *Post* timeframe based on the start of the second wave of pandemic (i.e., after September 7th, 2020). According to column (9) – (12), compared to the comments in the control group, the comments in the treatment groups experienced a dramatic change in the topic composition since the start of the second wave. The prevalence of topic 2, which was closely related to the services from meal kit companies, had an increase that was 13% greater in the treatment group than in the control group after September 7th, 2020. Trading off the greater emphasis on the company service, the comments in the treatment groups had less discussion on the topics related to food experience (topic 1), food type and quality (topic 3), and the convenience provided by the meal kit products (topic 4). Compared to the comments for the affordable meal kit companies, since the second wave of the pandemic started, the companies in the treatment group suffer 6.8%, 4.1%, and 3.6% more reduction for the discussion related to food experience, food type, and quality and product convenience respectively.

Table 5. Regression Estimations

Panel A: March 2020	Post = After March 13th, 2020							
	Measurement of the pandemic: *Post*				Measurement of the pandemic: *Post x log(cases+1)*			
Dependent Variable	Topic 1 Experience (1)	Topic 2 Service (2)	Topic 3 Quality (3)	Topic 4 Convenience (4)	Topic 1 Experienc (5)	Topic 2 Service (6)	Topic 3 Quality (7)	Topic 4 Convenience (8)
Post	0.022***	-0.009	-0.014**	0.000	-0.089***	0.171***	-0.073***	-0.032*
	(0.008)	(0.007)	(0.007)	(0.006)	(0.026)	(0.021)	(0.023)	(0.018)
Treat x *Post*	-0.022***	0.018***	-0.000	0.004	0.154***	-0.382***	0.161***	0.110***
	(0.005)	(0.004)	(0.004)	(0.003)	(0.023)	(0.019)	(0.021)	(0.016)
Treat x log(cases+1)					0.014	-0.004	-0.014	0.007
					(0.010)	(0.008)	(0.009)	(0.007)
Treat x log(cases+1)					0.024***	-0.033***	0.008*	0.006
					(0.005)	(0.004)	(0.005)	(0.004)
Treat x *Post* x log(cases+1)					-0.046***	0.075***	-0.015	-0.026***
					(0.010)	(0.008)	(0.009)	(0.007)
Observations	51,497	51,497	51,497	51,497	51,497	51,497	51,497	51,497
R-squared	0.069	0.176	0.034	0.076	0.070	0.184	0.035	0.077
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Company fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: September 2020	Post = After September 7th, 2020							
	Measurement of the pandemic: *Post*				Measurement of the pandemic: *Post x log(covid)*			
Dependent Variable	Topic 1 Experience (9)	Topic 2 Service (10)	Topic 3 Quality (11)	Topic 4 Convenience (12)	Topic 1 Experienc (13)	Topic 2 Service (14)	Topic 3 Quality (15)	Topic 4 Convenience (16)
Post	0.022*	-0.061***	0.016	0.036***	-0.236**	0.310***	-0.032	-0.072
	(0.012)	(0.010)	(0.011)	(0.009)	(0.103)	(0.084)	(0.093)	(0.072)
Treat x *Post*	-0.068***	0.130***	-0.041***	-0.036***	0.336***	-0.862***	0.411***	0.212***
	(0.006)	(0.005)	(0.006)	(0.004)	(0.060)	(0.049)	(0.054)	(0.042)
Treat x log(cases+1)					-0.000	-0.003***	0.001	0.003***
					(0.001)	(0.001)	(0.001)	(0.001)
Treat x log(cases+1)					0.042**	-0.056***	0.004	0.017
					(0.018)	(0.015)	(0.016)	(0.013)
Treat x *Post* x log(cases+1)					-0.064***	0.158***	-0.072***	-0.041***
					(0.010)	(0.008)	(0.009)	(0.007)
Observations	51,497	51,497	51,497	51,497	51,497	51,497	51,497	51,497
R-squared	0.071	0.186	0.035	0.077	0.072	0.192	0.036	0.078
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Company fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Besides the pandemic outbreak, the pandemic's progression can also shift meal kit customers' focuses toward the products. In column (5)-(8) and column (13)-(16), we had the log transformed daily new confirmed cases (log(cases+1)) interact with *Treat* x *Post* to more accurately measure for the progression of the pandemic. In this case, the variable of interest is *Treat* x *Post* x *log(cases+1)* and the its coefficient estimates the shifts of customer interests as the more confirmed cases of COVID-19 were reported. The results are very similar to the estimations discussed above. We found that, as more COVID-19 cases were confirmed each day, compared to the comments for affordable meal kits in the control groups, the comments in the treatment placed a higher emphasis on the company service and discussed less about features related to food quality and experience and product convenience. We also test the robustness of the results by utilizing the one week lagged confirmed cases as the measurement of the pandemic progression since the post of the comment may be delayed or it could take time for customers to digest the progression of the pandemic before they changed their real preferences or behaviors regarding meal kits. The estimations of the robustness check were presented in Table 6 and conclusions were similar to those derived from Table 5.

Table 6. Robustness check: Lagged COVID-19 confirmed cases

	Post = After March 13th, 2020			
Dependent Variable	Topic 1 Experience (1)	Topic 2 Service (2)	Topic 3 Quality (3)	Topic 4 Convenience (4)
Post	-0.090***	0.154***	-0.055**	-0.032*
	(0.026)	(0.021)	(0.023)	(0.018)
Treat x *Post*	0.151***	-0.384***	0.165***	0.111***
	(0.023)	(0.019)	(0.021)	(0.016)
Treat x log(cases+1)	0.012	-0.002	-0.009	0.001
	(0.013)	(0.011)	(0.012)	(0.009)
Treat x log(cases+1)	0.023***	-0.029***	0.005	0.005
	(0.005)	(0.004)	(0.005)	(0.004)
Treat x *Post* x log(cases+1)	-0.043***	0.074***	-0.021*	-0.019**
	(0.013)	(0.011)	(0.012)	(0.009)
Observations	51,497	51,497	51,497	51,497
R-squared	0.070	0.184	0.036	0.077
	Post = After September 7th, 2020			

Dependent Variable	Topic 1 Experience (5)	Topic 2 Service (6)	Topic 3 Quality (7)	Topic 4 Convenience (8)
Post	-0.219**	0.358***	-0.083	-0.093
	(0.098)	(0.080)	(0.088)	(0.069)
Treat x *Post*	0.316***	-0.825***	0.398***	0.203***
	(0.059)	(0.048)	(0.053)	(0.041)
Treat x log(cases+1)	-0.000	-0.002***	0.001	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
Treat x log(cases+1)	0.039**	-0.065***	0.013	0.021*
	(0.017)	(0.014)	(0.015)	(0.012)
Treat x *Post* x log(cases+1)	-0.061***	0.153***	-0.070***	-0.039***
	(0.009)	(0.008)	(0.008)	(0.007)
Observations	51,497	51,497	51,497	51,497
R-squared	0.072	0.192	0.036	0.078

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

The role of comment topics on customer review ratings

One of the most important metrics from the customer review is the review ratings, reflecting the customers' satisfaction toward the product and/or the company. These metrics also suggest the potential of a product/company to keep an existing customer in the future (Mudambi and Schuff 2010). Therefore, review ratings have been widely used as an indicator to understand customer preferences (Radojevic et al., 2017; Sriv & Sorenson, 2010).

The previous analysis has demonstrated a shift in customer interests for meal kit products from features associated with the product itself (i.e., food quality, experience, and convenience) to features associated with the service provided by the companies (i.e., delivery, customer service). In this section, we continued to investigate, besides the prevalence of the topics, whether the role of each topic on determining the customer review ratings also altered after the pandemic. For example, before the pandemic, the comments with a higher percentage of words discussing customer services and meal kit delivery might be associated with lower ratings as customers didn't value good service but complaint bad services. However, after the pandemic, more customers value the human contact services provided by meal kit companies. Therefore a comment with a higher prevalence on discussion related to services may enjoy higher ratings.

To understand the changes in customers' preferences towards four identified features/topics, we compared the changes of the roles of the four topics in influencing customer rating with the following equation.

$$\begin{aligned}Rating_{cit} = {} & a_0 + \theta_1 T1_{cit} + \theta_1 T2_{cit} + \theta_1 T3_{cit} + \delta_1 Treat_c \times Post_t + \delta_1 Treat_c + \delta_1 Post_t \\ & + \beta_1 Treat_c \times Post_t \times Topic1_{cit} + \beta_2 Treat_c \times Topic1_{cit} + \beta_3 Post_t \times Topic1_{cit} \\ & + \beta_4 Treat_c \times Post_t \times Topic2_{cit} + \beta_5 Treat_c \times Topic2_{cit} + \beta_6 Post_t \times Topic2_{cit} \\ & + \beta_7 Treat_c \times Post_t \times Topic3_{cit} + \beta_8 Treat_c \times Topic3_{cit} + \beta_9 Post_t \times Topic3_{cit} \\ & + \gamma' X_{cit} + Month_t + Comp_{ci} + \varepsilon_{cit}\end{aligned} \quad (2)$$

$Rating_{cit}$, from 1 (bad) to 5 (excellent), measures how the reviewer evaluated the meal kit service that she received from the specific company. $T1_{cit}$, $T2_{cit}$, $T3_{cit}$ are the prevalence of topic 1 (convenience), topic 2 (service), and topic 3 (food quality). Given that the sum of prevalence from four topics is equal to 1, the prevalence of topic 4 is omitted. Then, we had *Treat* x *Post* interacted with each topic to investigate changing roles of each topic on customer ratings since the pandemic started. We estimated equation 2 using an ordered logit regression and the estimated coefficients were presented in Table7.

Table 7. Logit estimations of topics prevalence on customer ratings

		Panel A: Post = March 2020		
Dependent Variable: Ratings	(Omitted topic: Topic 4)	Topic 1 Experience	Topic 2 Service	Topic 3 Quality
		0.645***	-2.694***	-3.090***
		(0.139)	(0.144)	(0.129)
x *Treat*	-0.064	0.215	0.065	0.190
	(0.142)	(0.172)	(0.176)	(0.163)
x *Post (March)*	0.376***	-0.205	-1.228***	-0.285*
	(0.142)	(0.175)	(0.174)	(0.160)
x *Treat* x *Post*	-0.069	0.124	2.241***	-0.074
	(0.168)	(0.231)	(0.225)	(0.217)
Observation	51,497			
		Panel B: Post = September 2020		

Dependent Variable: Ratings	(Omitted topic: Topic 4)	Topic 1 Experience	Topic 2 Service	Topic 3 Quality
		0.434***	-2.878***	-3.294***
		(0.110)	(0.124)	(0.101)
x *Treat*	-0.166	0.351***	0.190	0.323***
	(0.110)	(0.131)	(0.146)	(0.123)
x *Post (September)*	-0.272**	0.128	-0.957***	0.034
	(0.112)	(0.108)	(0.127)	(0.094)
x *Treat* x *Post*	0.327***	-0.053	2.058***	-0.249**
	(0.064)	(0.109)	(0.138)	(0.097)
Observation	51,497			

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Before the COVID-19 pandemic, comments with more discussion on the experience provided by the meal kit products, compared to a topic about convenience, were associated with higher customer review ratings, while the comments with more discussion on food quality and company services were more likely to have a lower rating. This suggests that customers appreciate the experience and adventure from the meal kit service before the pandemic and complained about food quality and company services before the pandemic. The comments in the treatment group valued the food experience and the food quality significantly more than the comments in the control group (i.e., affordable meal kits).

Since the pandemic outbreak in March, the customers in the control groups valued the convenience provided by the meal kit service significantly more so that a similar percentage of discussion on the convenience feature could lead to a higher rating than before. On the other hand, a similar length (%) of discussion on the company service suggested an even lower score. This could be caused by a decreased interest in a positive service experience or an enhanced dislike of an unsatisfied service experience. Without further sentiment analysis, we are unable to differentiate the two. Unlike the control groups, the comments in the treatment groups showcased an improved appreciation of the service provided by meal kits companies and the convenience offered by the meal kits. This preference changes consistent with the existing literature that convenience directly impacts consumer decisions during COVID-19 (Brewer & Sebby, 2021; Widiyanto & Wibowo, 2021).

Discussion and limitation

Using the online review comments and the text mining technique, we estimated the impacts of the COVID-19 pandemic on customer interest focuses on meal kit services by comparing the changes in the prevalence of four underlie topics derived from customer comments between the treatment and control groups before and after the pandemic. Our empirical results showed that, as the pandemic occurred (or as more COVID-19 cases confirmed each day), comments from the conventional meal kit group placed a higher emphasis on the service aspect and lower emphasis on features related to food quality, experience, and the product convenience, compared to the comments for affordable meal kits in the control groups. In terms of the review ratings, unlike the control groups, the comments in the treatment groups showcased an improved appreciation in both the service and the convenience aspects. Our empirical results indicate that people who can afford expensive meal kits pay more attention to the service offered by the companies during the pandemic and are also more likely to give higher ratings when their experiences with the service are positive. With these changes in customers' preferences, the food delivery companies should adapt quickly to maintain existing customers and even attract new ones. At the same time, people who purchase cheap meal kits, i.e., those in the control group, are more likely to receive low incomes before and during the pandemic. Our results suggest that they do not shift their preferences as much as the high-income group. This may be due to many reasons, such as insufficient awareness of pandemic-related sanitary requirements or simply financial stress they suffer. Consequently, the cheap meal kit companies may lack enough incentives to improve their service that fit the pandemic-related standards. This could further increase the gap between different income groups regarding the overall quality of food delivery people receive.

We have a few limitations in this study. First of all, our data has only a few characteristics at the customer review level and no information is available for customers themselves. One important requirement for the DID coefficient to be unbiased is that time-varying unobservables do not affect the outcome variable differently for the treatment and control groups. However, given out limited data, we could not exclude this possibility or conduct further tests. We can only refer to groups based on which meal kit, the conventional or the affordable one, customers purchased before and during the pandemic, yet one single customer may shift between treatment and control groups in our case. Second, we may need further evidence to test our hypothesis on what factors are associated with customer ratings on the meal kits. With the LDA method, we are able to discover underlying topics and corresponding prevalence in each review comment. However, customers' attitudes or tones reflected in these comments are still undiscovered. There are techniques like sentiment analysis that can detect positive or negative attitudes in the text information to further validate our hypotheses, and we plan to include them in our future work.

Appendix

Table A1 Count of comment by month

	Freq.	Percent	Cum.
2019	12,142	23.58	23.58
2020	39,355	76.42	100.00
January	3,862	7.50	7.50
February	5,504	10.69	18.19
March	5,475	10.63	28.82
April	7,022	13.64	42.45
May	3,327	6.46	48.92
June	4,952	9.62	58.53
July	4,524	8.78	67.32
August	4,108	7.98	75.29
September	3,968	7.71	83.00
October	3,299	6.41	89.41
November	3,802	7.38	96.79
December	1,654	3.21	100.00

References:

Acosta. 2020. *New Acosta Report Details How COVID-19 Is Reinventing How America Eats*. Acosta. https://www.acosta.com/news/new-acosta-report-details-how-covid-19-is-reinventing-how-america-eats.

Angelucci, M., Angrisani, M., Bennett, D., Kapteyn, A., and Schaner, S. 2020. Remote Work and the Heterogeneous Impact of COVID-19 on Employment and Health. https://doi.org/10.3386/w27749

Blei, D. M. 2012. Probabilistic topic models. *Communications of the ACM* 55(4):77–84. https://doi.org/10.1145/2133806.2133826

Blei, D. M., Ng, A., and Jordan, M. I. 2003. Latent dirichlet allocation. *The Journal of Machine Learning Research* 3:993–1022. https://doi.org/10.5555/944919.944937

Brewer, P., and Sebby, A. G. 2021. The effect of online restaurant menus on consumers' purchase intentions during the COVID-19 pandemic. *International Journal of Hospitality Management* 94:102777. https://doi.org/10.1016/j.ijhm.2020.102777

Cao, J., Xia, T., Li, J., Zhang, Y., and Tang, S. 2009. A density-based method for adaptive LDA model selection. *Neurocomputing* 72(7-9):1775–1781. https://doi.org/10.1016/j.neucom.2008.06.011

Chang, H. H., and Meyerhoefer, C. D. 2020. COVID-19 and the Demand for Online Food Shopping Services: Empirical Evidence from Taiwan. *American Journal of Agricultural Economics* 103(2):448–465. https://doi.org/10.1111/ajae.12170

Deveaud, R., SanJuan, E., and Bellot, P. 2014. Accurate and effective latent concept modeling for ad hoc information retrieval. *Document Numérique* 17(1):61–84. https://doi.org/10.3166/dn.17.1.61-84

Durbin, D. 2019. *Blue Apron latest to suffer in tough meal kit market*. AP NEWS. https://apnews.com/article/d0b241c26425443ca181010bb886ad8c.

Dyer, T., Lang, M., and Stice-Lawrence, L. 2017. The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics* 64(2-3):221–245. https://doi.org/10.1016/j.jacceco.2017.07.002

Ellison, B., McFadden, B., Rickard, B. J., and Wilson, N. L. 2020. Examining Food Purchase Behavior and Food Values During the COVID‐19 Pandemic. *Applied Economic Perspectives and Policy* 43(1):58–72. https://doi.org/10.1002/aepp.13118

Gentzkow, M., Kelly, B., and Taddy, M. 2019. Text as Data. *Journal of Economic Literature* 57(3):535–574. https://doi.org/10.1257/jel.20181020

Grashuis, J., Skevas, T., and Segovia, M. S. 2020 Grocery Shopping Preferences during the COVID-19 Pandemic. *Sustainability* 12(13):5369. https://doi.org/10.3390/su12135369

Grün, B., and Hornik, K. 2011. topicmodels: An R Package for Fitting Topic Models. *Journal of Statistical Software* 40(13). https://doi.org/10.18637/jss.v040.i13

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. https://doi.org/10.1016/j.jretconser.2018.02.006

Hong, L., and Davison, B. D. 2010. Empirical study of topic modeling in Twitter. *Proceedings of the First Workshop on Social Media Analytics - SOMA '10*. https://doi.org/10.1145/1964858.1964870

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., and Zhao, L. 2018. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications* 78(11):15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Judkis, M. 2017. *The meal-kit industry is at a crossroads. Will it ever figure out what we really want?* The Washington Post. https://www.washingtonpost.com/lifestyle/food/the-meal-kit-industry-is-at-a-crossroads-will-it-ever-figure-out-what-we-really-want/2017/10/06/74f239cc-842a-11e7-ab27-1a21a8e006ab_story.html.

Lee, A. 2020. Wuhan novel coronavirus (COVID-19): why global control is challenging? *Public Health* 179. https://doi.org/10.1016/j.puhe.2020.02.001

Leon, R. de. 2020. *How the coronavirus pandemic delivery surge created a lifeline for Blue Apron meal kits*. CNBC. https://www.cnbc.com/2020/05/22/how-coronavirus-pandemic-delivery-surge-gave-new-life-to-blue-apron.html

Mercatus. n.d. *eGrocery Adoption: The New Reality for Grocery Shopper Behavior*. Mercatus. https://info.mercatus.com/egrocery-shopper-behavior-report.

Mudambi, S.M. and Schuff, D. 2010. Research note: What makes a helpful online review? A study of customer reviews on Amazon. com. *MIS quarterly* :185-200.

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., and Agha, R. 2020. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery* 78:185–193. https://doi.org/10.1016/j.ijsu.2020.04.018

Radojevic, T., Stanisic, N., and Stanic, N. 2017. Inside the Rating Scores: A Multilevel Analysis of the Factors Influencing Customer Satisfaction in the Hotel Industry. *Cornell Hospitality Quarterly* 58(2):134–164. https://doi.org/10.1177/1938965516686114

Sriv, A., and Sorenson, P. G. 2010. Service Selection Based on Customer Rating of Quality of Service Attributes. *2010 IEEE International Conference on Web Services*. https://doi.org/10.1109/icws.2010.32

SUBTA. 2020. *The Evolution of Meal Kits and COVID's Impact on the Market: Blog*. SUBTA. https://subta.com/the-evolution-of-meal-kits-and-covids-impact-on-the-market/.

Troy, M., and Acosta, G. 2018. *What Is a Meal Kit Anyway?* Retail Leader. https://retailleader.com/what-meal-kit-anyway.

Yin, G. and Chen, J., 2020, October. Improving Causal Inference with Text as Data in Empirical IS Research: A Machine Learning Approach. *The 48th International Conference on Information Systems.*

Widiyanto, G., and Wibowo, F. X. P. 2021. Analysis of the Effect Product Quality, Trustworthiness, Convenience, Perceptions of Usefulness and Price on Purchase Intention During the Covid Pandemic 19. *Primanomics : Jurnal Ekonomi & Bisnis* 19(1):181. https://doi.org/10.31253/pe.v19i1.516