How Does Crime Affect Business Rating in Philadelphia

Yaqi Du

Department of Economics, University of Toronto

April 12, 2024

1. Introduction

In today's digital world, the way consumer shares their opinions about business has shifted to an online format. People tend to provide feedback through online ratings or reviews. These online ratings aggregate individual feedback into a collective judgment and assign businesses an average rating, which becomes an indicator of overall customer satisfaction. Moreover, these online ratings not only represent each consumer's personal taste; the overall score they combined reflects the demand for the business in the market. Thus, the average star ratings for each business have become a key indicator of consumer preferences and offer a detailed look into the market behavior. Although it is clear that the quality of the service that businesses provide directly affects these ratings, the impact of the external elements, such as the safety of the area surrounding the business, is rarely mentioned. This study aims to fill this gap by exploring how local crime rates influence customer satisfaction, as reflected by online ratings on platforms like Yelp.

The economic consequences of crime have a severe effect on consumer preferences which consequently decreases the performance of businesses. The pioneering study by Fe (2022) offers a detailed look at how different types of crime affect consumer decisions when visiting a business. The results reveal that consumers are apparently unwilling to patronize the businesses located in areas perceived as unsafe, especially when these places often experience public crime. Interestingly, the study outlines a more pronounced impact of property-related offenses compared to violent crimes on consumer preferences.

From another side to discuss the impact of crime, Bressler (2012) underlines the critical need for crime prevention that ensures business profit, or at least its sustainability. Focusing on the escalating costs related to the increasing criminal activities, Bressler addresses what potential

financial crisis businesses face when there is no effective prevention in place. The cost of crime is not just a statistic number but a crucial threat to over every unprotected business.

The studies discussed above reveal an important connection between crime and consumer preference or business outcome. However, they do not explore the direct effects of crime on consumer satisfaction. Specifically, the question of how the perceived safety problem of a business's surroundings influences consumer satisfaction remains unanswered. This paper aims to fill this research gap by examining how crime influences consumer satisfaction as manifested through online feedback mechanisms.

Moreover, the previous studies identify some mediators beyond crime that may impact business ratings. According to Mario (2017), income equality acts as a socio-economic stressor that affects human behavior, potentially leading to an increase in violent crime. Similarly, Bruce (1998) finds income inequality could be a powerful predictor of violent crime. On the other hand, Casey (2016) examines the effect of wealth disparity on entrepreneurship. Furthermore, Harpriya et al. (2020) point out the positive influence of family income on business performance. This paper broadens these discussions, and examines the complex interactions between crime rates, business ratings, and other contributing factors, thereby enriching our understanding of the multifaceted dynamics of the online business rating.

The subsequent sections of this paper will delve into four parts: data analysis, summary statistics, and findings. It shows the intricate dynamics that shape consumer satisfaction and ultimately reveals the true impact of crime on the success of businesses within the digital marketplace.

2. Data

2.1 Yelp Data

My study utilizes Yelp's comprehensive dataset to examine the impact of urban crime rates on consumer satisfaction, with the last update to the data in 2021. The main function of Yelp is to allow users to search for businesses by ratings, categories, or locations, which is instrumental in our economic analysis. Initially, our dataset included 150,345 observations across mainly eight metropolitan areas, showing a skewed distribution concentrating in certain cities. This skewed distribution led me to focus my efforts on Philadelphia. With its substantial 14,569 observations, Philadelphia owns the largest subset among all the cities, which makes it present a prime opportunity for a particular examination of the interplay between consumer satisfaction and urban crime rates.

My decision to focus on Philadelphia is due to the depth of its data, which enables me to conduct a detailed investigation into how external environmental factors affect consumers' online ratings. This targeted approach helps us explore the complex dynamics that impact consumer perception of a business, which are connected to business success.

2.2 Crime Data

Incorporating the crime and population statistical data from OpenDataPhilly has significantly enriched our analysis of the effect of crime rate on online business ratings in Philadelphia. OpenDataPhilly provides a detailed record of the crime incidents in 2021. It includes the specific location (longitude and latitude) and the types of crime, which are classified through Uniform Crime Reporting (UCR) codes. The UCR codes enable us to categorize crimes specifically into violent crimes, which consist of four offenses as defined by the FBI: murder and

non-negligent manslaughter, rape, robbery, and aggravated assault. This specific categorization of crime, combined with the population data can lead us to conduct sophisticated analysis of the different types of crime rates in various neighborhoods, thus understanding the interplay between local crime and consumer from a multifaceted viewpoint.

2.3 Economic Factor Data

To enrich the investigation of the economic influences on crime rates and business ratings, I incorporated the Gini index and median family income data from Zipatlas at the zip code level. Zipatlas is a website that provides a comprehensive collection of zip code characteristics. By web-scrapping the data, we are able to gain in-depth socio-economic insights into the local communities.

The Gini index is known for its ability to reflect income inequality within a specific region, while median family income indicates community wealth. The inclusion of such factors establishes a solid framework to explore the economic business environment.

By utilizing the Zipatlas data, the study captures the diversity and disparities experienced in Philadelphia's economy. Linking these economic measures with consumer behavior and business outcomes makes my research also underscores the importance of economic health and purchasing power to business success, as suggested by Wang et al. (2018).

Analyzing these economic indicators alongside crime rates and business ratings offers a holistic view of the factors influencing business success in urban areas.

2.4 Data Aggregation

Similar to another empirical study in the previous literature Fe(2022), a critical challenge I am facing is how to define local neighborhoods. This issue is crucial, as it fundamentally influences the design of our analytical model and the potential accuracy of its outcomes.

There is a wide range of possible options for defining a neighborhood according to the given geographic dimension. Nevertheless, there is still a lack of evidence for which dimensionality level can best reflect the impact of the crime rate. Usually, the influence of crime on business is constrained in a small geographic area, thus choosing smaller neighborhoods might further enhance our data, making our analysis more accurate and detailed. However, considering the spatial spillover effects, we should be cautious about using the over precise categorizations, such as the Census tracts. Using such neighborhood definitions might bypass the broader span of the crime-impact areas that are next to them.

Balancing these considerations, I have opted to use the zip code as the unit of analysis for this study. This approach ensures that a set of data is comprehensive but still relevant geographically enough to reflect the relationships between crime rates and certain economic performances in particular areas.

However, in the implementation of the regression analysis, one did not go to the extent of using the inner merge as a key strategic methodological choice. Generally, an inner merge would have reduced our dataset to only those observations that have matching data across all specified variables, potentially reducing the dataset to only 46 zip code-level observations. Instead, a left join has been made following the incorporation of the socioeconomic factors at the ZIP code level to ensure all the 14,569 businesses in Philadelphia are kept within this analysis. For that reason, socio-economic data corresponding to the zip code of the location of the businesses was also added in the additional columns for each business.

3. Summary Statistics

3.1 Summary Table of Key Variables

	crime_rate	violent_crime_rate	stars
count	46.000000	46.000000	14569.000000
mean	8.811376	0.915320	3.623035
std	4.329999	0.612426	0.960691
min	2.512606	0.142438	1.000000
25%	5.627982	0.456131	3.000000
50%	7.868035	0.822544	4.000000
75%	10.817776	1.112996	4.500000
max	22.400210	2.625807	5.000000

The table above shows the summary statistics for our key variables. The statistics for crime rate and violent crime rate are aggregated by 46 distinct zip codes, offering a localized view of these variables. Meanwhile, ‘stars' is derived from the broader Yelp dataset in Philadelphia, providing additional insight into consumer ratings of the total 14569 businesses.

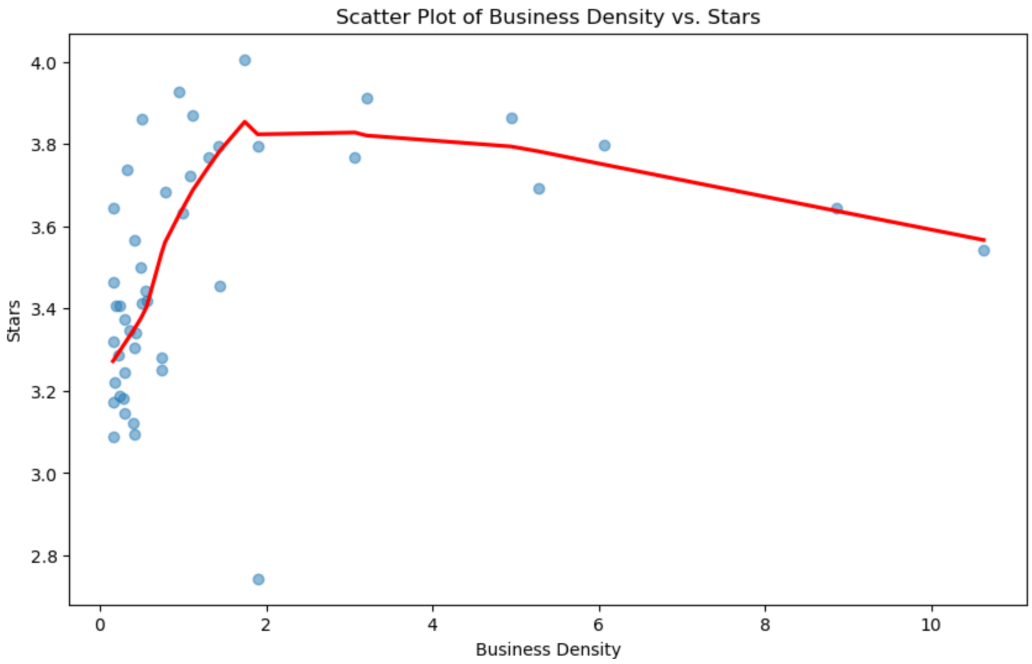
The mean crime rate of 8.81 per 100 people shows a large disparity with the extremely high standard deviation value of 4.32 in different areas' safety levels. This means that the actual effect of crime on local businesses and subsequently on local customer ratings will be very different between the two zip codes. Areas characterized by a high crime rate tend to lose many potential

customers, who of course might be bad for business, while low crime rates help in building a stronger economic environment.

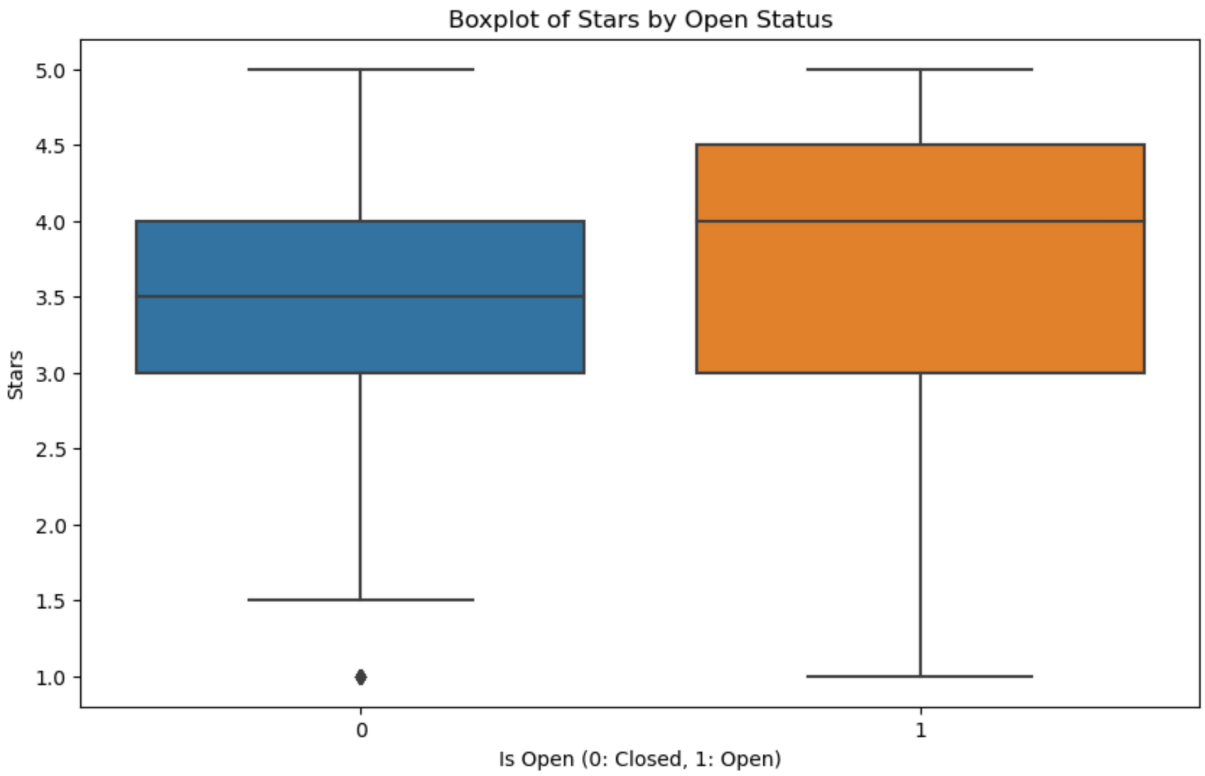
The violent crime rate, with a mean of 0.92, indicates that such crimes are relatively rare. Nonetheless, the existence of violent crimes is likely to have deeper psychological impacts on the perceived safety of a community, and it could, in a domino effect, influence consumer behavior and business performance even further than the general crime level would as Rosenthal and Ross (2010) indicate in their research. The lower standard deviation compared with the overall crime for violent crime rates gives less variability, but again, the presence of violent crime is a key issue for businesses.

An average of 3.62 and a median of 4 on business rating generally point to the levels of customer satisfaction being good. As seen, despite the challenges of the business environment, which include a high level of competition and a high rate of crime, a significantly large number of businesses either meet customers' expectations or surpass them. The implication is that quality of service, product, or experience may override some of the ill effects connected with increased crime rates or cut-throat competition.

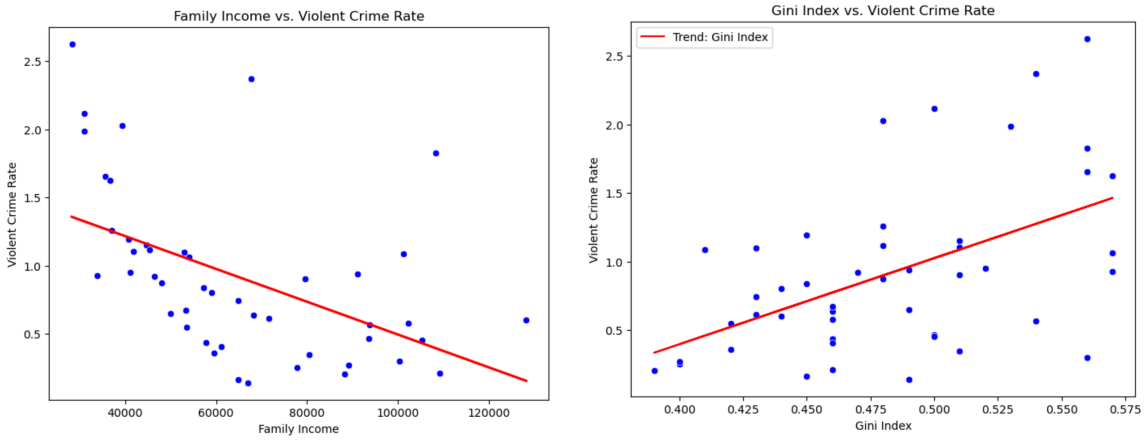
3.2 Visualization



The scatter plot of business density vs. stars on zip code level shows that there is an intricate relationship: it does increase to some density levels, and after that is attained, it keeps increasing up to a point with business density and average star rating on Yelp. This is probably to be caused by the benefits brought about by competition, which increase quality and customer service. On the other side, when business density goes up more than the threshold of 2, the ratings start falling, which is an indication that such competition will eventually lead to saturation of the market or over-competition, with its incumbent serious waste in resources and diminishing returns. The evidenced pattern in the scatter plot, furthers the research question of what influences Yelp ratings if it is not the crime rates that do so. It points out to be the competitive market that the business is in. This is a very critical condition, which emphasizes the fact that the impact on competition has overwhelmed the effect of oversaturation, proving that the rating is affected by more than crime; the businesses themselves have an impact on it from within and by the dynamics of their competitiveness.



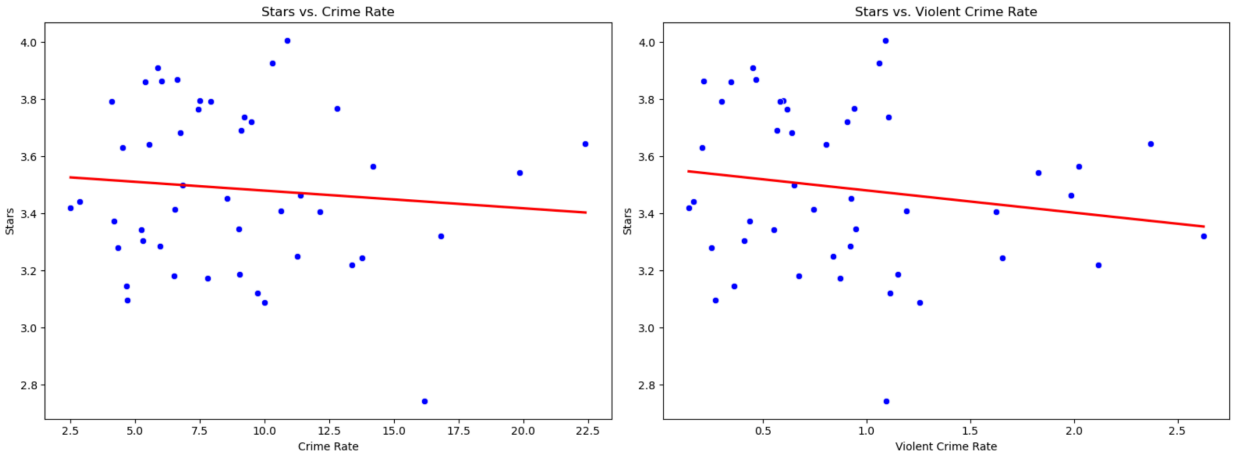
In fact, the tale from the boxplot of Yelp star ratings with regard to business open status tells an interesting story: generally, business open status tends to feature a higher median rating in comparison to businesses being closed. This may suggest some positive relationship between operational status and consumer perception. On the other hand, closed businesses display a remarkable spread in ratings, with the outliers skewed toward the lower side, hence pointing to potential poor influences in case some may be leading to closing the business. Moreover, the smaller spread of ratings for closed businesses indicates an even customer experience but not necessarily a good one. One big spotlight on open status as a variable may perhaps reveal many of the multifarious dynamics at play in Yelp ratings, besides crime. While not being causal in nature, the data does tend to support that operational status might be a large factor in determining higher consumer-based ratings of businesses, therefore still pointing the need to look at other factors in trying to determine exactly what affects Yelp ratings.



The zip code level scatter on the left indicates that an inverse relationship between family income and violent crime. In actuality, the line of best fit moving down indicates that it will be higher family incomes where there is a tendency for the rates of violent crime to be lower. This implies the economic booms of society can enhance stability, hence, reducing the number of crime incidents due to possibly reduced financial tension between the citizens, which aligns with the findings of the research by Hsieh and Pugh (1993). This inverse relationship is something that has to be taken care of in its regressions because it carries crucial implications in regard to the socio-economic environment in its effects on violent crime.

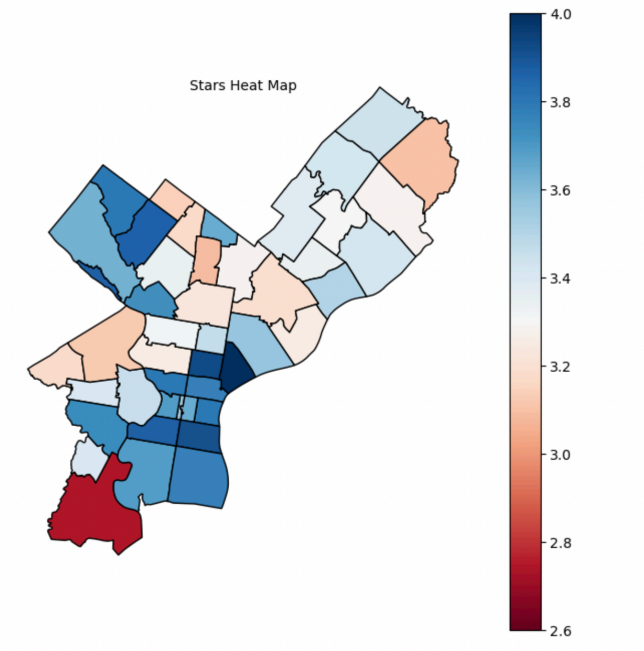
On the other hand, the right scatterplot clearly shows a positive relation between the Gini Index with violent crime rates because it contains an upward-sloping regression line. This goes in line with the hypothesis that higher income inequality in a region may actually increase social tensions and, hence, tend toward more violent crimes.

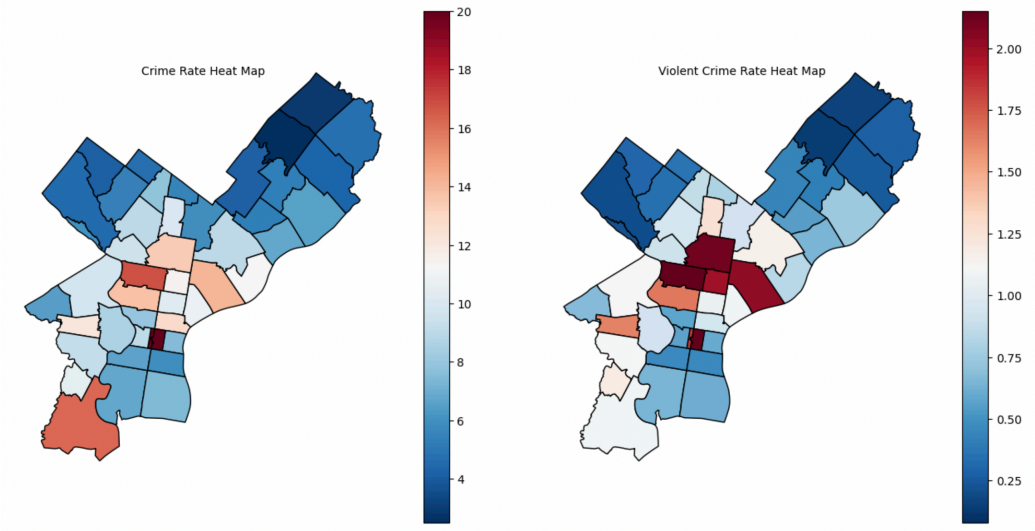
These two scatter plots illustrate a complex relationship between economic variables and crime rates. Moving forward into the regression analysis, more particularly into the combined effect of these relationships in relation to our dependent variable, business rating on Yelp, will be very important in the succeeding sections.



The scatterplot of the average business rating by zip code against crime rates on Yelp roughly reveals distinct patterns for overall crime rates and violent crime rates. The trend line for the overall crime rates is so close to flat, it suggests a light linear relationship between the crime rates and Yelp ratings. By and large, businesses have middle-of-the-line ratings, regardless of a high or low local crime rate. This implies other factors, perhaps some relating to business density and operational status already mentioned, play even more important roles in affecting the consumer ratings than the general crime rate. On the other hand, the line of the trend for the violent crime rate is having a more negative correlation, meaning that as the rate of violent crime goes high, the general trend of the Yelp business rating goes down. This suggests that, compared to the level of general criminality, violent crime is likely to impact more on the consumers' perceptions and satisfaction, largely because of the increased extent of insecurity or concern that it raises. While overall crime does not seem to predict strongly the business performance in Yelp, the existence of violent crime does predict a stronger negative effect on consumer reviews. This has therefore demonstrated the differing effects that different types of crime can have on consumer behavior and business ratings.

3.2 Mapping Key Variables





The analysis of heat maps illustrates the average Yelp ratings for businesses across various zip code areas and the frequency of violent and overall crimes reported per 100 people. It offers a sophisticated visual tool for examining the geographic correlation between business ratings and crime rates. Utilizing a color gradient, the maps convey the intensity of crime rates and the level of consumer ratings, with cooler colors indicating lower business ratings and crime rates.

The heat maps provide visual analysis that contrasts and compares zip code areas. They show that, even in the areas with high overall crime rates, the business ratings still persist at an average of 3.2 to 3.4. This suggests that, on average, consumer satisfaction does not necessarily correlate with the local overall crime rate. Interestingly, regions with lower overall crime rates, especially in the northwest segments, reveal a trend to higher business ratings. These observed results do show that overall crime might have a negative effect on the business rating, but the relationship is not straightway. Higher crime rates simply do not turn towards lower Yelp ratings; rather, it suggests other potential factors that have an impact on consumer perceptions and ratings.

The heat maps further prove that, in general, high violent crime areas do not line up with areas showing high Yelp business ratings, which further strengthens the point that the negative relationship of high violent crime rates is with consumer satisfaction.

These alignment with previous findings confirms the complexity of the relationship between crime rates and business ratings on Yelp. This insight indicates the need to consider a much broader set of socio-economic and environmental variables when assessing the impact of crime on business reputations.

4. Results

4.1 Empirical strategy

My goal is to explore the impact of crime on business ratings on Yelp, and I mainly focus on the two types of crime: overall crime and violent crime. Additionally, I will control the other confounding like business density, operational status, and economic environment to discern the true effect of crime and the potential interaction effects more accurately it may have with these

variables. We denote the vector β to represent the effect of our independent variables (e.g. crime rate).

Building on the discussion in Section 3, I will use the simple linear regression to discuss the independent effect of crime and business factors on Yelp ratings. This approach brings insights into the direct effect of these variables.

Furthermore, when examining the economic variables, simple linear regression runs the risk of introducing bias. In the previous analysis, is a potential connection between crime and the economic environment. According to a study by Mbuthia (2019), it was found that family income does not directly correlate with Yelp ratings. However, it exhibits a strong negative correlation with crime. Therefore, it's reasonable to treat the family income as an instrumental variable, considering its indirect effects on business ratings through its relationship with crime.

4.1 Main Result

In this section, I present the core findings of my paper. There I focus on the independent effect of the variables on Yelp business rating, which is structured around this linear equation:

$$Stars_i = \beta_0 + \beta_1 X1_i + \beta_2 X2_i + \ldots + \beta_n Xn_i + u_i$$

					Dependent variable: stars
	(1)	(2)	(3)	(4)	(5)
Gini					0.281
					(0.190)
business_density		0.051***	0.054***	0.052***	0.051***
		(0.005)	(0.005)	(0.005)	(0.005)
const	3.646***	3.742***	3.639***	3.738***	3.608***
	(0.019)	(0.021)	(0.026)	(0.028)	(0.092)
crime_rate	0.041***	-0.012*	-0.012*	-0.013*	-0.012*
	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
is_open			0.135***	0.095***	0.096***
			(0.020)	(0.020)	(0.020)
is_restaurant				-0.159***	-0.160***
				(0.018)	(0.018)
violent_crime_rate	-0.381***	-0.090*	-0.098**	-0.081*	-0.093*
	(0.040)	(0.048)	(0.048)	(0.048)	(0.049)
Observations	10962	10962	10962	10962	10962
R^2	0.008	0.018	0.023	0.030	0.030
Adjusted R^2	0.008	0.018	0.022	0.029	0.029
Residual Std. Error	0.900 (df=10959)	0.896 (df=10958)	0.894 (df=10957)	0.891 (df=10956)	0.891 (df=10955)
F Statistic	46.686*** (df=2; 10959)	67.938*** (df=3; 10958)	63.114*** (df=4; 10957)	67.072*** (df=5; 10956)	56.263*** (df=6; 10955)
Note:					*p<0.1; **p<0.05; ***p<0.01

This table shows the effects of the overall crime rate(crime_rate) and violent crime rate(violent_crime_rate) on Yelp ratings, adjusting for various confounders. The result from the base model, as illustrated in column (1), reveals a significant and substantial negative correlation between the violent crime rate and Yelp ratings. However, when it comes to the overall crime rate, it exhibits an intriguing positive relationship with Yelp ratings, suggesting an increase in stars as the crime rate rises. This result appears to contradict the common sense and the consumer preference theory.

But as I control the business density, the dynamics shift towards expected patterns. Even though t a marked reduction in the magnitude of the violent crime rate's coefficient from -0.381 to 0.090, the influence of the overall crime rate on stars becomes negative. Meanwhile, business

density is shown to positively affect ratings, which aligns with the competition theory that increased competition may drive quality improvements, thus enhancing consumer satisfaction.

Upon incorporating operational status in column (3), the findings confirm that open businesses receive a higher star. And the impact of the violent crime rate becoming more pronounced. Column (4) includes a dummy variable(is_open) indicating whether a business is a restaurant or not. The analysis shows that the adverse effect of the violent crime rate on stars decreased again as we control the category of the business, and being a restaurant is associated with lower ratings. In column (5), I add the Gini index (Gini) to examine the social-economic influence. However, the results reveal a positive but statistically insignificant effect on stars. This may suggest that the Gini index is not a reliable predictor of Yelp star ratings.

Overall, both the R-squared and adjusted R-squared values increase with the introduction of additional variables. This indicates that the model's predictive accuracy is enhanced. After a thorough comparison of all the models above, I conclude that the fourth one is the best the best predictive performance for Yelp ratings, not only because it has a superior statistical outcome, but also due to the significance of all its variables.

4.2 Explore The Interaction Effect

In the table below, I explore the interaction effects among variables based on the preferred model I discussed in Section 4.1:

$$Stars_i = \beta_0 + \beta_1 ViolentCrimeRate_i + \beta_2 CrimeRate_i + \beta_3 BusinessDensity_i + \beta_4 IsOpen_i + \beta_5 IsRestaurant_i + u_i$$

				Dependent variable: stars
	(1)	(2)	(3)	(4)
Gini				2.454***
				(0.457)
business_density	0.058***	0.033***	0.053***	0.064***
	(0.005)	(0.007)	(0.005)	(0.006)
const	3.644***	3.811***	3.827***	2.494***
	(0.043)	(0.033)	(0.042)	(0.229)
crime*gini				-0.019
				(0.120)
crime*open			0.049***	
			(0.011)	
crime_rate	-0.007	-0.014**	-0.052***	-0.010
	(0.007)	(0.007)	(0.011)	(0.058)
density*open		0.027***		
		(0.006)		
is_open	0.097***	0.002	-0.013	0.102***
	(0.020)	(0.030)	(0.046)	(0.020)
restaurant	-0.159***	-0.160***	-0.160***	-0.157***
	(0.018)	(0.018)	(0.018)	(0.018)
violent*all	-0.008***			
	(0.003)			
violent*gini				-2.994***
				(0.958)
violent*open			-0.424***	
			(0.097)	
violent_crime_rate	0.049	-0.072	0.258***	1.506***
	(0.066)	(0.048)	(0.091)	(0.474)
Observations	10962	10962	10962	10962
R^2	0.030	0.031	0.031	0.036
Adjusted R^2	0.030	0.031	0.031	0.035
Residual Std. Error	0.890 (df=10955)	0.890 (df=10955)	0.890 (df=10954)	0.888 (df=10953)
F Statistic	57.303*** (df=6; 10955)	58.897*** (df=6; 10955)	50.760*** (df=7; 10954)	50.823*** (df=8; 10953)

The analysis begins with column (1), where the model assesses an interaction term between the violent crime rate and the overall crime rate(violent*all). This interaction is intended to capture the additional marginal effect of the violent crime rate. The result shows a negative and significant interaction coefficient, which indicates that the combination of overall crime and

violent crime exacerbates their negative impact on Stars. This is likely due to increased consumer concerns for their safety.

Column (2) examines the interaction between business density and open status (density*open). From the economic perspective, this interaction reflects market dynamics. As expected, there is a significant and positive relationship between the interaction term and Yelp rating. This finding indicates that the benefit of business density on the rating is enhanced when the business is open, likely because areas with more opening businesses attract more customers, fostering intense market competition, which could lead to higher quality and, thus, higher ratings.

In column (3), the model introduces the interaction of crime and business operational status considering the externalities on the local economic. The interaction term of violent crime and open status is negative, which suggests that the negative effect of violent crime on stars is more pronounced for the businesses that are open. Interestingly, when we look at the overall crime, the interaction is positive. It shows that consumers' ratings are less impacted by the overall crime for a business that is opening. This might indicate that though overall crime rates are high in some areas, they do not necessarily deter consumer satisfaction if the areas are not perceived as dangerous.

Column (4) includes the interaction of the Gini index and crime, and it shows that violent crime significantly worsens business ratings in areas with greater income inequality. This result aligns with my previous hypothesis that the Gini index is closely linked to violent crime.

In summary, this analysis explores some key interactions of among the dependent variables. Although the statistics results show that incorporating the interaction terms into regression does not necessarily improve the predictability and accuracy, these insights still contribute to a deeper

understanding of the complex dynamics that influence consumer satisfaction and business ratings on Yelp.

4.3 IV Regression

In Section 3.1 I explore that median family income might have a negative relationship with the violent crime rate. Namely, there is likely a two-way relationship between violent crime rate and stars: Richer families often reside in safer environments, potentially leading to lower violent crime rates. However, according to Mbuthia (2019), we can determine that the income of a neighborhood is not directly correlated to business rating on Yelp. The simple linear regression here might return a biased result. Specifically, it could show higher violent crime rates with better business ratings due to analytical bias.

Given these considerations, median family income could be considered as a potential instrumental variable. Utilizing it as such could provide clearer insights into the causal relationships between violent crime rates, business ratings, and economic status.

Model Comparison

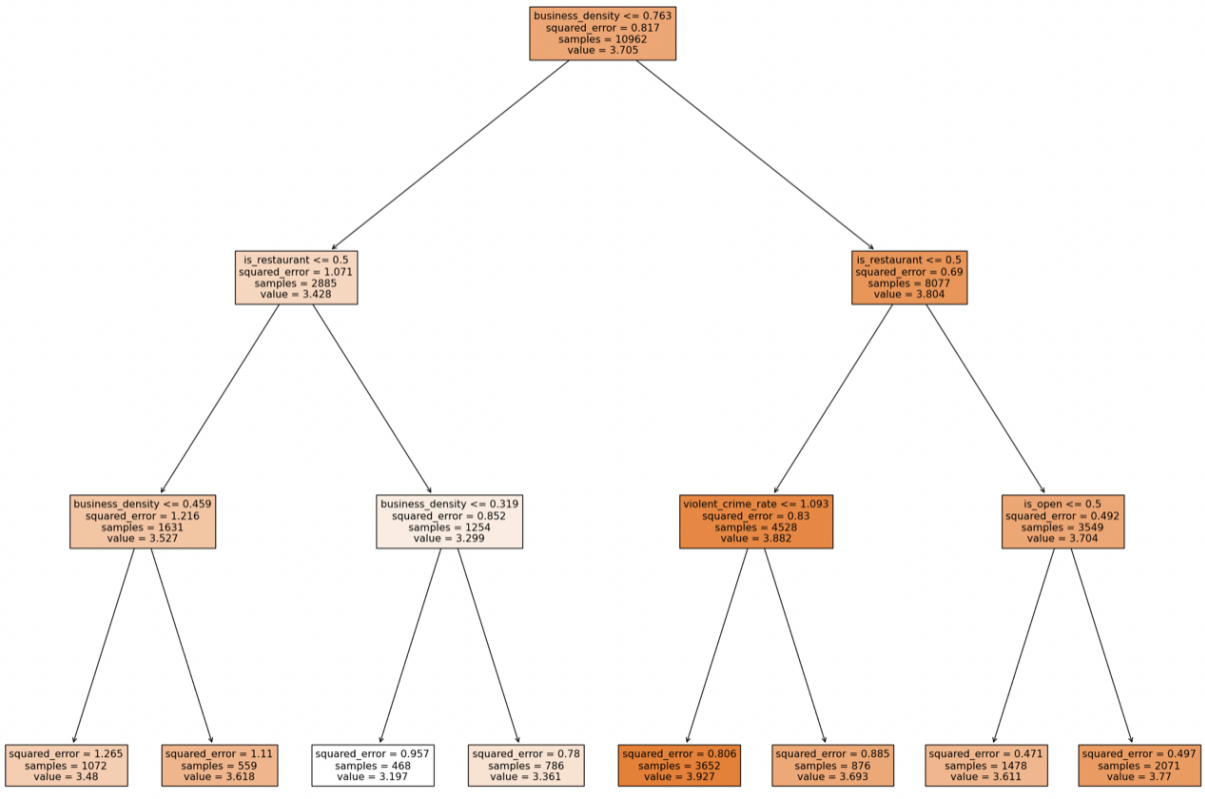
	OLS	IV2SLS
Dep. Variable	stars	stars
Estimator	OLS	IV-2SLS
No. Observations	10962	10962
Cov. Est.	robust	robust
R-squared	0.0345	-0.2535
Adj. R-squared	0.0343	-0.2537
F-statistic	380.58	326.52
P-value (F-stat)	0.0000	0.0000
==========	==========	==========
const	3.1513	4.3970
	(87.812)	(117.68)
violent_crime_rate	0.0364	-0.7663
	(2.6632)	(-18.070)
Income	6.586e-06	
	(19.107)	
==========	==========	==========
Instruments		Income

From the model comparison table, it is evident that in the Ordinary Least Squares (OLS) model, the violent crime rate effect on the business rating is positive when controlling for income, meaning that the businesses receive a better reputation if they are located in dangerous areas. This result contradicts to common sense and my previous analyses where all the coefficients related to violent crime are negative.

However, as I regard the income as an instrumental variable and applied Instrumental Variables to Two-Stage Least Squares (IV2SLS), the results shift significantly. The new model suggests that higher violent crime rates actually lead to lower stars. This adjustment helps to mitigate the bias present in the OLS model and enhances the validity of our conclusion about the effect of crime on Yelp business rating.

4.3 Regression Tree

To identify which variable exerts the most influence on Yelp stars, in this section, I delve into the predictive power of the regression tree, based on the preferred variables from the most accurate model discussed in Section 4.1.



As the regression tree highlights, the business density is the most significant factor in predicting. This may imply that market competition is crucial for improving consumer satisfaction. Following business density, the next split is on the business category. The tree shows that non-restaurant businesses tend to have higher average stars, reflecting the different expectations of consumers between these categories.

Further down the tree, the nodes include variables such as violent crime rate, open status, and business density again. As we can observe from the tree, businesses in areas with higher violent crime rates generally receive lower stars, whereas the opening businesses tend to have higher ratings. These results are consistent with our OLS regression results, confirming the negative impact of violent crime on business ratings and the positive effect of operational status.

Moreover, the regression tree also shows how the variables interact — for instance, how business density affects ratings differently in restaurants versus non-restaurant businesses. This nuanced understanding helps businesses identify which factor is more critical for their reputation under certain situations.

5. Conclusions and Future work

5.1 Conclusions

Extensive research has been conducted on the impact of crime on business outcomes and consumer preferences. However, there is much less attention paid to how crime affects consumer satisfaction, especially as reflected through online ratings and reviews. This paper fills part of the gap with a detailed analysis of the impact of different types of crime on Yelp business rating in Philadelphia. Central to my research is the idea that consumers' satisfaction with a business depends on crime type and the economic environment in which the business is located.

Our main findings reveal that violent crime has a significant negative effect on business ratings. In addition, the main results suggest that overall crime does not necessarily have a negative effect on business ratings, or that its negative effect is much less pronounced compared to violent crime. This provides additional evidence that when evaluating a business, consumer may take the violent crime as their primary consideration.

By exploring the interaction effect of crime and other factors related to the surrounding environment of the business, I find that the effect of violent crime rate is intensified by the Gini index, indicating that the negative experiences brought about by violent crime are more pronounced in areas with higher income inequality. Interestingly, restaurants may have lower business ratings compared to non-restaurant businesses. This could be due to the quality of the food provided by restaurants being more tangibly assessed by consumers. I also ran a regression tree to identify which variable exerts the most influence on Yelp stars and found that business density is the most important factor, followed by the category of the business. Economic theory has concluded that market competition can improve the quality of goods that producers supply. This may suggest that for the consumer, quality is always the primary concern,

Our work indicates that consumers take crime, especially violent crime, into consideration when rating the service provided by a business on Yelp. We conclude that our findings are consistent with the argument that consumers' satisfaction may be decreased due to the perception of violent crime, potentially leading to business failure. These insights are helpful for police maker formulating urban economic strategies and provide valuable information to businesses when choosing their location.

5.2 Future Work

There are still many avenues for further analysis that have been left for the future due to our main Yelp data only being updated to 2021, which means that our analysis is only confined to the year 2021. If we are able to access the Yelp data that includes the change in stars over time, we are able to draw a more comprehensive study. In my OLS model, I didn't include the fixed effect because our dataset is not panel data. Thus, if the dataset could be enriched, the predictive function would be more accurate.

Additionally, in the study, I noticed the coefficient of the Gini index remains positive in both independent and interaction models. Although it doesn't show significance, this unusual outcome still warrants investigation. Does it really mean the area with higher inequality will lead to a higher business reputation? What is the reason behind such a phenomenon? Is the Gini Index a mediator variable or an instrument? To answer this question, further study is needed to examine the detailed link between the Gini Index, business rating, and crime.

References

Bressler, M. (n.d.). *Journal of Management and Marketing Research The Impact of Crime on Business*. https://www.aabri.com/manuscripts/09202.pdf

Coccia, M. (2017, April 11). *General Causes of Violent Crime: The Income Inequality*. Papers.ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2951294

Fe, H., & Sanfelice, V. (2022). How Bad Is Crime for Business? Evidence from Consumer Behavior. *Journal of Urban Economics*, *129*, 103448. https://doi.org/10.1016/j.jue.2022.103448

Frid, C. J., Wyman, D. M., & Coffey, B. (2016). Effects of wealth inequality on entrepreneurship. *Small Business Economics*, *47*(4), 895–920. https://doi.org/10.1007/s11187-016-9742-9

Harpriya, Sharma, R. K., & Sah, A. N. (2020). Impact of demographic factors on the financial performance of women-owned micro-enterprises in India. *International Journal of Finance & Economics*. https://doi.org/10.1002/ijfe.2133

Highest Median Family Income in Philadelphia by Zip Code | Zip Atlas. (n.d.). Zipatlas.com. Retrieved April 12, 2024, from https://zipatlas.com/us/pa/philadelphia/zip-code-comparison/highest-median-family-income.htm#map

Hsieh, C.-C., & Pugh, M. D. (1993). Poverty, Income Inequality, and Violent Crime: A Meta-Analysis of Recent Aggregate Data Studies. *Criminal Justice Review*, *18*(2), 182–202. https://doi.org/10.1177/073401689301800203

Kennedy, B. P., Kawachi, I., Prothrow-Stith, D., Lochner, K., & Gupta, V. (1998). Social capital, income inequality, and firearm violent crime. *Social Science & Medicine*, *47*(1), 7–17. https://doi.org/10.1016/s0277-9536(98)00097-5

Lowest Inequality / Gini Index in Philadelphia by Zip Code | Zip Atlas. (n.d.). Zipatlas.com. Retrieved April 12, 2024, from https://zipatlas.com/us/pa/philadelphia/zip-code-comparison/lowest-inequality-gini-index.htm#data

Mbuthia, J. (2019, August 21). *Is there a correlation between a restaurants ratings and the income levels of a neighborhood ?*Medium; The Startup. https://medium.com/swlh/is-there-a-

correlation-between-a-restaurants-ratings-and-the-income-levels-of-a-neighborhood-5fe41165e4f1

MCGLOIN, J. M., & ROWAN, Z. R. (2015). A THRESHOLD MODEL OF COLLECTIVE CRIME. *Criminology*, *53*(3), 484–512. https://doi.org/10.1111/1745-9125.12077

Nau, R. (2019). *What's a good value for R-squared?* Duke.edu. https://people.duke.edu/~rnau/rsquared.htm

Rosenthal, S. S., & Ross, A. (2010). Violent crime, entrepreneurship, and cities. *Journal of Urban Economics*, *67*(1), 135–149. https://doi.org/10.1016/j.jue.2009.09.001

Sribney, W., Wiggins, V., Drukker, D., & StataCorp. (n.d.). *FAQ: Negative and missing R-squared for 2SLS/IV | Stata*. Www.stata.com. https://www.stata.com/support/faqs/statistics/two-stage-least-squares/

Wang, Y., Xu, X., He, J., Chen, C., & Ren, K. (2018). Impact of Purchasing Power on User Rating Behavior and Purchasing Decision. *Lecture Notes in Computer Science*, 413–422. https://doi.org/10.1007/978-3-319-93818-9_39

Yelp. (2021, February 16). *Yelp Dataset*. Www.kaggle.com. https://www.kaggle.com/datasets/yelp-dataset/yelp-dataset/data