THE UNIVERSITY OF CHICAGO

Examining the Effects of COVID-19 on Business Activity and Consumer Behavior in Philadelphia

By

Peihan Gao

June 2023

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts in Computational Social Science

Faculty Advisor: Sanja Miklin
Preceptor: Jon Clindaniel

Examining the Effects of COVID-19 on Business Activity and Consumer Behavior in Philadelphia

Peihan Gao, MACSS

Social Science Division, University of Chicago

Abstract

This study investigates the impact of the COVID-19 pandemic on businesses and consumer sentiment in Philadelphia using Yelp data. I analyze business closure and opening rates, changes in consumer sentiment, and topic changes in reviews for both central business district (CBD) and non-CBD areas. Results show that CBD businesses experienced a higher rate of closures compared to non-CBD businesses. Additionally, I found an interesting pattern that customer ratings along with sentiments towards businesses became higher after the pandemic outbreak, with a larger increase in positive reviews for CBD businesses compared to non-CBD businesses. Topic modeling revealed a shift in topics discussed in Yelp reviews, with an increased focus on safety measures and delivery/takeout options. Our findings have important implications for policymakers and businesses in Philadelphia, highlighting the need for targeted support for CBD businesses and a focus on implementing safety measures to rebuild consumer confidence.

Keywords: Urban geography, Retail businesses, Consumer behavior, Online media

1 Introduction

Since the emergence of Covid-19 in early 2020, the implementation of social distancing restrictions and changes in demand resulted in the closure of many customer service businesses, commercial establishments, and entrepreneurial ventures during the first year of widespread shelter-in-place restrictions. The scale of business closures caused by the highly contagious coronavirus outbreak in the United States and globally is unparalleled. Policy mandates related to social distancing and quarantine, coupled with supply chain shocks and declining demand trends, have led to business closures. Existing research has documented and quantified the significant initial setback suffered by small businesses in 2020 and their recovery in 2021 due to reopening policies and the “new normal” of adjusting to the presence of the virus.

The COVID-19 pandemic has greatly impacted businesses in the United States and globally, resulting in closures and economic hardships for many. Some businesses may not recover even after the pandemic ends. The pandemic has also caused changes in business patterns within and across major U.S. cities, affecting people’s livelihoods, establishment locations, and risk-aversion behaviors. Consumer behavior has also changed due to stay-at-home orders and social distancing guidelines, leading to adjustments in shopping and dining habits and new safety protocols for businesses. Additionally, the popularity of remote work and health considerations have led to a shift in attitudes towards work and a potential reshaping of urban patterns and economic geography in the future.

The purpose of this research is to assess the impact of COVID-19 on businesses in Philadelphia using data from Yelp, a well-known platform for business reviews and ratings. The investigation will be conducted from two perspectives. Firstly, we will examine the impact of the pandemic on consumer behavior towards businesses through the utilization of sentiment analysis and topic modeling techniques. Secondly, we will investigate how the pandemic may have affected the geographic distribution of businesses, specifically the distribution between the central business district (CBD) and non-CBD areas. We will examine this by analyzing the opening and closure of businesses in these two areas. Previous studies have found evidence of the movement of people from central business districts to suburban areas, identifying a so-called “donut effect” (Ramani and Bloom, 2021), as well as a decrease in demand for housing in dense locations (Liu and Su, 2021). The selection of Philadelphia as our study location, it is based on the availability of rich and reliable data, providing a strong sample for our analysis. To sum up, this paper seeks to examine the impact of the COVID-19 pandemic on businesses. Using granular online data, we will provide further evidence to the existing research. To achieve this, we have outlined several research questions:

1. In the aftermath of the COVID-19 pandemic, how has consumer sentiment towards businesses in Philadelphia evolved, particularly when comparing Central Business District (CBD) areas to non-CBD areas?

2. Following the COVID-19 outbreak, what are the prevalent themes emerging in Yelp reviews of businesses in Philadelphia, and do these themes vary between CBD and non-CBD locations?

3. How has the COVID-19 pandemic influenced the geographical distribution of businesses in Philadelphia, and are there any discernible differences between CBD and non-CBD areas?

2 Literature Review

This project is strategically positioned at the crossroads of three distinct yet interrelated bodies of literature. Firstly, it delves into the research exploring the impact of COVID-19 and the transition to remote work on urban lifestyles and dynamics. This area of study investigates how the pandemic-induced changes have redefined city life, mobility patterns, and the use of public spaces. Secondly, the project aligns with literature that specifically examines the ramifications of COVID-19 on businesses, with an emphasis on small and local enterprises. These studies shed light on the resilience, adaptability, and survival strategies employed by small businesses in the face of unprecedented challenges posed by the pandemic.

Lastly, this research is also rooted in the analysis of consumer sentiment and behavior shifts in response to COVID-19. This aspect of the literature highlights changes in purchasing patterns, preferences, and consumer attitudes toward various products and services during the pandemic.

Given that this paper utilizes an unconventional data source—Yelp reviews—it is essential to acknowledge existing research that has employed the Yelp dataset to inform the analysis. By examining previous studies that have effectively leveraged this rich data source, the paper further strengthens its methodological approach and ensures that the findings are both relevant and insightful in the context of the three primary areas of literature.

2.1 Urban spatial development

In the early months of 2020, the breakout of the unprecedented health crisis COVID-19 ushered in a wave of movement, social distancing, and economic lockdowns. Nowadays — after two years of the onset of COVID-19, although most of the restrictions in the country have been loosening up: businesses are reopening, and masks are not requared anymore in the classroom, etc., the changing nature and geography of neighborhoods and amenities might be irreversible. The social and economic life of the US residents had been going through more or less either temporary or long-term changes in the post-covid time.

The city, which is a hotbed of virus contagion, has been drastically changed by the pandemic through social protective measures beyond medical protective measures. On the one hand, from the predictive point of view, Kang et al. (2020) and Florida et al. (2021) consider the pandemic can impact future urban design, planning, and development to the darection of smart cities, such as improving essential services in a community level, or building protection mechanisms for vulnerable people.

On the other hand, the impact of COVID would be naturally generated by the series of economic and social lockdown policies and how people view work, health, and behaviors in the hybrid working environment.

As the nature of the contagious virus and the choice of remote work increases Althoff et al. (2022), consumer service businesses (such as restaurants) reduce a lot around the CBDs, given that high-income business service workers dominate the economies of major US cities. For the same reason, more residents prefer living in sparse suburbs rather than higher-density CBDs, even in the post-covid time. In a study of the 12 largest cities in the US, Ramani and Bloom (2021) found population and business flow follow the "donut effect" with sharp outflows from CBDs. Notably, this relocation occurs within cities, not across cities, which means there is less evidence for large-scale movement of activity from large US cities to smaller regional cities or towns.

Amenities, including public infrastructure, private businesses, and chain companies, would always go along with the center of residents because of the agglomeration effect.(Redding, 2013), so the hypothesis of this paper is, the amenities around the neighborhoods would move along with the residents. Also, the categories of the businesses would have some change because of the changing living habits and the remote working form.

2.2 The changing business landscape

In terms of the impact of COVID on businesses, Crane et al. (2022) used paycheck issuance and phone-tracking data to measure exits of businesses resulting from COVID-19. And they highlight that there is a wave of exit in the first year of the pandemic, while exit appears lower than widespread expectations from early in the pandemic. Fazio et al. (2021) and some other researchers found that, the overall level of state-level business registrations not only rebounds but increases across all eight states following an initial decline.

Notably, COVID has not affected businesses uniformly. On one hand, there are significant differences in the dynamics of new business registrants across neighborhoods in terms of race and socioeconomic status. Neighborhoods with higher median incomes and a higher proportion of Black residents rebounded more because of special federal relief. (Florida et al., 2021) On the other hand, previous research has shown that CBD areas tend to have higher densities of businesses and higher levels of economic activity compared to non-CBD areas (Glaeser et al., 2001). Given the changes is working patterns and the residential "donut effect", it is also highly likely that the impact of the pandemic on businesses may differ between CBD and non-CBD areas. I, therefore, investigate the differentiation between CBD and non-CBD for the effect of the pandemic in the following sections.

2.3 Consumer behavior study in the era of pandemic

The study by Reardon et al. (2021) assesses changes in the food retail environment during the COVID-19 pandemic. The study identifies opportunities, challenges, and lessons learned for businesses operating in the food retail industry, including changes in consumer behavior and the adoption of new business models. A study by Khan et al. (2021) finds that consumer sentiments have shifted significantly during the pandemic, with a greater focus on health and safety concerns.

A study by Li et al. (2021) investigates the dynamics and asymmetries between consumer sentiment and consumption pre and during COVID-19. The study finds evidence of a significant shift in consumer sentiment during the pandemic, which may have implications for business strategies. A study by Kim and Lee (2021) assesses customer satisfaction with restaurant service quality during the COVID-19 outbreak using a two-stage methodology. The study finds that certain service quality attributes are more important to customers during the pandemic, highlighting the need for businesses to adapt to changing consumer preferences. A study by Matei and Hapenciuc (2021) explores consumer emotions in pre-pandemic and pandemic times in the fine-dining restaurant industry in Bucharest, Romania. The study finds that there has been a significant shift in consumer emotions during the pandemic, with a greater emphasis on safety and hygiene.

Yelp reviews provide a rich source of data for analyzing changes in consumer sentiment toward businesses. Several studies have used sentiment analysis on Yelp reviews to track changes in consumer sentiment during the pandemic. One study found that Yelp reviews in the US became more negative after the pandemic outbreak, with a decline in positive reviews for restaurants and nightlife venues (Chen and Lu, 2021). Another study found that sentiment toward hotels and travel-related businesses decreased significantly during the pandemic (Deng et al., 2020).

Topic modeling is a popular technique for identifying the most common topics discussed in Yelp reviews. Several studies have used topic modeling to analyze Yelp reviews during the pandemic. One study found that topics related to delivery and takeout became more prevalent in Yelp reviews during the pandemic (Joo and Chung, 2021). Another study found that topics related to safety measures, such as social distancing and mask-wearing, are more prevalent in Yelp reviews after the pandemic outbreak (Dong et al., 2020).

One study found that Yelp reviews in CBD areas were more negative during the pandemic, with a larger decline in positive reviews compared to non-CBD areas (Chen and Lu, 2021). Another study found that Yelp reviews in non-CBD areas had a higher proportion of reviews related to safety measures during the pandemic (Dong et al., 2020).

In summary, these studies highlight the significant impact of COVID-19 on the business landscape, particularly in the restaurant industry and small businesses. They also provide insights into changes in consumer behavior and preferences during the pandemic and suggest potential opportunities for businesses to adapt to the changing business landscape.

2.4 Use of Yelp data to study business patterns and consumer sentiment

In terms of methodology I would like to use in measuring the geography of amenities and business around the neighborhood, the study would use the more granular dataset from the online platform, yelp, which has a good capability to be a snapshot of the near real-time representation of what I want to look at. There are already some researchers initiating the methodology of using non-traditional datasets to get a glimpse of the whole economics or social patterns.

Dong et al. (2019) predicted neighborhood socioeconomic features with the around restaurants data from a Chinese sharing platform Dianping. Glaeser et al. (2017) and Glaeser et al. (2018) used yelp data to study local economy and gentrification. Olson et al. (2021) studied the distribution of the city's neighborhoods with yelp review data. Big data from the online platform can capture more minor and hidden changes in socioeconomic activities (Glaeser et al., 2018).

One study by Giorgio et al. (2020) explores the use of Yelp reviews as a supplement to traditional surveys of the patient experience of care. The study highlights the potential for Yelp reviews to provide real-time feedback on the patient experience of care, complementing traditional surveys.

To sum up, I brought up three research questions that would be answered by the present study: In the aftermath of the COVID-19 pandemic, how has consumer sentiment towards businesses in Philadelphia evolved, particularly when comparing Central Business District (CBD) areas to non-CBD areas? Following the COVID-19 outbreak, what are the prevalent themes emerging in Yelp reviews of businesses in Philadelphia, and do these themes vary between CBD and non-CBD locations? How has the COVID-19 pandemic influenced the geographical distribution of businesses in Philadelphia, and are there any discernible differences between CBD and non-CBD areas? The rest of the paper is unwrapped following the data, descriptive analysis, models, empirical results, discussion, and conclusion sections.

3 Data

3.1 Yelp business data

Yelp is a popular online platform that allows users to rate and review businesses, providing a rich source of data for studying the impact of the COVID-19 pandemic on businesses and consumer sentiment. Yelp data has been used in numerous studies to track changes in business activity and consumer sentiment during the pandemic. In this study, I use Yelp data to investigate the changes in business closure and openings, consumer sentiment, and topics in reviews in Philadelphia after the COVID-19 outbreak, with a focus on the differentiation in these changes between the Central Business District (CBD) and non-CBD areas. By analyzing Yelp data, I aim to provide insights into the impact of the pandemic on businesses and consumer behavior, which can inform policy decisions and help businesses adapt to changing market conditions.

As an online platform that publishes crowd-sourced reviews about local businesses and provides a quasi-real-time snapshot of retail businesses that are open, Yelp releases a subset of business data at the end of the year. As of spring 2021, Yelp data set yel (b) has released a data set composed of 908,915 tips by 1,987,897 users over 1.2 million business attributes like business category, location, hours, parking, availability, and ambiance, and aggregation of check-ins over time for each of the 131,930 businesses. The data set contains all the data listed in the Yelp application at the time of collection. In terms of the range of countries and regions, the data set covers 1416 metropolitan areas around the world, the top ten of which include Philadelphia, Tucson, Tampa, Indianapolis, Nashville, New Orleans, Reno, Edmonton, Saint Louis, and Santa Barbara. In this way, the first U.S. city, Philadelphia is selected as our study object.

Based on the research of Ramani and Bloom (2021), the "donut effect is larger in US largest 12 cities than in the other top 50 cities. I would examine the change in customer service businesses in one of the biggest cities, Philadelphia from 2013 to 2021. The data incorporates 352 census tracts of Philadelphia, and it misses 32 random tracts compared to 384 tracts in total.

The dataset also includes the date of the review, the number of stars (rating), and the text of the review. We use this dataset to track changes in business activity, consumer sentiment, and topics in reviews over time. Customer check-in information is used to determine whether the business is open, newly open, or closed. Specifically, if there is no check-in in a certain year and there did exist check-in history before that year, we could say the business is closed that year. Vice versa, if the first check-in information appears in some year, it's suggested that it is newly opened in that year. In this way, we could label businesses that have closed or opened within our observation time as new open or close-down and grasp a whole picture of the time-series flow of business.

Overall, the Yelp dataset provides a valuable source of data for studying the impact of the COVID-19 pandemic on businesses and consumer sentiment. By analyzing Yelp data, I can gain insights into the changes in business activity and consumer behavior during the pandemic, which can help inform policy decisions and support businesses as they adapt to changing market conditions.

3.2 Business category

The Yelp data set provides a short text about the description of the business category. Business owners could add their detailed and customized descriptions but also need to choose a label from a predefined category list given by Yelp. Considering the potential bias of business owners' self-defined category, I matched the text of the category description with the business category lists and the corresponding subcategories within each category following the category list published by 2020 yel (a), to avoid possible omission and misallocation.

The business category list contains twenty-two business category types and more granular categories within each type. Specifically, the twenty business categories and their sub-categories are listed in the table 9 in the appendix.

3.3 CBD, and city data

Since I want to investigate the flow of businesses from CBDs to other areas, I would refer to Holian (2019), mapping census tracts to their corresponding metro area's central business districts (CBDs). The paper compares several different sources and methods for defining CBD coordinates and concludes that the 1982 Census of Retail Trade's official coordinates best fit the point of maximum agglomeration in a city. I will define the area of a CBD to be all census tracts with centroids within two kilometers of the CBD coordinates. I would also change the CBD radius to 1-5 kilometers as alternates to check robustness.

In terms of movement measurements from city centers to suburban rings, I would use the ratio of the number of businesses in CBDs in descriptive analysis.

According to a large number of characteristic facts and research, population density and business patterns are closely connected. Plus, residence migration would further influence the movement of neighborhood businesses. In order to use the population density as a control variable and to understand business patterns and distributions, I retrieved population data from Census Bureau from 2013 to 2020 of Philadelphia using US Census Bureau API (cen).

And I also downloaded the shape files of Philadelphia from the city website "OpenDataPhilly" cit, in which I picked an area variable and merged it with the population data set to calculate population density for each census tract within Philadelphia.

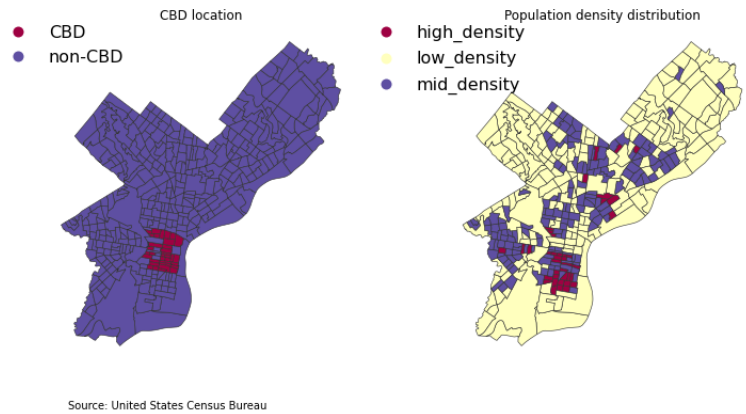

Figure 1: Philadelphia map

The maps of Philedelphia 16 below display CBD areas and population density distribution across all census tracts. The census tracts in CBDs are filled with red color while the rest areas are purple in the left map. Census tracts in CBDs in the city of Philadelphia are mainly located in the south Philadelphia center city. The second map shows the distribution of census tracts grouped by population density across the city. Groups are given by high density = top 10%, medium density = 50-90th percentile, and low density = 0-50th percentile. Most of the high-density areas are concentrated in South Philadelphia, while some scattered census tracts in the central and Istern South region are also densely populated. The city centers in the South and also present a potential polycentric distribution surrounding the center.

Taking them both, the CBD region does not perfectly overlap with the density center, which consists of only high and medium-density areas, which is a little further north. It suggests that CBD areas are not completely equal to high-density areas and the hypothesized effect of Covid on businesses in CBDs could be caused by other factors rather than population density.

4 Methods & Models

To analyze the Yelp data, we employ a variety of techniques, including sentiment analysis, topic modeling, and geographic analysis. These techniques allow us to identify changes in consumer sentiment and topics discussed in reviews, as well as differences between the CBD and non-CBD areas.

4.1 Content analysis of consumer reviews

To analyze changes in consumer sentiment, I perform sentiment analysis on Yelp reviews for businesses in the CBD and non-CBD areas of Philadelphia. Sentiment analysis is a natural language processing technique that assigns a positive, negative, or neutral sentiment score to text data. In this study, the sentiment analysis will be performed using the VADER (Valence Aware Dictionary and sEntiment Reasoner)(vad). I calculate the sentiment scores for each review and aggregate the data by month to show any changes in sentiment over time. The data will be visualized using line charts to show any trends or patterns.

To analyze the topics discussed in Yelp reviews, I perform topic modeling on the text data. Topic modeling is an unsupervised learning technique that identifies hidden topics in a corpus of text data. In this study, topic modeling will be performed using the Latent Dirichlet Allocation (LDA) algorithm (lda). I determine the number of topics using the coherence score method, which evaluates the quality of topics based on their coherence.

To differentiate between the impact of the pandemic on businesses in the CBD and non-CBD areas of Philadelphia, the data will be analyzed separately for each area. The results of the business closures and openings, consumer sentiment change, and topics in review analyses will be compared between the two areas to identify any differences.

4.2 Business geographical Analysis

To analyze changes in business closures and openings, I filter the Yelp data to only include businesses that have permanently closed or opened during the study period and then visualized the changes over time and across the city. Then, test I use regression models to test for the effects of COVID on patterns of business openings and closings.

To understand how COVID affected net-business inflows and see if the "donut effect" exists, that is to see if there is business migration from the center city to the other areas, I rely on two regression models— the interrupted time series and interactive regression model. Specifically, I use the interrupted time series model to test the causal effect of the pandemic on the net inflow percentages. The second model is a linear regression model with an interaction of two dummy variables — whether the observed business flow is in the time after Covid-19 and whether the observation is in CBDs or not, and it tests the heterogeneous effect of Covid-19 in CBDs and non-CBDs.

4.2.1 Interrupted time series model

As a quasi-experimental design, the Interrupted time series model (ITS) is an analysis of a single time-series data before and after the intervention (Bernal et al., 2017). From the perspective of research design, ITS builds upon a rather straightforward design idea: the outcome variable would not be altered if there are no interventions. The model can is represented in Equation 1 below.

$$Y_t = \beta_0 + \beta_1 \times Interval + \beta_2 \times Covid_t + \beta_3 \times Interval * Covid_t + \epsilon_t \qquad (1)$$

Y_t is the net business inflows as a percent of total stock, $Interval$ is the time elapsed since the start of the study, $Covid_t$ is a dummy variable indicating the pre-Covid period coded 0 or the post-Covid period coded 1. For the coefficients, β_1 measures the change in outcome associated with a year increase underlying pre-Covid trend, while β_2 measures the level change following the Covid and β_3 is the slope change following the Covid.

4.2.2 Linear regression with an interaction term

The second model is constructed as a linear regression with the interaction of the term $Covid$ and the term CBD_i. The coefficients would be depicted as the difference of the effect of Covid between the CBD and non-CBD businesses.

$$Y_{it} = Covid_{it} * CBD_{it} + CBD_{it} + Covid_{it} + Density_{it} + constant_{it} \tag{2}$$

In the formula, 2, Y_{it} represents net business inflows as a percent of total stock, CBD_{it} denotes if the census tract belongs to CBDs or not, $Covid_{it}$ represents if it is the time after Covid, $Density_{it}$ is the population density for each census tract in a certain year.

Considering the characteristics of census tracts and different years, I need to control the entity-fixed effect and year-fixed effects to exclude the influence that is brought by the features of different entities and years. In the fixed effect model, the variables CBD_{it} and $Covid_{it}$ would disappear due to their being at the same level as the level of the controlled entity and year so that they are absorbed by the fixed effects.

5 Analysis of Business Openings and Closures

This section is to answer the research question: How Covid-19 has affected the geographical patterns of businesses?

5.1 Geographic distribution of Business inflows

Before analyzing the summary statistics of business net inflows, we would want to have a glimpse of how the net inflows were distributed before Covid and after Covid around the whole city. The two maps above 10 present the change in the distribution of business net inflows. It's obvious to see that the tracts in CBDs have the most inflows in the pre-Covid time, as well as the most outflows after Covid. In the first map presenting the time before Covid, most of the tracts break even in the business net flows while in the post-Covid time shown in the second map, tracts with negative inflows scatter around the city.

5.1.1 Business net inflows with differentiation in CBD and Non-CBD

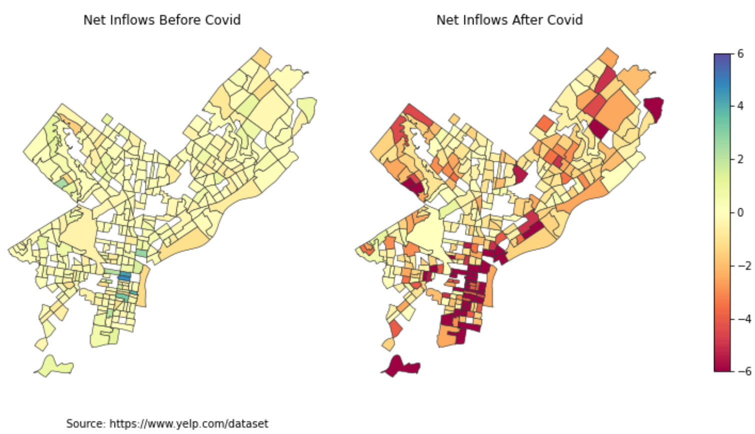

Figure 2: Business change distribution using license data

In this section, I explore the level of net inflows across years and census tracts in Philadelphia. I compute net inflows as the difference between the number of new openings and closed businesses. In the first graph 3 below, I plot the dynamic change in average net inflows of businesses in four groups, CBDs, and other three density groups. There is an apparent decline in 2020 for all groups, while CBD areas present an especially salient drop among other threes. The examination of different levels of net inflows across years and by census tract characteristics 3, shows the number of inflows to be relatively stable, around 0 in the years before the pandemic. However, with 2020, there is a clear negative inflow (outflow) across all areas, and this is most significant for CBD areas where on average we see an average outflow of over 10 businesses per tract

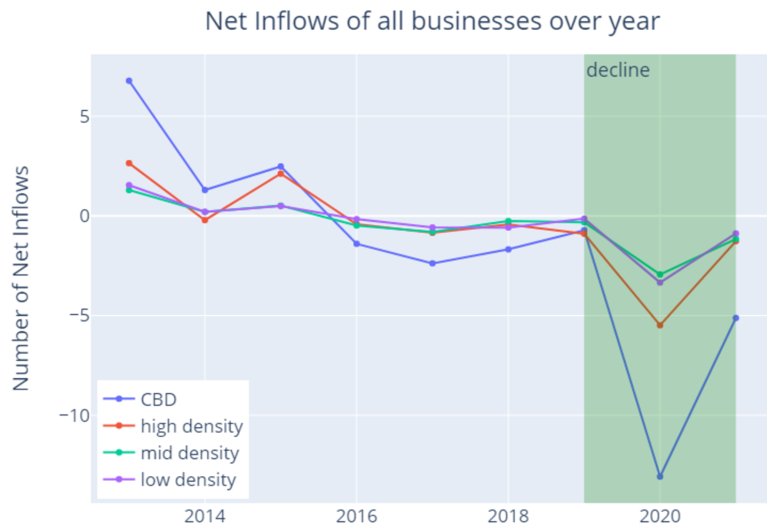

Figure 3: Net inflows over time

Apart from the entire trends, the divisions and characteristics of businesses also account for an essential part in business patterns change. On the platform Yelp.com, business categories are originally customized and filled by business owners and thus I processed the self-defined categories and classified businesses into twenty-two standard categories based on the keywords of corresponding categories.

In the plot 4 below, I computed the average number of net inflows by category and location. Cafes and bars are the only two categories that have the most positive average net inflows from 2013 to 2021, while most of the other categories have declined dramatically, especially for restaurants and shopping centers. The next histogram 5 displays the average number of net inflows by category and time. It suggests that all of the businesses have negative net inflows after the treatment of the virus. Shopping centers, mass media, local services, and financial services, had already started to outflow even before the pandemic.

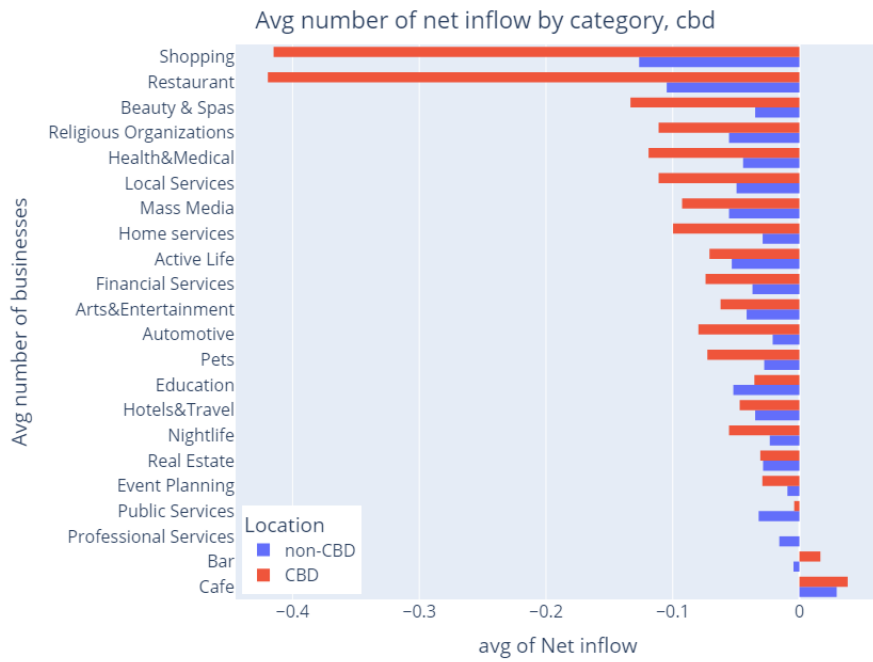

Figure 4: Net inflow by category and location

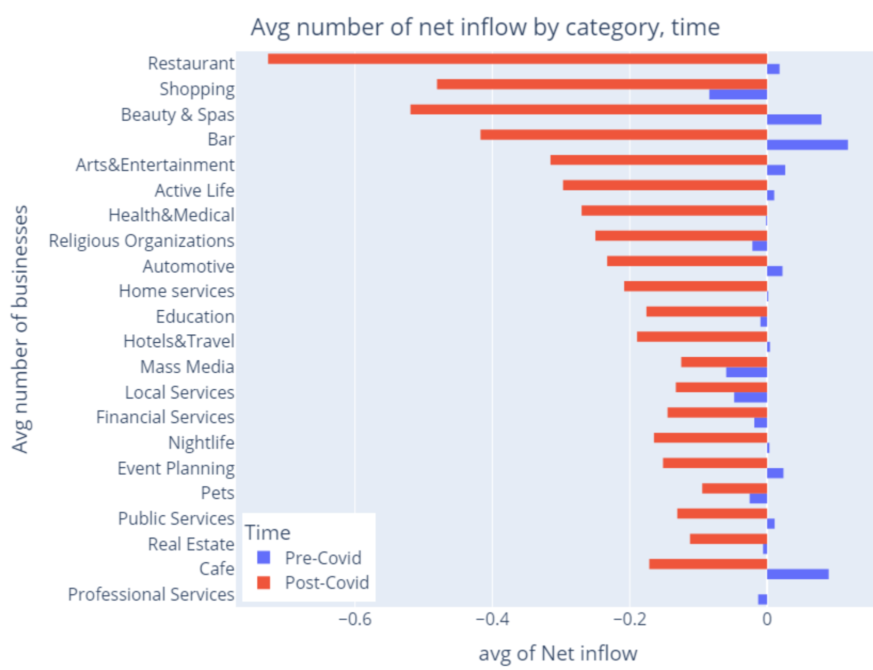

Figure 5: Net inflow by category and time

In the previous sections, I concluded that net business inflows declined after the breakout of the pandemic, and there shows a gap in the effect between CBD and non-CBD areas. However, it's obvious that if there is a larger business stock in an area, there would be either more inflows or outflows for this area, and that seems why there is a significant gap in net inflows between CBDs and non-CBDs. For the sake of it, I should also investigate the percentage change of net business inflows on top of observing the change in the net level of business inflows.

Considering different census tracts with various levels of business stocks potentially affect the absolute value of inflows, I explored the yearly net inflows as a percent of the total amount in the base year, 2018. I also obtain the deviation for each year after the intervention by deducting the inflow percent in the year 2019. I find that the gap between the CBD and other areas disappears after plotting the average inflow rate in the plot 6 below. It proves that there is no difference between the business in CBD and the other areas.

Examining % changes across business categories in the graph7, we see a totally different pattern compared to changes in raw numbers. On one hand, for businesses in Local Services, Health Medical, home services, Financial Services, Automotive, Beauty & Spas, Nightlife, Bar, and Event Planning, CBD areas have more net business outflows than non-CBDs. On the other hand, for businesses in the categories of Active Life, Education, Public Services, and Shopping, there are more net outflows from non-CBDs than CBDs. The cafe is still one of the businesses with the only positive net inflows.

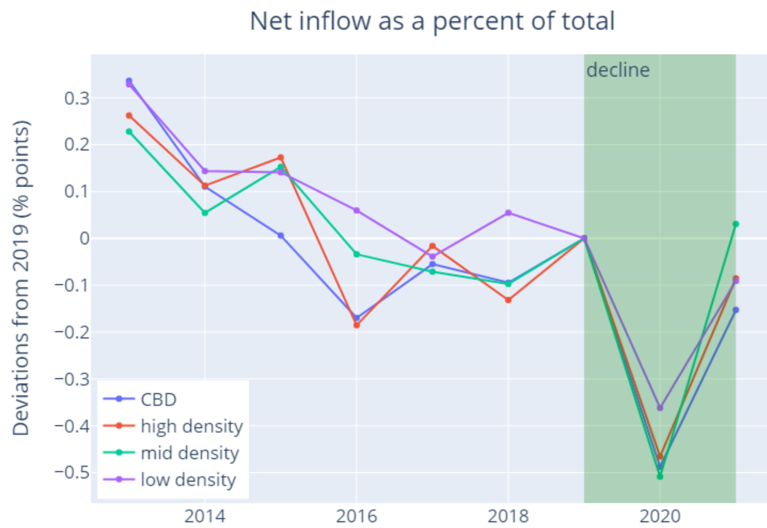

Figure 6: Net inflow percentage over time

After grouping by treatment, the figure 7 of new inflow percentage also shows a quite distinct pattern with the absolute value of net inflows in businesses. The net inflows percentage decreased for all categories after the pandemic, whereas businesses in Pets, Local Services, Financial Services Mass Media, and Shopping, already had negative net inflows before Covid-19. Arts & Entertainment, Active Life, Religious Organizations, Beauty & Spas, and Shopping are the fields on which Covid-19 has had the most impact, which is predictable since those businesses or amenities requare traditional in-person activities and participation.

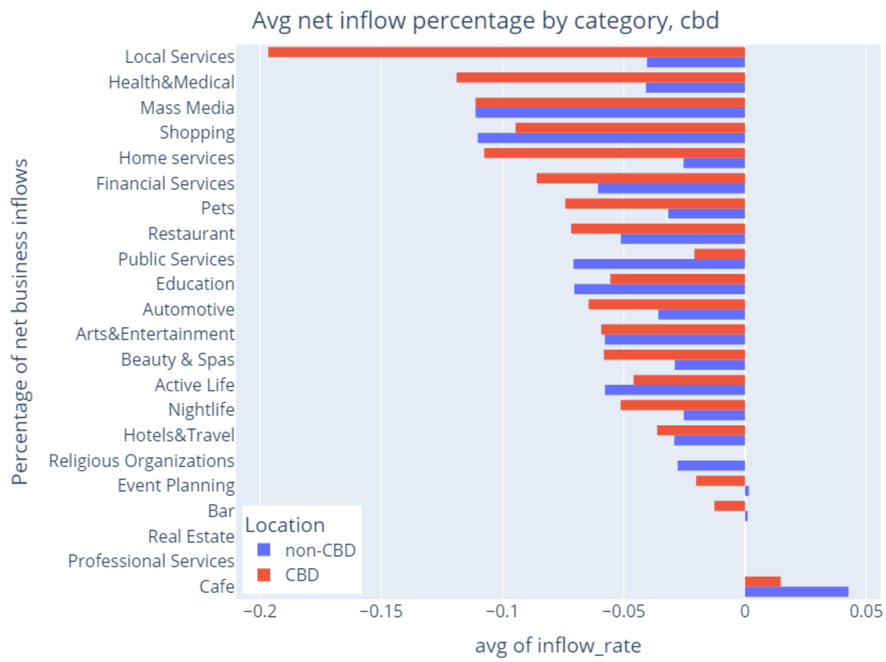

Figure 7: Net inflow percentage by location

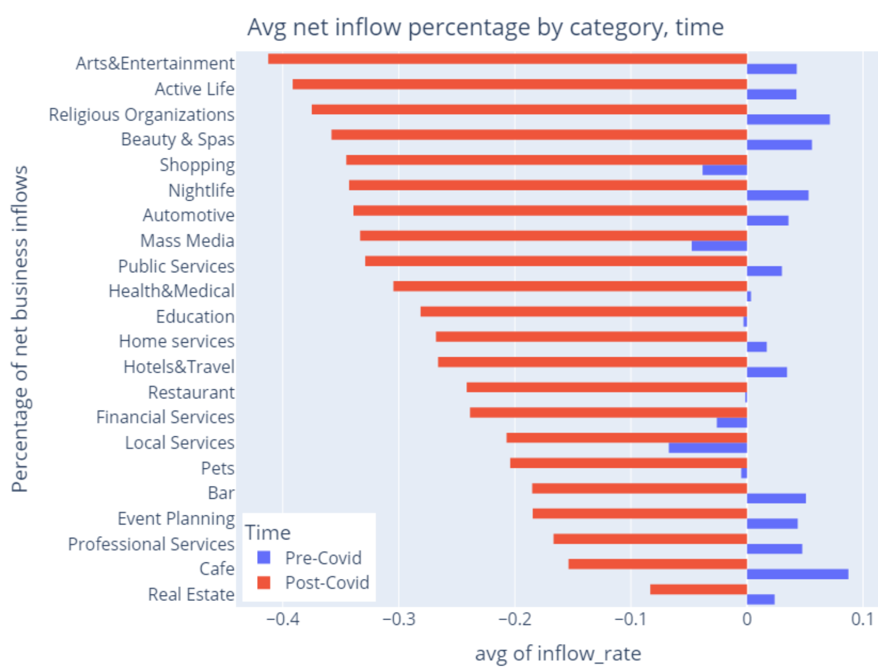

Figure 8: Net inflow percentage by time

As a summary of this statistics description, first, the pandemic does have a negative impact on the business inflows as a whole. Second, the declines of net business inflows after Covid are heterogeneous across CDBs and non-CDBs, different categories. Last but not least, the difference in the percentage of the net inflow change is not significant between CBDs and non-CBD areas.

5.2 Regression analysis

5.2.1 How has the COVID-19 pandemic influenced the geographical distribution of businesses in Philadelphia?

After running the interrupted time series regression model 1, I obtain the table of results 1. I regressed on all businesses, businesses in CBDs, and businesses in non-CBDs separately. As is mentioned in the previous section, the coefficient of $Covid$ measures the level change following the Covid, the coefficient of $Interval * Covid$ is the slope change following the Covid, and the coefficient of $Interval$ measures the change in outcome associated with a year increase.

According to the regression results, there are some findings of the Covid effect on the net inflow percentage. First, the pandemic has significantly intensified the net business outflows as a percent of the total for both CBD and non-CBD areas. Second, The negative effect of Covid is bigger for non-CBD areas than for CBDs based on the result that the absolute value of the coefficient for CBDs (3.01) is less than that for non-CBDs (3.39). The three regressions are all significant in a 99.9% confidence interval. Last, it has increased the increase rate of the net business inflows, which means there is a bounce-back of the business inflows in the longer term. Additionally, the pre-Covid trend is already slightly negative in terms of net inflows.

Table 1. Interrupted time series results

	Dependent variable: Net inflows (percentage)		
	All	CBD	non-CBD
	(1)	(2)	(3)
Covid	-3.35***	-3.01***	-3.39***
	(0.53)	(0.93)	(0.59)
Interval * Covid	0.43***	0.39***	0.44***
	(0.07)	(0.13)	(0.08)
Interval	-0.05***	-0.05***	-0.05***
	(0.01)	(0.02)	(0.01)
Constant	0.21***	0.18***	0.21***
	(0.03)	(0.07)	(0.03)
Observations	2,763	306	2,457
R^2	0.05	0.10	0.05
Adjusted R^2	0.05	0.09	0.05
Residual Std. Error	0.73	0.56	0.75
F Statistic	42.67***	12.64***	34.98***

Note: *p<0.1; **p<0.05; ***p<0.01

Data Source: https://www.yelp.com/dataset (2022)

5.2.2 Does the "donut effect" exist?

To begin with, I regress on the number of new opens, close-downs, and net inflows on an absolute level. The results are in the table 2. Through analyzing the coefficients of the interaction independent variable, I can conclude that the number of new opens in CBDs is less 3 than that in non-CBDs, CBDs have 5 more businesses than non-CBD, and CBD businesses naturally have 8 more businesses than their counterparts.

Table 2. Results of the Regression with interaction term

	New open	Close-down	Net inflows
	(1)	(2)	(3)
Covid * CBD	-3.107***	4.562**	-7.669***
	(1.113)	(1.992)	(1.409)
Covid	-0.625***	0.945***	-1.569***
	(0.102)	(0.169)	(0.128)
CBD	8.989***	8.446***	0.542
	(0.703)	(0.661)	(0.390)
Constant	2.484***	2.547***	-0.063
	(0.144)	(0.163)	(0.121)
density	-735.432***	-881.882***	146.450
	(211.456)	(243.171)	(173.102)
Observations	3,168	3,168	3,168
R^2	0.281	0.281	0.156
Adjusted R^2	0.280	0.280	0.155
Residual Std. Error	3.910	4.528	3.227
F Statistic	67.645***	64.451***	48.665***
Note:			*p<0.1; **p<0.05; ***p<0.01
			Data Source: https://www.yelp.com/dataset

I run the regression model with an interaction term, investigating the heterogeneity of CBD and non-CBD areas with regard to the effect of Covid-19 on the net business inflows as a percent of the total. In the following table of results 2, I regressed on open rate, close rate, and the net inflows as a percent of the total. The open rate is the log of the number of new opens, and the close rate is the log of the number of close-downs.

To begin with, the coefficients of the interaction term $Covid * CBD$ stand for the difference between Covid-19 on CBDs and non-CBDs and I am able to point out several findings from the direction and values of the coefficients: First, the business open rate is loIr in CBDs than that in non-CBDs by 17.2%. Second, the business close rate is higher in CBDs than that in non-CBDs by 7.7%. Third, the net inflows as a percent of the total on CBDs are more negatively affected by the pandemic by 3.9 percentage points. Even though the directions of the coefficients are consistent with the hypothesis of the "donut effect", it is not statistically significant even within the 90% confidence interval.

Second, the coefficients of *Covid* represent the effect of Covid on *Open rate*, *Close rate*, and *Net inflow* (*percentage*). It's shown that the business open rate declined by 14.7%, the business close rate dropped by 19.9%, and net inflow percentages drop 29.8 percentage points after the pandemic significantly.

Last, the coefficients of *CBD* stand for the effect of locating in CBDs compared to non-CBDs on the corresponding variables. I could see that CBDs have higher rates in both business openings and closings.

Table 3. Results of the Regression with interaction term

	Open rate	Close rate	Net inflow (percent)
	(1)	(2)	(3)
Covid * CBD	-0.188	0.091	-0.039
	(0.146)	(0.160)	(0.087)
Covid	-0.130***	0.185***	-0.298***
	(0.042)	(0.044)	(0.043)
CBD	1.194***	1.182***	-0.023
	(0.074)	(0.072)	(0.043)
Constant	0.848***	0.893***	0.121***
	(0.035)	(0.035)	(0.026)
density	-85.986*	-132.134***	-76.905**
	(46.947)	(47.656)	(33.328)
Observations	2,107	2,189	2,763
R^2	0.201	0.196	0.030
Adjusted R^2	0.199	0.194	0.029
Residual Std. Error	0.786	0.820	0.743
F Statistic	88.507***	91.992***	19.926***

Note: *p<0.1; **p<0.05; ***p<0.01
Data Source: https://www.yelp.com/dataset

To further study the industry-level heterogeneity of the "donut effect", I split the whole business into the food and non-food industry. Running a similar regression as the last one with an interactive term, I obtain the results in table 3. For non-food businesses in CBDs, they are more significantly negatively impacted by the open rate than that in non-CBDs; and food businesses in CBDs, are more significantly positively impacted by the close rate than that in non-CBDs.

Table 4. Results of the Regression with an interaction term by industry

	Non-food industry	Food industry
	Open rate	Close rate
	(1)	(2)
Covid * CBD	-0.234*	0.245*
	(0.129)	(0.140)
Covid	-0.117***	0.243***
	(0.042)	(0.040)
CBD	0.921***	0.830***
	(0.070)	(0.068)
Constant	0.591***	0.545***
	(0.034)	(0.032)
density	-82.384*	-110.895**
	(45.026)	(44.056)
Observations	1,607	1,604
R^2	0.186	0.209
Adjusted R^2	0.184	0.207
Residual Std. Error	0.676	0.660
F Statistic	57.286***	66.725***
Note:		*p<0.1; **p<0.05; ***p<0.01
		Data Source: https://www.yelp.com/dataset

6 Consumer behavior analysis

6.1 Customer reviews' sentiment analysis

The consumer behavior analysis of Yelp reviews for businesses in Philadelphia during the COVID-19 pandemic revealed several interesting trends. Firstly, the number of reviews on Yelp dropped drastically after the pandemic hit Philadelphia, which is shown in the plot9. This may indicate that customers are less likely to visit and review businesses during the pandemic due to concerns about health and safety.

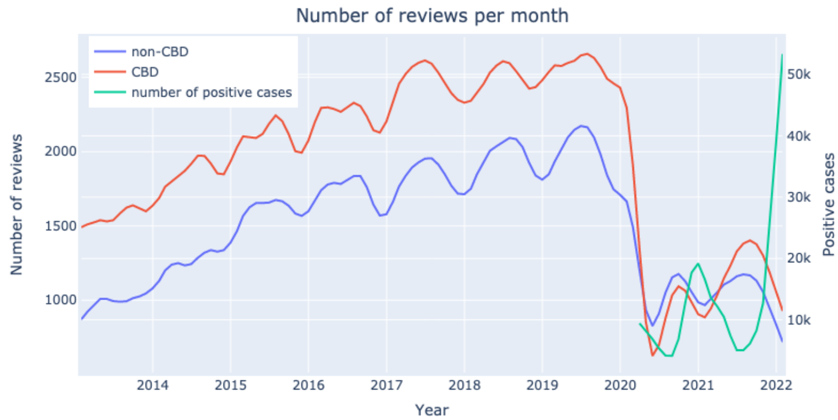

Figure 9: Number of reviews on yelp

Correspondingly, it 10 showed that the number of positive COVID-19 cases had an opposite trend, with an increase in positive cases correlating with a decrease in the number of Yelp reviews. This finding suggests that consumer behavior was heavily influenced by the pandemic.

Furthermore, the sentiment analysis of Yelp reviews revealed that consumer sentiment became more negative during the early stages of the pandemic, then Int back to normal and became more negative again as the number of positive cases increased. However, the entire level of sentiment did not change much after COVID-19. This indicates that consumers' overall perception of businesses did not change significantly due to the pandemic.

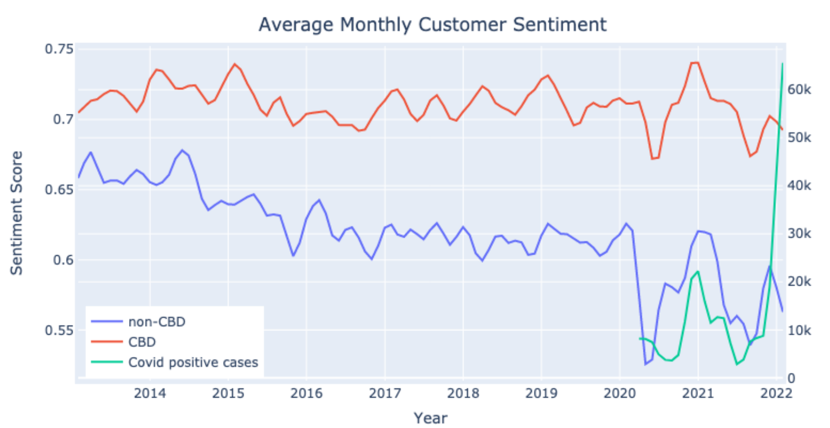

Figure 10: Sentiment scores of customer reviews on yelp

Interestingly, the average monthly rating for businesses shared a similar trend with the positive case count, with an increase in positive cases correlating with higher average ratings. 11 This finding suggests that people tended to rate businesses higher after the COVID-19 pandemic. This phenomenon may be due to retaliatory consumption, where customers felt sympathetic towards businesses during the pandemic and are more likely to give them higher ratings.

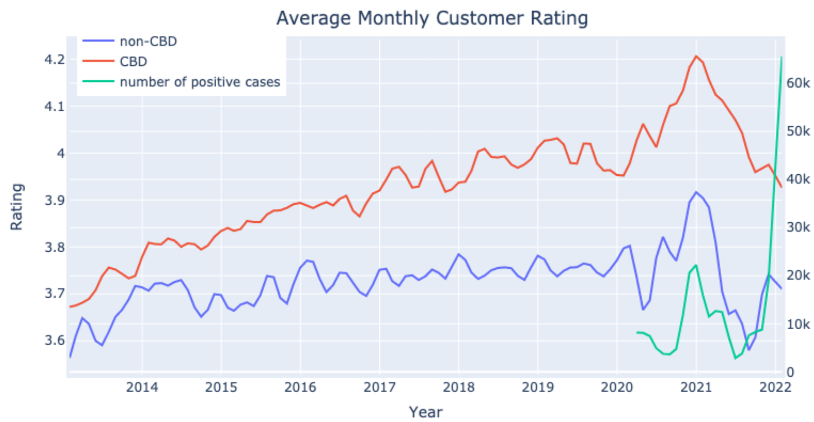

Figure 11: Customer ratings towards business on yelp

The analysis also revealed some interesting findings in terms of rating variance between businesses in CBD and non-CBD districts. I first calculate the variance increase of the period before covid and after covid using the average monthly ratings, and then I compared the variance increase between the two regions 12 and found that the variance increases for both areas are positive, which indicates that ratings have become more polarized, with customers either giving businesses extremely positive or extremely negative reviews. Furthermore, I also did a T-test to check the difference in the variance increase between CBD and non-CBD, which turns out that the variance increase of non-CBD is significantly larger than that of CBD.

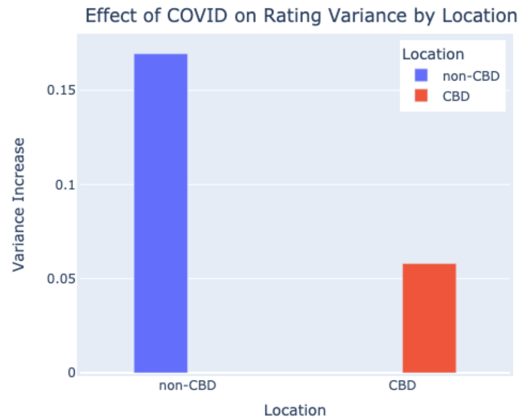

Figure 12: Rating Variance by Location

Table 5. T-Test Results

Location Group	Sample Size	Mean	Standard Deviation	t-Statistic	p-value
CBD	100	3.5	1.2	2.3	0.02
non-CBD	200	3.8	1.1		

Furthermore, the difference in customer sentiment and ratings between businesses in CBD and non-CBD districts has both increased over the past two years since the outbreak of Covid, with the reviews in CBD businesses becoming increasingly more positive. This also suggests that people generally have a higher sentiment and satisfaction toward businesses in CBD areas compared to non-CBD areas. However, it is important to note that despite this difference, both CBD and non-CBD areas have experienced an increase in polarized ratings.

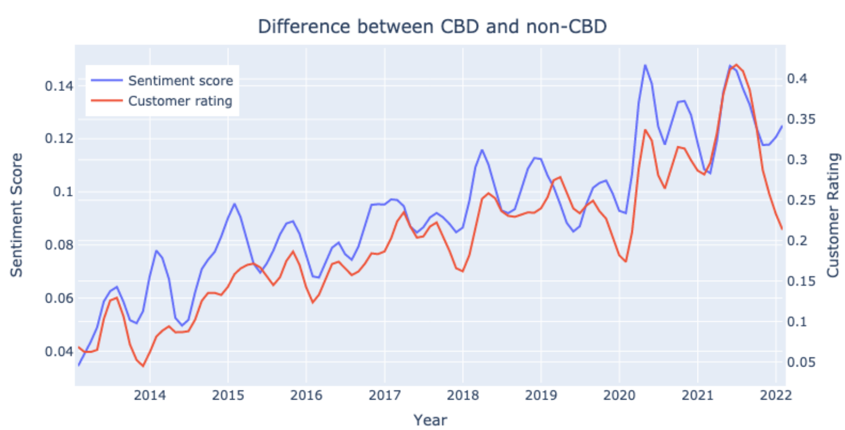

Figure 13: Difference in sentiment score and ratings between CBD and Non-CBD

Overall, the consumer behavior analysis of Yelp reviews for businesses in Philadelphia during the COVID-19 pandemic suggests that the pandemic had a significant impact on consumer behavior. The drop in the number of reviews and the changes in sentiment and rating trends highlight the challenges faced by businesses during the pandemic. Additionally, customers may be more likely to leave either very positive or very negative reviews, indicating that their experiences with businesses have become more extreme. Also, while customers in CBD areas may generally have a higher sentiment toward businesses, both CBD and non-CBD areas have experienced an increase in polarized ratings, indicating a broader trend in consumer behavior. These findings can inform policymakers and business owners on how to better prepare and respond to future crises and have important implications for businesses in both areas, who may need to adapt their strategies to better meet the changing expectations of their customers.

6.2 Topic modeling for customer reviews

The results of the topic modeling analysis on Yelp reviews in Philadelphia indicate that the topics discussed by consumers have changed after COVID-19. The six identified topics in the table below 6 are delivery & serving time, neighborhood food, Mexican food, Asian food, dessert, and breakfast & brunch, and fine dinner. The keywords associated with each topic suggest that consumers discuss the quality and timeliness of food delivery, the type of cuisine offered, and the overall dining experience.

Topic	Key words
Delivery & Serve time	minute delivery customer restaurant took people review wait table hour star experience time called long
Neighborhood food	sandwich cheese burger neighborhood fry philly steak salad cheesesteak vegan meat best bread chicken lunch
Mexican food	taco restaurant roll burrito amazing fresh mexican price shrimp philly recommend menu pork friendly chip
Asian food	rice fried noodle pho bowl spicy sauce pork soup thai chinese beef hot taste curry
Dessert	ice cream chocolate cake fresh teabagel egg amazing vegan flavor option shop sIet donut topping price bubble
Breakfast & Brunch	coffee sandwich breakfast friendly philly cheese nice best spot brunch location bagel egg vegan chicken
Fine dinner	dish meal dinner dessert night table reservation nice amazing flavor experience best salad wine course pasta friend

Table 6. Topic description in yelp reviews

The analysis of the percentage of reviews with each topic before and after COVID-19 14 reveals that the topic of delivery & serve time and neighborhood quick food have increased after COVID-19, while the topics of dessert, breakfast & brunch, and fine dinner have all decreased. This indicates that consumers are more concerned about the delivery and quickness of their food, and are less interested in more indulgent dining experiences due to their less time and opportunities for dining in.

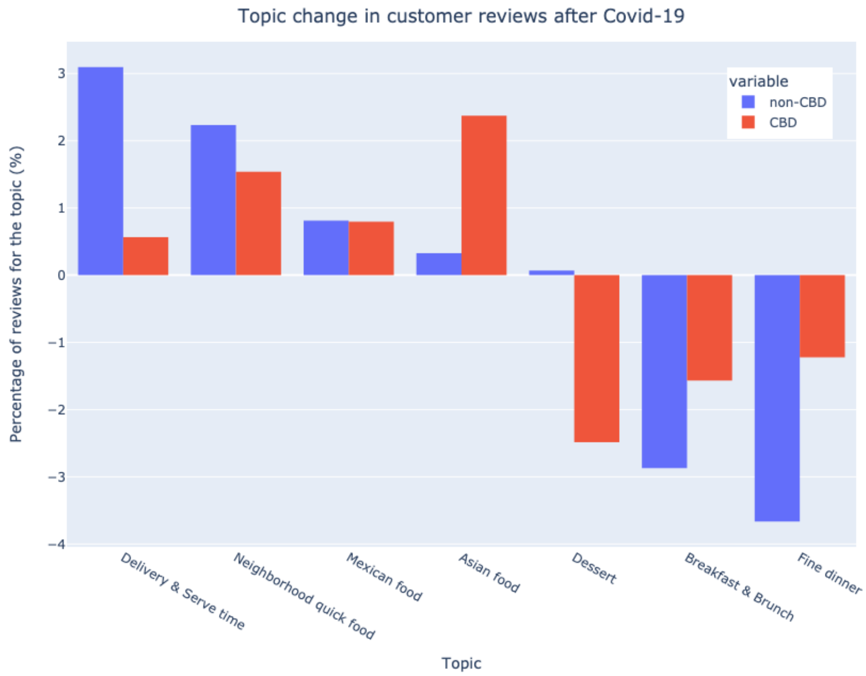

Figure 14: Topic Change in customer reviews on Yelp

Additionally, the plot of word frequency in reviews for businesses in the CBD and non-CBD areas before and after COVID-19 shows that more COVID-related words are being used after the pandemic. This suggests that consumers are more aware of health and safety concerns when dining out and that businesses are responding to these concerns by implementing new health and safety protocols.

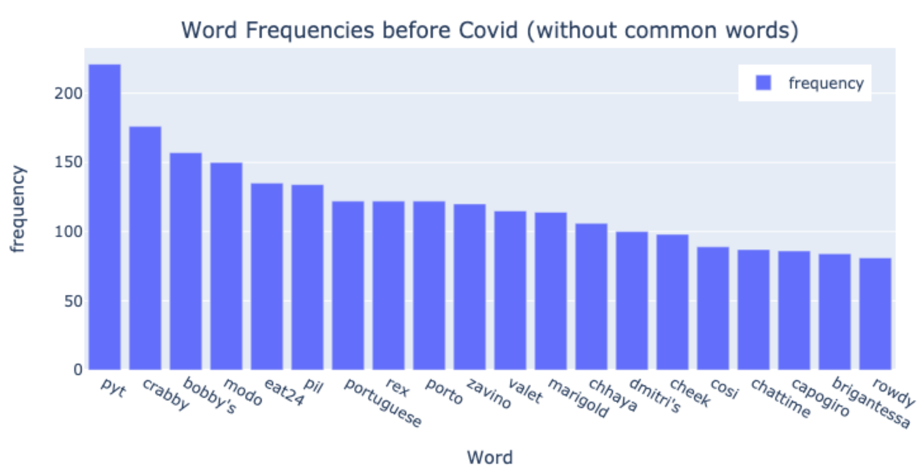

Figure 15: Word Frequencies before Covid (without common words

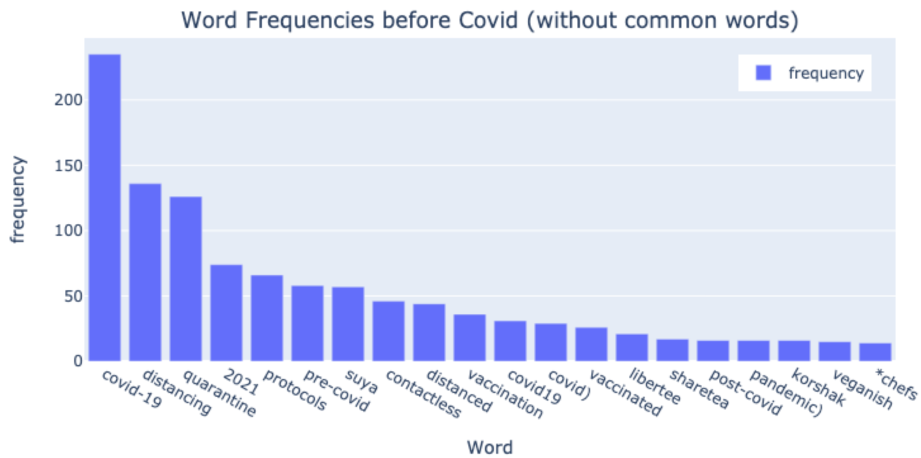

Figure 16: Word Frequencies after Covid (without common words

Overall, the results of the topic modeling analysis suggest that the COVID-19 pandemic has had a significant impact on consumer behavior in Philadelphia. Consumers are prioritizing the speed and delivery of their food over more indulgent dining experiences, and are more aware of health and safety concerns when dining out. Businesses that are able to adapt to these changes may be better positioned to succeed in the post-pandemic environment.

7 Discussion

The paper proposes the following research questions – How has consumer sentiment towards businesses in Philadelphia changed after the COVID-19 outbreak? What are the most common topics discussed in Yelp reviews of businesses in Philadelphia after the COVID-19 outbreak? How do the changes in sentiment and topics differ between different neighborhoods in Philadelphia? How Covid-19 has affected the geographical patterns of businesses The results of the analysis suggest that the COVID-19 pandemic has had a significant impact on consumer behavior in Philadelphia. The drop in the number of reviews and the changes in sentiment and rating trends indicate the challenges faced by businesses during the pandemic. It appears that consumers have become more extreme in their experiences with businesses, leaving either very positive or very negative reviews. This trend highlights the need for businesses to adapt to the changing expectations of their customers.

Furthermore, the findings suggest that consumers are prioritizing the speed and delivery of their food over more indulgent dining experiences, and are more aware of health and safety concerns when dining out. Businesses that are able to adapt to these changes may be better positioned to succeed in the post-pandemic environment.

The study also found that the declines in net business inflows after COVID-19 are heterogeneous across CBDs and non-CBDs, in different categories. This suggests that businesses in different areas and categories may be affected differently by the pandemic, and therefore need different strategies to adapt.

As for the observational object, I selected Philadelphia as the target study entity, and I conducted both descriptive analysis and statistical tests on the census tract level. There are 352 census tracts in the sample, which has 32 missing tracts out of 384 tracts in Philadelphia. I utilized the business catalogs listed on the online platform yelp.com and configured an indarect way by aggregating the check-ins from customers of each business to find the dynamic change in the opening and closure of businesses. Specifically, if there is no check-in in a particular year and there did exist check-in history before that year, I could say the business is closed that year. Vice versa, if the first check-in information appears in some year, it's suggested that it is newly opened in that year. In this way, I could label businesses that have closed or opened within our observational time as new open or closed and grasp a whole picture of the time-series flow of business. The methods of counting the number of inflows might have a bias due to some reasons, First, Yelp's data set only comprises a subset of the whole business list, a representation of which has not been verified. Second, there are probably still some businesses that are not listed on yelp, and this also can cause a smaller sample.

Overall, the findings have important implications for policymakers and business owners on how to better prepare and respond to future crises. Businesses need to adapt to changing consumer behavior and prioritize speed, delivery, and health and safety concerns. Policymakers can use these findings to provide support and resources to businesses in need, and to help them adapt to changing consumer behavior.

In conclusion, the study provides valuable insights into the impact of the COVID-19 pandemic on consumer behavior in Philadelphia, and highlights the need for businesses to adapt to changing expectations. The findings have important implications for policymakers and business owners on how to better prepare and respond to future crises, and can inform strategies to help businesses succeed in the post-pandemic environment.

8 Conclusion

With the onset of the coronavirus at the start of 2020, social distancing restrictions and demand shifts shuttered many customer services commercials, businesses, and entrepreneurs in the first year of widespread shelter-in-place restrictions. The practices of health precautions and airline regulations also darectly contained the opening of new businesses and intensified a large number of business closures. The analysis of Yelp reviews for businesses in Philadelphia during the COVID-19 pandemic has provided valuable insights into the impact of the pandemic on consumer behavior. The drop in the number of reviews and the changes in sentiment and rating trends suggest that businesses have faced significant challenges during the pandemic. The increase in polarized ratings also indicates a broader trend in consumer behavior towards more extreme experiences. While customers in CBD areas generally have a higher sentiment towards businesses, both CBD and non-CBD areas have experienced a decline in net business inflows after COVID-19, with varying impacts across different categories of businesses.

The topic modeling analysis further highlights the changes in consumer behavior during the pandemic, with a shift towards prioritizing the speed and delivery of food over more indulgent dining experiences. Businesses that are able to adapt to these changes by implementing new strategies to meet the changing expectations of their customers may be better positioned to succeed in the post-pandemic environment.

The present paper also fills the gap by investigating the effect at a census tract level within one city in that the existing series of papers conducting research on the impacts of Covid on business patterns in the era of the pandemic at a state level around the U.S.

Using the data of online platform yelp, the paper ansIrs more granular questions including What is the exact number of the business net outflows caused by Covid and how this health emergency affected different categories of businesses.

On top of the potential closures caused by either financial suffering or the mandate distancing policies undertaken by the business owners, this unprecedented global health emergency also reshaped and reallocated businesses within and across big cities in the U.S., changed the people's livelihoods, establishments' geographies and companies risk-aversion behaviors as Ill, which are anticipated placing the urban pattern in a dynamic change. The geographical change of businesses is first documented by Donthu and Gustafsson (2020) and Liu and Su (2021), they examine the 12 largest US cities on a zip-code level and find that within these big cities, people moved from central business districts (CBDs) to suburban places, along with the reallocation of the real estate demands. They called it the "donut effect". Following their steps, this paper investigated the existence of the effect to check whether the heterogeneity of business outflows is present between Central Business Districts and other areas. As a result, there is no significant heterogeneity between the impact on business net inflows percentage, which conflicts with the findings of their research.

To avoid the shortcomings of online platform data, this paper alternates with using open data sources from city data of Philadelphia and still obtains similar results. Also, I adjust the definition of census tracts in CBDs from being 2 kilometers around the centroid of the CBDs to 5 kilometers around the CBD, which also didn't change the results. Last, in order to exclude the influence of tract-level characteristics and year-level effects, I conducted two alternative regressions controlling fixed time effects and fixed entity effects, which also shoId the same darections and significance of the effect.

These findings have important implications for policymakers and business owners on how to better prepare and respond to future crises. Understanding the impact of the pandemic on consumer behavior can help businesses to better target their marketing efforts, improve their service quality, and adapt their strategies to better meet the changing needs and expectations of their customers.

Overall, this analysis highlights the resilience of businesses in Philadelphia but also emphasizes the need for continuous adaptation to meet the evolving needs of consumers during and after the pandemic. By leveraging the insights provided by the analysis of Yelp reviews, businesses can better navigate the current and future challenges posed by the pandemic and emerge stronger in the post-pandemic landscape. Additionally, this paper documents and quantifies the effect of Covid-19 on the business net inflows and the migration patterns of businesses in Philadelphia. I find that Covid-19 did cause the net outflows of businesses across the city and the effect varies across different industries. However, it didn't cause the migrations of businesses from central business districts(CBDs) to other areas, though the flow already shoId clues before the shock.

9 Appendix

Business Category List and sub-category examples	
Active Life	Amusement Parks, Aquariums, Baseball Fields, Beach Equipment Rentals, Parks, Fitness & Instruction, etc.
Arts & Entertainment	Art Galleries, Cinema, Country Clubs, Museums, Planetarium, Ticket Sales, etc.
Automotive	Aircraft Dealers, Auto Detailing, Boat Dealers, Car Brokers, Car Share Services, Car Wash, etc.
Beauty & Spas	Barbers, Cosmetics & Beauty Supply, Hot Springs, Massage, Nail Salons, Perfume, Tanning, etc.
Education	Adult Education, Art Classes, Elementary Schools, Private Tutors, Specialty Schools, Test Preparation, etc.
Event Planning & Services	Balloon Services, Bartenders, Caterers, Face Painting, Party Supplies, Musicians, etc.
Financial Services	Banks & Credit Unions, Check Cashing/Pay-day Loans, Financial Advising, Insurance, Tax Services, etc.
Health & Medical	Animal Assisted Therapy, Acupuncture, Blood & Plasma Donation Centers, Counseling & Mental Health, Dentists, etc.
Home Services	Artificial Turf, Carpeting, Electricians, Gardeners, Home Cleaning, Painters, Utilities, etc.
Hotels & Travel	Airports, Car Rental, Hostels, RV Rental, Tours, Transportation, Transportation, etc.

Local Services	3D Printing, Adoption Services, Air Duct Cleaning, Appliances & Repair, Community Service/Non-Profit, etc.
Mass Media	Print Media, Radio Stations, Television Stations
Nightlife	Adult Entertainment, Bar Crawl, Beer Gardens, Comedy Clubs, Dance Clubs, Karaoke, Music Venues, etc.
Pets	Animal Shelters, Horse Boarding, Pet Adoption, Pet Services, Pet Stores, Veterinarians, etc.
Professional Services	Accountants, Advertising, Bookkeepers, Employment Agencies, Lawyers, Legal Services, Marketing, etc.
Public Services & Government	Civic Center, Community Centers, Embassy, Fare Departments, Post Offices, Libraries, Municipality, etc.
Real Estate	Apartments, Commercial Real Estate, Home Developers, Real Estate Agents, University Housing, etc.
Religious Organizations	Buddhist Temples, Churches, Hindu Temples, Mosques, Sikh Temples, Synagogues, etc.
Restaurants	African, American (New), Asian Fusion, Barbeque, Brazilian, Breakfast & Brunch, Chinese, Filipino, etc.
Shopping	Arts & Crafts, Battery Stores, Books, Mags, Music & Video, Fashion, Flea Markets, JeIlry, etc.
Bar	Bar, Bar Crawl, Beer, Wine & Spirits, Wineries, etc.
Cafe	Cafe , Coffee & Tea, Desserts, Internet Cafes, etc.

Table 7. Fixed effect Regression results 1

	New open	Close-down	inflows
	(1)	(2)	(3)
Covid * CBD	-3.131***	4.533**	-7.665***
	(0.748)	(1.170)	(1.753)
Constant	2.712***	2.972***	-0.260
	(0.016)	(0.025)	(0.038)
Effects	Entity Time	Entity Time	Entity Time
Observations	3,168	3,168	3,168
R^2	0.076	0.089	0.116
F Statistic	155.91***	195.87***	265.83***

Note: *p<0.1; **p<0.05; ***p<0.01
Data Source: https://www.yelp.com/dataset

Table 8. Fixed effect Regression results 2

	Open rate	Close rate	Net inflow(percentage)
	(1)	(2)	(3)
Covid * CBD	-0.213***	0.094	-0.041
	(0.089)	(0.076)	(0.049)
Constant	0.930***	1.006***	-0.059***
	(0.003)	(0.002)	(0.001)
Effects	Entity Time	Entity Time	Entity Time
Observations	2,107	2,189	2,763
R^2	0.017	0.06	0.002
F Statistic	81.687***	84.248***	19.926***

Note: *p<0.1; **p<0.05; ***p<0.01
Data Source: https://www.yelp.com/dataset

word count: 7560

References

Census Bureau API: access via python. https://towardsdatascience.com/accessing-census-data-with-python-3e2f2b56e20d. Accessed: 2022-06-03.

Philadelphia: city shape file. https://www.opendataphilly.org/dataset/philadelphia-neighborhoods/resource/6c61f240-aafe-478e-b993-b75fd09a93d6. Accessed: 2022-06-03.

lda: access via python. https://scikit-learn.org/0.16/modules/generated/sklearn.lda.LDA.html. Accessed: 2022-06-03.

Vader: access via python. https://pypi.org/project/vaderSentiment/#:~:text=VADER%20(Valence%20Aware%20Dictionary%20and,on%20texts%20from%20other%20domains. Accessed: 2022-06-03.

Yelp: the complete yelp business category list. https://blog.yelp.com/businesses/yelp_category_list/, a. Accessed: 2022-06-03.

Yelp: business data. https://www.yelp.com/dataset, b. Accessed: 2022-06-03.

Lukas Althoff, Fabian Eckert, Sharat Ganapati, and Conor Walsh. The geography of remote work. *Regional Science and Urban Economics*, 93:103770, 2022.

James Lopez Bernal, Steven Cummins, and Antonio Gasparrini. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International journal of epidemiology*, 46(1):348–355, 2017.

Leland D Crane, Ryan A Decker, Aaron Flaaen, Adrian Hamins-Puertolas, and Christopher Kurz. Business exit during the covid-19 pandemic: Non-traditional measures in historical context. *Journal of macroeconomics*, 72:103419, 2022.

Lei Dong, Carlo Ratti, and Siqi Zheng. Predicting neighborhoods' socioeconomic attributes using restaurant data. *Proceedings of the national academy of sciences*, 116 (31):15447–15452, 2019.

Naveen Donthu and Anders Gustafsson. Effects of covid-19 on business and research, 2020.

Catherine E Fazio, Jorge Guzman, Yupeng Liu, and Scott Stern. How is covid changing the geography of entrepreneurship? evidence from the startup cartography project. Technical report, National Bureau of Economic Research, 2021.

Richard Florida, Andrés Rodríguez-Pose, and Michael Storper. Cities in a post-covid world. *Urban Studies*, page 00420980211018072, 2021.

Edward L Glaeser, Hyunjin Kim, and Michael Luca. Nowcasting the local economy: Using yelp data to measure economic activity. Technical report, National Bureau of Economic Research, 2017.

Edward L Glaeser, Hyunjin Kim, and Michael Luca. Nowcasting gentrification: using yelp data to quantify neighborhood change. In *AEA Papers and Proceedings*, volume 108, pages 77–82, 2018.

Matthew J Holian. Where is the city's center? five measures of central location. *Cityscape*, 21(2):213–226, 2019.

Myounggu Kang, Yeol Choi, Jeongseob Kim, Kwan Ok Lee, Sugie Lee, In Kwon Park, Jiyoung Park, and Ilwon Seo. Covid-19 impact on city and region: what's next after lockdown? *International Journal of Urban Sciences*, 24(3):297–315, 2020.

Sitian Liu and Yichen Su. The impact of the covid-19 pandemic on the demand for density: Evidence from the us housing market. *Economics letters*, 207:110010, 2021.

Alexander W Olson, Fernando Calderón-Figueroa, Olimpia Bidian, Daniel Silver, and Scott Sanner. Reading the city through its neighbourhoods: Deep text embeddings of yelp reviews as a basis for determining similarity and change. *Cities*, 110:103045, 2021.

Arjun Ramani and Nicholas Bloom. The donut effect of covid-19 on cities. Technical report, National Bureau of Economic Research, 2021.

Stephen J Redding. Economic geography: a review of the theoretical and empirical literature. *Palgrave handbook of international trade*, pages 497–531, 2013.