Essays on Risk Avoidance Behaviors

Junwei Mao

A dissertation
submitted in partial fulfilment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Xu Tan, Chair

Fahad Khalil

Jing Tao

Program Authorized to Offer Degree:

Economics

University of Washington

Abstract

Essays on Risk Avoidance Behaviors

Junwei Mao

Chair of the Supervisory Committee:
Xu Tan
Department of Economics

This dissertation is a collection of essays on individual's risk avoidance behaviors. Risk avoidance behavior is a key topic in behavioral economics, and it is also a significant impactor of market dynamics.

Chapter 1 investigates the labor participation activity in the Chicago taxi industry to provide a new method for estimating the wage-risk tradeoff. I propose a cross-period specification to estimate the value of a statistical life (VSL) in a solo industry setup. I find that taxi drivers in Chicago city exhibit lower risk premium than that in other cross-sectional analysis, however, the hazard compensation to keep their workload is similar to some regulated COVID-19 hazard pay in other industry. I also find evidence supporting that taxi drivers make their work decisions across the days while most existing literature is focused on the in-day decision making process. It also compliments existing literature on COVID-19 related hazard pay with quantitative analysis.

Chapter 2 studies the impacts brought by the extensive rise of carjacking cases to the ridesharing service in Chicago. I use a first-differenced model and spatial analysis to unveil the change in labor supply of ridesharing services in response to the rising risk from carjacking crimes. Evidence suggests that ridesharing drivers strategically avoid areas with higher carjacking crime risks, and this risk avoidance behavior is sensitive to hours in the day and holidays.

Chapter 3 explore the indirect economic impacts generated by the prolonged 2015 harmful algal bloom event in California, U.S.A. During this event, harmful algae produced high levels of domoic acid (a neurotoxin) resulting in fishery closures for all commercial Dungeness crab fisheries in California lasting roughly 4.5 months. To estimate the impacts of the event, we investigate whether Dungeness crab prices, both at the ex-vessel and consumer levels, were negatively impacted by the event after the closures were lifted and the crab was declared safe to eat. We find ex-vessel prices fell by at least 9.6% but consumer prices were not impacted. We put forth three competing theories to explain these outcomes and discuss the efficiency and distributional implications of each alternative.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Table Number Page

ACKNOWLEDGEMENTS

I am filled with deep gratitude towards many who have supported, guided, and inspired me throughout the long journey of completing this dissertation. First and foremost, I would like to express my sincere gratitude to my academic committee. I am particularly indebted to Professor Xu Tan, whose mentorship and guidance have significantly contributed to my research. Her patience, understanding, and ability to motivate have been sources of strength and inspiration throughout this process. I can't thank her enough for helping me at the most pivotal moment of my academic journey, and I am highly motivated and inspired by her profound demonstration of support and faith in my potential. I also extend my thanks to the other committee members, Professor Fahad Khalil, Professor Jing Tao, and Professor Hema Yoganarasimhan, for their insightful comments and suggestions that have greatly enriched this work. They are always reachable and supportive. I'm deeply grateful for their advice and time on my research.

My sincere appreciation also goes to my former committee members, Professor Pat Bajari, Professor Yuya Takahashi, and Professor Yanqin Fan. Their initial guidance and contributions laid a crucial foundation for my research, and I am thankful for the time and effort they invested in my progress.

Professor Sunny Jardine deserves special mention for our fruitful and enjoyable joint project. Her collaboration on our project has been nothing short of invaluable. Working alongside her has not only been an academically enriching experience but also a personal pleasure.

I'm also grateful to my friends for providing a supportive and enjoyable environment throughout this journey. I would like to thank my friends Shuo Jiang, Thor Dodson, Kwongyu Wong, Matthew Daniels, and Gani Nurmukhametov. Your friendship and shared moments have been a source of joy and relief during the most challenging times. Your intellectual stimulation and inspiration are also precious gifts.

Finally, to my parents, your unwavering support, love, and belief in me have been the bedrock of my resilience and perseverance. I am deeply grateful for your sacrifices and for always being my source of strength and encouragement.

Chapter 1

TRADEOFF BETWEEN FATALITY RISK AND WAGE: EVIDENCE FROM THE CHICAGO TAXI INDUSTRY IN THE COVID-19 PANDEMIC

1.1 Introduction

Driving a taxi is traditionally considered a dangerous occupation. The risk of violence in the workplace has been highlighted in numerous occupational public health research. However, during the period of the COVID-19 pandemic, taxi drivers have one more risk to be taken into consideration. In 2020, 23.67% of the total deaths from taxi drivers and chauffeurs are deaths from COVID-19, which is the highest among all census occupation titles in 46 states. Similarly, Billingsley, Sunnee, et al (2022) find that compared with other occupations, taxi and bus drivers have elevated mortality rates from COVID-19 in Sweden. Nafilyan et al. (2022) find that taxi drivers have the highest age-standardized COVID-19 mortality rates in the UK.

How do taxi drivers perceive the risk of working in the pandemic? It is shown by Dryhurst, Sarah, et al. (2020) that people perceive more risk if they have had direct personal experience with the virus. James, Kenneth, et al. (2021) find similar results that taxi drivers in Kingston and St. Andrew Jamaica perceive themselves to be at occupationally related risk of COVID-19, and they gather their information largely from media.

The increased risk of working for essential workers leads to heated debate about COVID-19 hazard pay. For example, the city of Seattle finally passed the Grocery Employee Hazard Pay Ordinance, which requires grocery businesses in Seattle to pay hazard pay of $4 per hour to their employees during the pandemic. The taxi industry, however, is another story. A taxi driver is the business owner after the taxi medallion is acquired, and the driver decides the effort to put into the work. In this case, drivers cannot expect compensation from the employer for working under additional risk. They cannot charge a higher price for trips either since it is strictly regulated. Alternatively, they can choose to reduce their working hours, or even quit driving during the pandemic. Traditional methods of estimating the tradeoffs between money and fatality risks do not work in this situation. I propose an alternative approach to estimate the mortality risk premiums using time varying fatality risks in this paper. I also calculate the value of a statistical life and welfare loss from the estimates to provide policy suggestions.

1.2 Related Work

This paper follows a long stream of work on the estimation of the value of a statistical life. Early research like Hamermesh (1977) and Viscusi (1979, 1980) start using a hedonic wage model to estimate workers' tradeoff between wage and perceived dangers in their working conditions. Other studies, such as Thaler and Rosen (1975), Brown (1980), and Leigh (1981) use the Society of Actuaries data set as a measurement of job-related fatality risk instead of risk perception from survey data. From early 1990s, the majority of the studies use the data collected by the government, either collected by the National Institute of Occupational Safety and Health (NIOSH), or from the Census of Fatal Occupational Injuries (CFOI), both led by the U.S. Department of Labor Bureau of Labor Statistics (BLS). The hedonic wage model and cross-sectional analysis are the main framework applied to the study of the mortality risk. One key issue with such kind of studies is the endogeneity and heterogeneity of risk preference across industries. Entrance to the industry with a certain mortality risk is a self-selection process of a worker, and those who with lower-level risk aversion are likely to concentrate on jobs with higher income and higher risk. With the selection bias in the wage-risk tradeoffs estimation, it's reasonable to expect a lower-than-average estimation of a statistical life in those industries. This paper provides an estimation of the wage-risk tradeoff based on time-varying fatality risk instead of cross-sectional variations, avoiding the heterogeneity of risk preference across industries. Additionally, traditional estimations based on the hedonic wage model assume relatively fixed working hours. There are more and more jobs with flexible hours and payments with the development of technology, and the hedonic wage model may not work in this scenario. This paper is among the few innovative studies conducting a wage-risk analysis on an industry with flexible working efforts self-picked by the workers.

Labor supply estimation is another stream of work followed by this paper. After getting their medallion with a fixed cost, taxi drivers self-select their efforts to put into the work. It's close to the case in which the principal sells the company to the worker. The estimation of the labor supply in the taxi industry is a unique stream of flexible labor supply estimation. Camerer et al. (1997) analysis the relationship between hourly income and hours worked on that day using the NYC cab data, and find a negative relationship. They conclude that taxi drivers have income-target labor supply, which means the income effect dominates and drivers stop working when their daily income hits an earning target. Chou (2002) find similar results from a study on taxi drivers in Singapore. However, it has been discussed heatedly later. Farber (2005) use the similar data set as Camerer et al. (1997) and shows that the negative relationship between hourly income and working hours are caused by misspecification, and the income effect on drivers does not significantly impact their stop decision. Stafford (2015) uses data from crab fishermen to analysis the flexible labor supply, and concludes that the negative elasticity found by Camerer et al. (1997) is possibly caused by measurement error in workers' self-selection decisions. This paper echoes the finding of a positive wage elasticity of taxi labor supply. What's more, this paper finds that taxi drivers also make their working decisions across days, in addition to the daily working hour decision analyzed in previous studies.

Lastly, this paper contributes to the emerging literature analyzing the economic impact of the COVID-19 pandemic on the labor market. Byrne, Stephen, et al. (2020) study the initial impact on the labor market in Ireland; Webster et al. (2021) discuss some early impact on a sample of enterprises from southern Europe; and Groshen, E. L. (2020) uses the data from the Bureau of Labor Statistics to investigate the impact of the pandemic on the US labor market. Almost all current work confirms a rapid job loss at the beginning of the pandemic in most countries, and it's followed by a slow recovery. However, while the exogenous initial impact is considerable and well documented, very little research has been done on responses to the pandemic from the labor force after the initial impact. This paper is focused on the medium-term impact of the pandemic instead of the initial impact. It analyzes the wage-risk trade off among taxi drivers to show how the labor force responds to the pandemic when public health restrictions and fatality risk are more of a routine rather than an exogenous shock.

1.3 Background

1.3.1 COVID-19 pandemic and the taxi industry in Chicago

The shelter-in-place public health order was issued and effective from March 18, 2020, in the city of Chicago. Public amenities like parks, beaches, and trails were closed later on March 26 according to the governor's Stay-at-Home executive order. Figure 1.1 shows the average daily new cases per 100 population and average daily new deaths per 10,000 population. There are three spikes in the number of cases and deaths by the end of 2021. The initial spike happened in early 2020, when the virus first hit the city and triggered restrictive public health orders. The second spike was in November and December 2020. It was the surge caused by large gathering during the holiday season together with the Delta variant, which was more transmissible and deadly. The last spike in the data was at the end of 2021. It was likely caused by holiday gatherings and the Omicron variant, which was even more transmissible but less deadly. To fight the virus and help resume the normal activity, COVID-19 vaccines received permissions at the end of 2020, and started to be delivered in 2021. Figure 1.1 also shows the percentage of residents who are fully vaccinated (2-doses) in the city of Chicago. The process of vaccine delivery was slow in the beginning of 2021, but increased rapidly in April and May 2021 when larger group of people are permitted to get vaccines. The city reached the 50% fully vaccinated milestone in July, and it finished another 10% in the rest of the year. With the help of vaccines, the average new case rate and death rate were low during 2021 until the last spike. Although the new case rate of the last spike was higher than the previous two, the death rate was much lower due to the protection from vaccines and the less deadly variant.

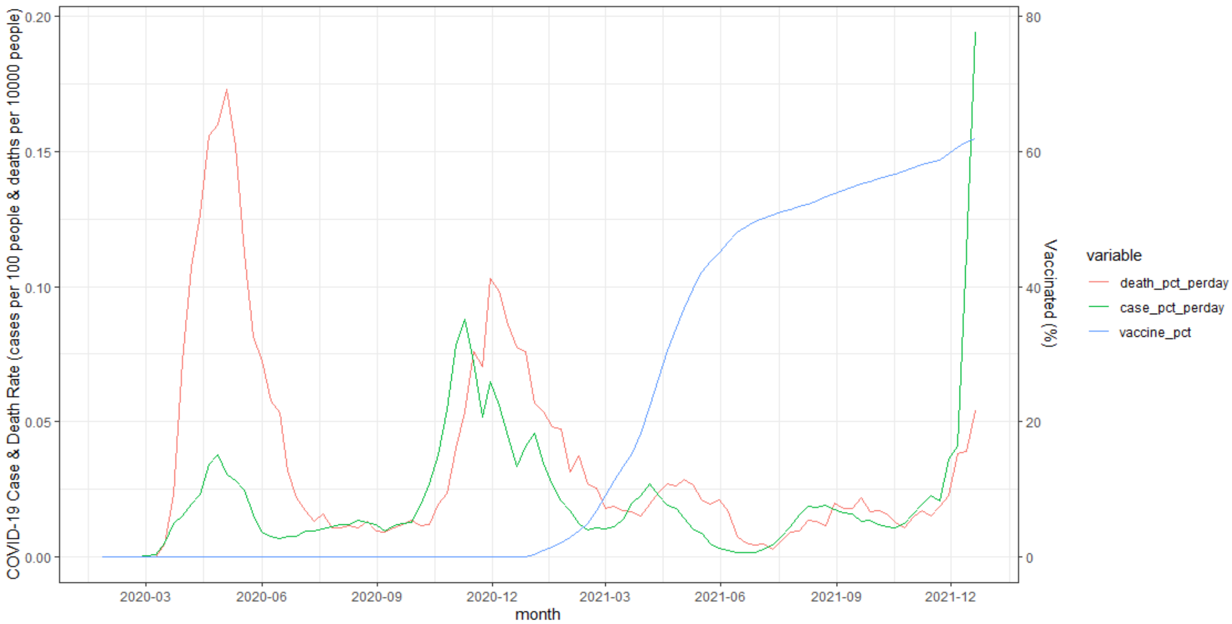


Figure 1.1: Daily case rate, death rate, and accumulated vaccinated rate in Chicago

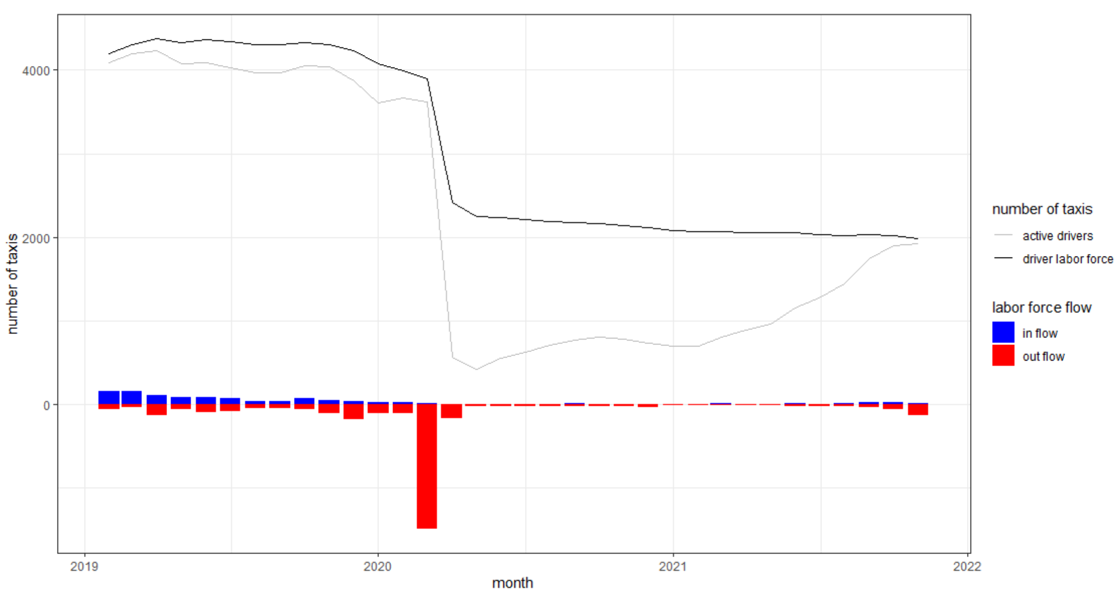


Figure 1.2: Number of active taxis and taxi labor force in Chicago

Note: Blue bars show the number of taxis who's first observed trip is in that month. Red bars show the number of taxis whose last observed trip is in that month.

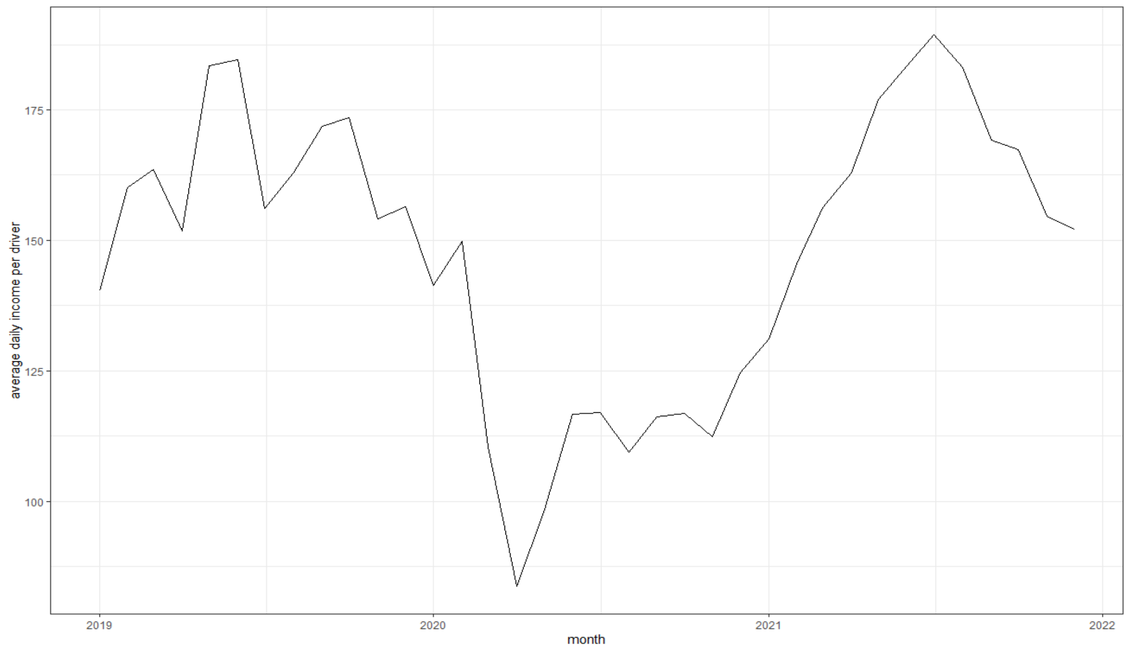


Figure 1.3: average daily income per taxi driver in Chicago

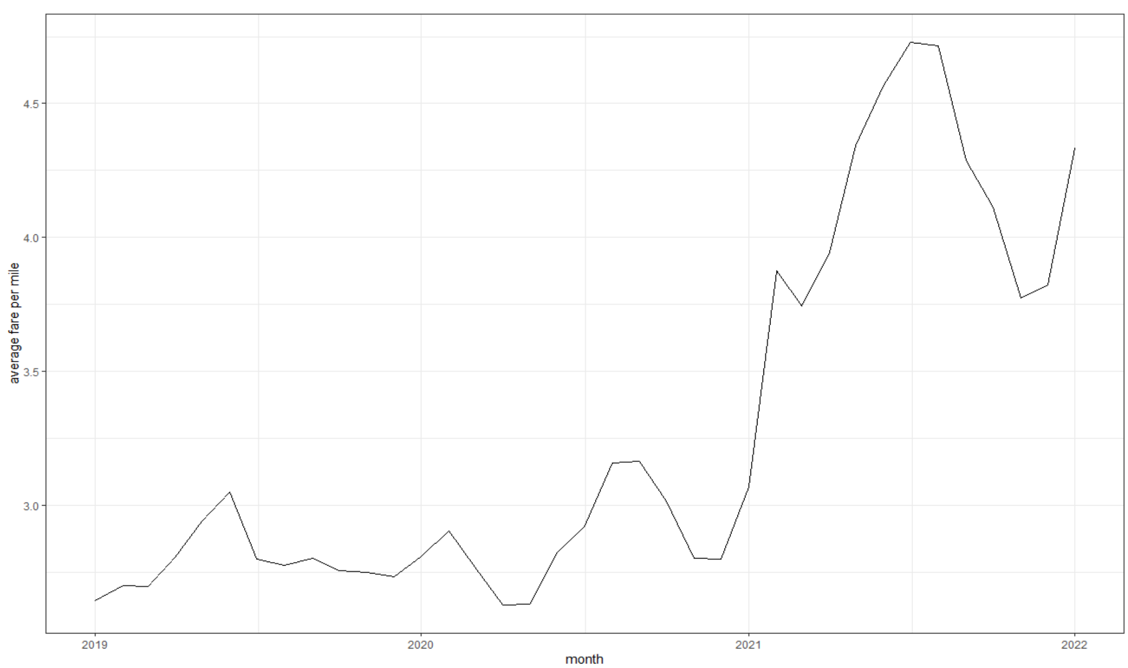


Figure 1.4: Average fare per mile for TNP trips in Chicago

The taxi industry in Chicago has been heavily challenged for years by competitors: transportation network providers (TNP), also referred to as the ride-hail or ride-share companies. The price of a taxi medallion has dropped from over $300,000 to less than $50,000 in recent years. However, the most severe blow in recent years was brought by the COVID-19 pandemic. Figure 1.2 shows the number of active drivers per month from 2019 to 2021. There were around 4000 drivers in 2019, but the number declined sharply by over 80% in March 2020 shelter in place order becomes in effect in the city of Chicago. Later, the number of active taxis was under 1000 until mid-2021, and it later recovers to around 2000 by the end of the year. The red bars show the outflow in the labor. Despite the initial impact of the pandemic, which removes over 1500 taxis out of the market, the number of taxis leaving the market is fairly low after that. It's reasonable to assume that the remaining drivers are more attached to the taxi market. The labor supply of remaining drivers shows time varying features despite the stable number of drivers. It can be the outcome of the market equilibrium where drivers reduce their activity because of lower profitability under restrictive public health orders. It can also be an indicator of drivers' response towards the risk of working during the pandemic. Since the price of taxi trips are fixed, when drivers' reservation wage increases in response to the rising COVID-19 risk, the ones with higher reservation wage choose to stop working. The market is taken by the rest of the drivers, and the average income increases. The incremental part of the income and, more importantly, the incremental part of the reservation wage implied by the absence from the market can be considered hazard pay to those remaining drivers for working under risky conditions. Figure 1.3 shows the average daily income per taxi from 2019 to 2021. The initial impact brings the daily income down to $90 from $150. It's steadily around $110 during the rest of 2020, and recovers to the normal level, or even exceeds it in the second half of 2021. The surge in daily income in mid-2021 is the joint result of the recovering transportation demand and the shortage of drivers.

Not only the number of taxis is cut in half during the pandemic, the number of TNP drivers also decreased dramatically during the pandemic. Unlike the fixed schedule price in the taxi industry, transportation network providers adjust their price according to the supply and demand in real time, so the trip fare reflects the tension of the market directly. Figure 1.4 shows the average fare per mile of all TNP trips in Chicago. The trend of the price surge in 2021 matches the trend of the daily income of taxis in 2021, as they are considered close substitutes. It also shows the difference between price schemes in the taxi industry and TNP industry. The unit price of TNP trips is much higher in 2021 as a response to the shortage of drivers, while the unit price is fixed for taxi trips.

The impact of the COVID-19 pandemic on the economy has been a heated topic. The majority of the study focuses on the treatment effect in the short period when the pandemic first hit the economy. The initial impact does have large effects, however, it's hard to conclude that the effects are solely caused by the pandemic for at least two reasons. The first reason is overreaction. There is a lot of literature pointing out overactions in various markets at the beginning of the pandemic. The taxi industry is likely one of them. The number of trips and daily income dropped dramatically in March 2020 with the initial impact, but quickly recovered in May, exhibiting overactions in the market. The second reason is simultaneous events. When the pandemic first hit the US, different events impact the economy simultaneously. For example, when the case rate and death rate first spike happened in early 2020, the shelter-in-place executive order became effective at the same time. It's difficult to distinguish the impact of the first COVID-19 surge from the impact of the executive order. To address this problem, I focus on a medium term of the pandemic instead, which is the period starting from June 2020 to the end of 2021. The choice of the starting point is based on the fact that the change in average driver daily wage was first stabilized, and the first surge of COVID-19 cases and deaths were mostly finished in July 2020. On one hand, it skips the initial impact to avoid potential overreaction; on the other hand, it contains a time series of covid impacts to help identify the specific pandemic effects. I also include data points from January 2019 to early March 2020 as a baseline without impacts from the pandemic.

1.3.2 Estimation of Hazard pay

There have been lots of discussions on hazard pay for essential workers during the pandemic. Most of them are qualitative research, debating whether essential workers should be given hazard pay. Quantitative research of hazard pay is closely related to the topic of the value of a statistical life (VSL) as both of them measure the tradeoffs between money and fatality risks. Traditional estimates of VSL are mostly cross-sectional analysis. Workers are assumed to seek the best wage-risk combinations among all job offers to maximize their utility, and the labor market equilibrium is formed by the observed wage-risk combinations, which can be concluded as a positive-slop wage-risk tradeoff curve, and VSL can be interpreted as the slope of this tradeoff curve. Hedonic wage analyses are usually used to trace out the acceptable wage-risk combinations. A typical hedonic wage model can be written as the following:

$$ln(w_i) = \alpha + H_i'\beta_1 + X_i'\beta_2 + \gamma_1 p_i + \varepsilon_i \tag{1.1}$$

where w_i is the wage rate of worker i, α is a constant term, H is a vector of worker i's personal characteristic variables, which often includes a variety of human capital measures like age, education, and experience. X is a vector of job specific characteristic variables for worker i, such as industry, management positions, and physical exertion associated with the job. p_i is the fatality risk associated with worker i's job. Estimate of γ_1 in this model measures $\frac{dln(w_i)}{dp_i} = \frac{1}{w} \cdot \frac{dw_i}{dp_i}$. Thus, $\widehat{VSL} = \hat{\gamma}_1 w$. For example, if w_i is measured as the hourly wage, and p_i is measured as the annual fatality per 10,000 workers, then $\widehat{VSL} = 10000\hat{\gamma}_1\overline{w}\overline{h}$ where h is annual working hours and $\overline{h} = 2000$ is a typical value used in the calculation.

Two of the major fatality risk measurements are data collected by the U.S. Department of Labor Bureau of Labor Statistics (BLS) and fatal occupational injuries data collected by the National Institute of Occupational Safety and Health (NIOSH). Such hedonic wage analyses require variation in jobs, wage rates, and associated fatality risk. In the case of the solo industry, the taxi industry for example, there is hardly any variation in job characteristics and associated fatality risk, and the measure of wage rate is tricky as the unit price of the service is fixed. So, traditional hedonic wage analyses are not applicable. The COVID-19 pandemic, however, brings time varying fatality risk to the solo industry case. It provides identification to tradeoffs between money and fatality risks.

1.4 Empirical Models and Strategies

1.4.1 Individual level model

The labor supply of a taxi driver is based on the participation decision and effort decision. I focus on drivers' weekly labor supply, namely the weekly participation decision and the number of active driving days. The labor supply model can be described by the following equation:

$$y_{it} = \alpha + \beta_1 \ln(w_{it}) + \gamma_1 p_t + \gamma_2 c_t + X_t' \beta_2 + H_i' \beta_3 + \varepsilon_{it} \tag{1.2}$$

where y_{it} is the measurement of the labor participation. One measurement is a binary indicator of whether driver i is active in period t. The other measurement is the number of driver i's active days in period t. w_{it} is the wage rate for driver i in period t, measured as average daily income. p_t is the COVID-19 daily death rate per 10,000 population, and c_t is the daily case rate per 100 population. X is a vector of period specific control variables, including a binary indicator of in-pandemic period, a binary indicator of vaccine protection, the number of public holidays within the period. It also includes a set of monthly dummies to control the seasonality. ε_{it} is the unobserved part of the labor decision. H is a vector of driver fixed effects.

When y_{it} is the binary indicator of labor participation, Equation (1.2) is a binary labor supply function. A driver decides to work in that period if the wage is higher than the reservation wage. γ_1 measures the impact of fatality risks to reservation wage. When y_{it} is the number of active days, the equation is a linear-log labor supply curve. γ_1 measures the direct impact of fatality risks on labor supplied.

1.4.2 Aggregate level model

I also estimate an aggregate level to measure the overall impact of the pandemic. The explanatory variables are similar to the individual level model since the total labor supply is simply an aggregation of individual labor participation decision.

$$\tilde{y}_t = \alpha + \beta_1 \ln(\widetilde{w}_t) + \gamma_1 p_t + \gamma_2 c_t + X_t' \beta_2 + \varepsilon_t \tag{1.3}$$

where $\tilde{y}_t$ is the measurement of labor supplied. One measurement is the log of number of active drivers in period t, and the other measurement is the log of number of driver-day combination served in period t. $\widetilde{w}_t$ is the overall driver average daily income in period t. Notation $\sim$ indicates that it is an aggregated variable. p_t, c_t, and X_t' are the same as in the individual level model. The aggregate level model is a constant elasticity labor supply model. γ_1 measures the shift in the supply caused by the fatality risk.

To measure the tradeoff between fatality risk and income, I assume drivers can be monetarily compensated to hold the same level of working efforts when they are faced with additional fatality risk. Then $\left.\frac{dln(w)}{dp}\right|_{dy=0} = -\frac{\gamma_1}{\beta_1}$ is the counterpart of the wage-risk tradeoff slope in Equation (1.1). If p is measured as daily fatality in 10,000 population and w is measured as the daily income, then

$$\widehat{VSL} = -\frac{10000\overline{w}\hat{\gamma}_1}{\hat{\beta}_1} \tag{1.4}$$

where $\overline{w}$ is the average daily income. The estimation of VSL is applicable to both the individual level model and the aggregate level model.

1.4.3 Self-selection issue

One issue in identifying the effects is the self-selection problem in the individual level model. The observed daily income is based on drivers who choose to participate in the market as their reservation wage is lower than the market income. The self-selection issue causes bias in the estimates. Following Stafford (2015), instead of observed daily income, I use predicted income in the individual level model to address this problem. Individual daily income is predicted by the following specification,

$$\text{For each individual taxi } i\text{: } w_i = \theta_i\widetilde{w} + \mu_i \tag{1.5}$$

Intuitively, individual driver's income is predicted as a proportion of the overall average income, and θ_i can be perceived as driver i's ability compared to the average level. The specification is estimated using observed daily income, and then $\widehat{w}_i = \hat{\theta}_i\widetilde{w}$ is generated for all drivers in all periods to estimate Equation (1.2).

1.4.4 Endogeneity of income

The specifications for individual level and aggregate level are indeed supply functions. Endogeneity of the wage rate is a natural problem of estimating a supply function. Using predicted income instead of observed income may address this problem partially as Camerer et al. (1997) argue that overall average wage rate can serve as an exogenous variable and instrument variable for individual wage rate. However, it's still debatable if it's a valid instrument variable. In the sense of estimating the impact from the fatality risk of COVID-19, the endogeneity of income should not affect the estimates on p_t, even if the estimates on $\ln(\widetilde{w}_t)$ is biased due to the endogeneity. However, biased estimates on the estimates on $\ln(\widetilde{w}_t)$ can cause problems in the welfare analyses. The ideal solution is to employ some demand shifters as IVs that work for both the individual level model and the aggregate level model. I choose three demand shifters as IVs to control for the endogeneity of income. One demand shifter is the number of L-train riders in the period. It indicates the shift in the overall commuting demand. The second variable is the average fare per mile from TNP tips in the period. Intuitively, it's the price of a close substitute to taxi trips. Since the fare per mile for taxi trips is fixed, this variable indicates the demand shift between taxi and TNP. The last instrumental variable is the number of "game days" in the period. The city of Chicago holds teams in all major sports leagues. The average attendance of a sports game in Chicago ranges from 20,000 to 80,000. The number of game days indicates event-related demand surge.

1.5 Data

The main data is the taxi trips records from the city of Chicago. I collect all trips that start and end within 77 community areas in Chicago between 2019 and 2021. Each record includes the time and geographic information at the beginning and end of the trip. A taxi id is provided so that individual taxi behaviors can be tracked. However, the data does not indicate whether a taxi is operated by an individual driver or a fleet constructed by multiple drivers. There are 5214 unique taxi id in the data, among which 4985 ids have activity in 2019, but only 2214 remains in 2021.

Another data used is the COVID-19 data reported by the city of Chicago, including daily cases, deaths, and vaccine doses given. For vaccine protection, because the vaccination status of an individual driver is unobservable, I set the 50% fully vaccinated milestone as the starting point of vaccine protection being effective for all drivers. There are two alternative vaccine protection points that I also test. One is the end of May 2021. Taxi drivers are classified as essential works in the transportation department and are scheduled to receive vaccines in phase 1C. According to the city of Chicago, vaccination of phase 1C was mostly finished in April and May. So, it's reasonable to assume that most taxi drivers received their doses by the end of May. Another possible vaccine protection point is early August 2021, when Chicago largely announced that it reached the milestone of 70% adults receiving at least one does. As James, Kenneth, et al. (2021) find that media is an important information source for taxi drivers to receive knowledge of COVID-19, the city's announcement of the vaccination milestone may generate public confidence in vaccine protection.

For supplementary data, the L train riders data comes from the report of the city of Chicago. It's collected on a monthly basis. I first calculate daily average riders, and then aggregate them into weekly data accordingly. So, if the beginnings and ends of two weeks are in the same month, then train riders are the same for those two weeks. In the week of which the beginning and the end are in different months, the number of train riders is the average weighted by the number of days from each month. The average fare per mile from TNP is from the TNP trip data collected by the city of Chicago. It's reported in a similar manner as the taxi trip data, but it does not include a vehicle identifier. In this case, individual driver behaviors cannot be tracked in the TNP trip data. The game day data is collected from the website of sports-reference.com. It provides per game information for major sports leagues. I collect information of games hosted in the city of Chicago for the following leagues: MLB (Chicago Cubs and Chicago White Sox), NBA (Chicago Bulls), NFL (Chicago Bears), and NHL (Chicago Blackhawks).

1.6 Results

1.6.1 Individual level model

I first estimate Equation (1.5) to build predictions of individual daily wage. Figure 1.5 shows the distribution of R^2 and $\hat{\theta}_i$. Overall, the simple prediction model works well, with majority of the regressions having R^2 beyond 0.8. The distribution of $\hat{\theta}_i$ also behaves close to a normal distribution centered at 1. $\hat{w}_{it} = \hat{\theta}_i \tilde{w}_t$ then is used in the following analyses.

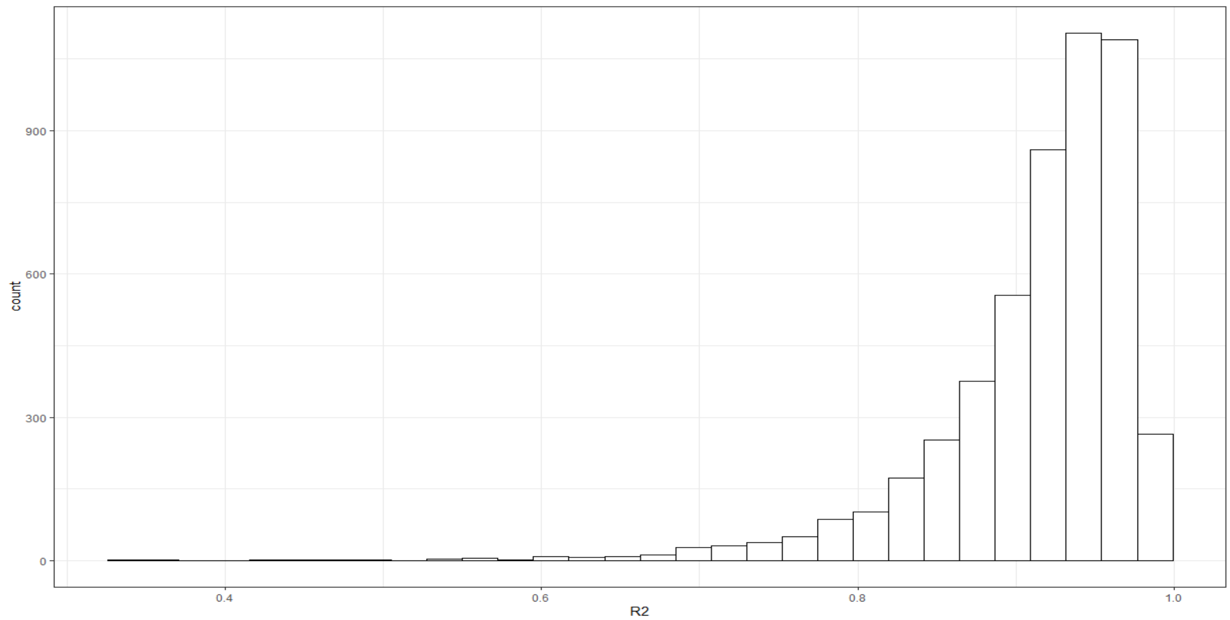
Table 1.1 shows results of estimation on Equation (1.2). The dependent variable in column (1) and (2) is the binary indicator of a driver being active. Specifications are estimated using probit probability function. The dependent variable in column (3) to column (6) is the number of active days provided by individual drivers. Column (3) and (4) show OLS results, and column (5) and (6) show IV regression results. Column (1), (3), and (5) include both pre-covid periods starting from January 2019, and in-covid periods starting from June 2020, while column (2), (4) and (6) only includes in-covid observations. Main coefficients of interest are consistent in their signs across all specifications. Specifically, the wage elasticity of supply is positive; the COVID-19 death rate has a negative impact on taxi labor participation. The case rate shows a positive correlation with labor participation in most specifications, which can be explained by the correlation between social activities and the spreading of COVID-19. Vaccine protection shows a significant positive effect on the level of activity, indicating that vaccination helps the labor market to build confidence and recover during the pandemic. Holiday shows negative effects on the taxi labor supply. Finally, the in-covid period indicator has a significant negative sign as expected, which echoes the well documented initial impact.

Comparing the results between OLS and IV regressions, the estimates on p_t do not vary a lot with or without IV. The estimate on $\ln(\hat{w}_{it})$, however, varies significantly on the in-covid observation, indicating possible endogeneity.

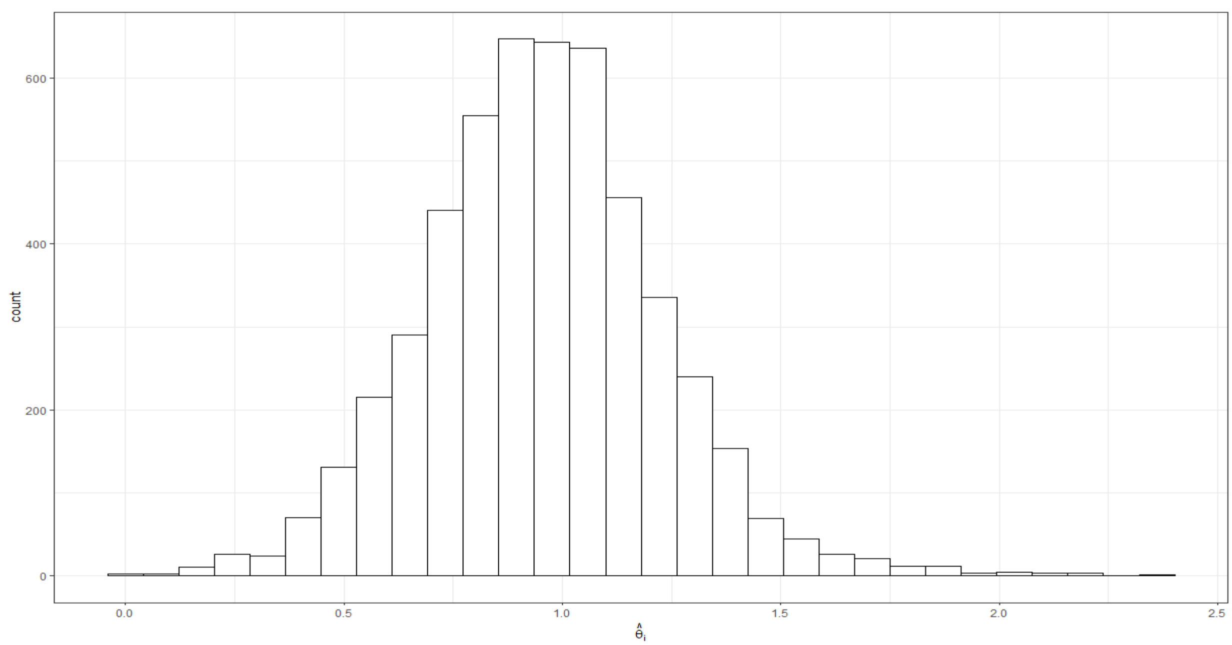
When comparing the results from the full observation and from the in-covid observation, there is a considerable difference between the estimates on death rate and case rate. It may be caused by individual heterogeneity and seasonality before and in the pandemic. Another problem with the individual level model is the model fitness, especially for the in-covid observation. The adjusted R^2 is around 0.064, indicating that most of the variance in the dependent variable is not explained by the model. It's also caused by the individual heterogeneity and seasonality, as the time invariant heterogeneity is being controlled. Without further detailed individual level data, it's very challenging to improve the fitness of the model. Despite the fact that there are fluctuations in the estimates, the results are consistent with the hypothesis that taxi drivers choose their working efforts across days in a week, and it's consistent with the literature that the quantity of labor supplied responses positively to changes in the wage rate.

Table 1.1: Regression results on individual level models

	Dependent variable					
	active		*active_days*			
	Probit	Probit	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(\hat{w}_{it})$	0.335***	0.5427***	0.9789***	0.6932***	0.9396***	0.9476***
	(0.0138)	(0.0261)	(0.0208)	(0.0317)	(0.0264)	(0.0351)
p_t	-1.7064***	-4.4147***	-0.5681**	-6.8184***	-0.6001***	-6.4289***
	(0.1550)	(0.1871)	(0.2215)	(0.2323)	(0.2219)	(0.2335)
c_t	1.4595***	0.897***	2.3208***	-0.4055**	2.2621***	0.0003
	(0.1032)	(0.1181)	(0.1641)	(0.1576)	(0.1658)	(0.1594)
vaccine_50%	0.4763***	0.3424***	0.7093***	0.6609***	0.7204***	0.5736***
	(0.0068)	(0.0109)	(0.0105)	(0.0135)	(0.0114)	(0.0145)
holiday	-0.0125***	-0.0183***	-0.114***	-0.0725***	-0.1156***	-0.0797***
	(0.0037)	(0.0050)	(0.0056)	(0.0064)	(0.0057)	(0.0064)
in_covid	-1.6833***		-3.5219***		-3.5264***	
	(0.0061)		(0.0090)		(0.0092)	
Adj-R^2			0.3777	0.0645	0.3777	0.0643



(a)



(b)

Figure 1.5: Distribution of individual level income prediction estimates

Table 1.2: OLS and IV regression results on aggregate level models

	Dependent variable *ln(# of active taxis)*			
	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
$\ln(\tilde{w}_t)$	0.4888***	1.0225***	0.8564***	1.3112***
	(0.0793)	(0.1216)	(0.1087)	(0.1405)
p_t	-4.3717***	-6.3036***	-4.0722***	-5.8616***
	(0.8434)	(0.8910)	(0.9142)	(0.9334)
c_t	2.1475***	1.8495***	2.6965***	2.3100***
	(0.6247)	(0.6044)	(0.6832)	(0.6374)
vaccine_50%	0.6074***	0.3637***	0.5029***	0.2645***
	(0.0399)	(0.0519)	(0.0471)	(0.0580)
holiday	0.0101	-0.0262	0.0251	-0.0344
	(0.0215)	(0.0246)	(0.0234)	(0.0257)
in_covid	-1.5290***		-1.4871***	
	(0.0344)		(0.0380)	
Adj-R^2	0.9783	0.9440	0.9746	0.9392

Note: Column (1) and (3) use pre- and in-covid data. Column (2) and (4) use only in-covid data.

Table 1.3: OLS and IV regression results on aggregate models

	Dependent variable *ln(active taxi days)*			
	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
$\ln(\tilde{w}_t)$	0.6561***	1.1642***	0.9906***	1.4372***
	(0.0807)	(0.1236)	(0.1090)	(0.1419)
p_t	-4.4497***	-6.6043***	-4.1773***	-6.1864***
	(0.8587)	(0.9050)	(0.9170)	(0.9430)
c_t	1.9700***	1.4781**	2.4695***	1.9135***
	(0.6360)	(0.6139)	(0.6853)	(0.6440)
vaccine_50%	0.5859***	0.3420***	0.4908***	0.2482***
	(0.0406)	(0.0528)	(0.0473)	(0.0586)
holiday	-0.0177	-0.0529**	-0.0041	-0.0606**
	(0.0219)	(0.0250)	(0.0235)	(0.0260)
in_covid	-1.5406***		-1.5024***	
	(0.0350)		(0.0381)	
Adj-R^2	0.9788	0.9464	0.9760	0.9424

Note: Column (1) and (3) use pre- and in-covid data. Column (2) and (4) use only in-covid data.

1.6.2 Aggregate results

Table 1.2 and Table 1.3 show the results of regressions on aggregate variables. In both tables, column (1) and (3) use full period data, and column (2) and (4) use only in-covid data. Column (1) and (2) show OLS estimates while column (3) and (4) show IV regression estimates. The dependent variable in Table 1.2 is the natural log of the number of active taxis, and the dependent variable in Table 1.3 is the natural log of the number of active taxi day combinations. The results are mostly consistent with each. Daily income has a positive wage elasticity of supply and fatality risk from COVID-19 reduce the labor supply. IV regressions still show that there is possibility that the daily income may be an endogenous variable as the estimates on $\ln(\widetilde{w}_t)$ differ between OLS and IV specifications. The estimates on death rate and case rate are more consistent across specifications despite that there are still small fluctuations. It's reasonable as personal heterogeneity and seasonality cancel out when the market labor supply is aggregated. The overall model fitness of the aggregate model is also much improved compared to individual level model. Given the lack of detailed individual driver data, the aggregate level model is considered a more reliable model, and the estimates are used for further analysis.

1.6.3 Check for vaccine protection

The vaccine protection period is set at the 50% fully vaccinated milestone given that individual vaccination records are not available. It's possible that the actual vaccine protection happens earlier or later than the milestone for taxi drivers. Here I consider two alternative vaccine protection starting time. The first one is the end of May 2021 because taxi drivers are allowed to take covid vaccines in Phase 1C, which starts on March 29, and most of the does are delivered in April and May according to the city announcement. So, the end of May is the first alternative vaccine protection starting for taxi drivers. Another milestone that can be used is the 70% 1-dose delivered claimed by the city in August 2021, which is largely announced. Table 1.4 shows the IV estimates on Equation (1.3) using alternative vaccine protection date. Column (1) and (4) are the same as column (5) and (6) in Table 1.3 as benchmarks. Column (1) to (3) use full data and column (4) to (6) use only in-covid data. By comparing estimates across specifications, all estimates keep their signs and most of the estimates are stable. The change in estimates is relative stronger when the

Table 1.4: IV estimates with alternative vaccine protection definition

	Dependent variable *active_days*					
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(\widetilde{w}_t)$	0.9906***	0.7285***	1.1549***	1.4372***	1.0146***	1.544***
	(0.1090)	(0.0999)	(0.0954)	(0.1419)	(0.1746)	(0.1021)
p_t	-4.1773***	-2.5546***	-4.1755***	-6.1864***	-4.6292***	-5.7054***
	(0.9170)	(0.8083)	(0.9140)	(0.9430)	(0.8569)	(0.8998)
c_t	2.4695***	2.1065***	2.4592***	1.9135***	1.2326**	2.0550***
	(0.6853)	(0.5826)	(0.6805)	(0.6440)	(0.5853)	(0.5935)
vaccine_50%	0.4908***			0.2482***		
	(0.0473)			(0.0586)		
vaccine_May		0.5798***			0.4170***	
		(0.0428)			(0.0729)	
vaccine_Aug			0.5062***			0.2632***
			(0.0447)			(0.0467)
holiday	-0.0041	-0.0200	0.0072	-0.0606**	-0.0490**	-0.0574**
	(0.0235)	(0.0199)	(0.0234)	(0.0260)	(0.0224)	(0.0245)
in_covid	-1.5024***	-1.6106***	-1.4722***			
	(0.0381)	(0.0361)	(0.0359)			
Adj-R^2	0.9760	0.9828	0.9760	0.9424	0.9583	0.9483

Table 1.5: IV estimates on aggregate models involving either COVID-19 death rate or case rate

	Dependent variable *ln(active taxi days)*			
	(1)	(2)	(3)	(4)
$\ln(\widetilde{w}_t)$	0.95***	1.1239***	1.3971***	1.7556***
	(0.1107)	(0.1234)	(0.1442)	(0.1882)
p_t	-2.993***		-5.857***	
	(0.8939)		(0.9941)	
c_t		1.5747**		1.9016**
		(0.7267)		(0.8698)
vaccine_50%	0.5239***	0.5154***	0.2646***	0.2262***
	(0.0471)	(0.0520)	(0.0599)	(0.0796)
holiday	-0.0050	0.0030	-0.0587**	-0.0655*
	(0.0242)	(0.0264)	(0.0273)	(0.0351)
in_covid	-1.4923***	-1.5913***		
	(0.0394)	(0.0350)		
Adj-R^2	0.9743	0.9697	0.9361	0.8950

Note: Column (1) and (2) use pre- and in-covid data. Column (3) and (4) use only in-covid data.

vaccine protection is assumed at the end of May. The 50% fully vaccinated milestone and 70% partial vaccinated milestone have similar effects. It provides evidence for the vaccine protection milestone setup that the protection effect reaches a steady level when the vaccination rate hit a certain milestone.

1.6.4 Check for multicollinearity

In the results above, I suggest that the positive effect from the case rate is rooted in the close relationship between social activities and virus infections. However, the high correlation between the case rate and the death rate may contaminate the estimates. One possible result of this multicollinearity problem is that the estimates on the case rate and death rate can go opposite directions while they influence the dependent variable in the same direction if the model is correctly specified. I test the aggregate model with either case rate or death rate is involved. The results are shown in Table 1.5. When involved separately, death rate and case rate do not perform distinctly from that when both are involved. Estimates for both variables hold the direction and the change in magnitude is small considering the high correlation between the two variables. The result suggests that multicollinearity does not contaminate the regressions and estimates for both variables are consistent.

1.6.5 Value of statistical life

One important use case of measuring effects from COVID-19 fatality risks in work is to estimate the hazard pay and the value of statistical life for the taxi industry. The key parameter of VSL estimation in the model is $-\frac{\gamma_1}{\beta_1}$. Table 1.6 shows the estimates of $-\frac{\gamma_1}{\beta_1}$ from different specifications.

Table 1.6: $-\frac{\hat{\gamma}_1}{\hat{\beta}_1}$ from different IV regression results

Specification	individual	aggregate	
Dependent variable	active days	active drivers	active driver-days
Full sample	0.6387	4.7550	4.2169
In-covid sample	6.7844	4.4704	4.3045

Estimates from the individual level model are volatile while those from aggregate level model are consistently lying around 4.5, between 4.2 to 4.8. Using Equation (1.4), these estimates translate to an estimated VSL between \$6.3 million and \$7.2 million with the average daily income per taxi being around \$150. Although this value is lower than the \$11.6 million suggested by the guidance of the department of transportation in 2020, it is still on the same scale and aligned with estimates from other studies. A reasonable reason for the smaller VSL in the taxi industry can be the result of a self-selection process. Working as a taxi driver, highlighted in a lot of occupational public health research, is considered a highly dangerous job. It's not surprising that surviving taxi drivers have a higher risk tolerance, thus lowering their expected mortality risk premiums. In terms of hazard pay, with the fatality risk being 0.05 death per 10000 people per day, a taxi driver would need a compensation between \$31.5 and \$36 per day to keep the normal workload. This hazard pay rate is close to the mandated hazard payment to grocery store workers in Seattle at \$4 per hour given 8 working hours per day.

1.6.6 Welfare analysis

The aggregate level model estimated above is indeed a market labor supply curve. The fatality risk from COVID-19 shifts the labor supply of taxi drivers inwards and reduces the overall social welfare. In a counterfactual situation, where taxi drivers receive proper protection from the fatality risk, the social welfare can be improved. Figure 1.6 illustrates the improvement of the welfare. With the constant elasticity labor supply model, the upper bound and lower bound of the welfare improvement are given by a perfectly elastic demand and a perfectly inelastic demand. Both the upper bound and the lower bound are proportional to the observed market value of the taxi industry, $\widetilde{w}_t y_t(\widetilde{w}_t)$. The ratios are closed form functions on the wage elasticity of labor, β_1, and the reduced fatality risk, Δp_t:

$$\Delta V_{upper} = \frac{1}{\beta_1 + 1}(e^{\gamma_1 \Delta p_t} - 1)\widetilde{w}_t y_t(\widetilde{w}_t)$$

$$\Delta V_{lower} = \frac{\beta_1}{\beta_1 + 1}\left(1 - e^{-\frac{\gamma_1 \Delta p_t}{\beta_1}}\right)\widetilde{w}_t y_t(\widetilde{w}_t)$$

Figure 1.7 shows the relationship between welfare improved and the fatality risk reduced, conditional on β_1 and γ_1. With a 0.05 reduction in the daily fatality rate in 10,000 people, the welfare increases between 10% and 15% of the market value, depending on the elasticity of the labor supply and demand.

The initial impact brings the strongest influence on the market, which is equivalent to increase the death rate by 0.2 to 0.3 per 10000 people. It can at least reduce the total welfare by 30% according to my estimation, and it indeed created a worse consequence. Vaccination, however, partially makes up for it. It improves welfare by around 10% to 20%, depending on the specification. The welfare loss caused by the fatality rate in the medium run is almost covered or even surpassed by the welfare improvement from vaccination. Not only protecting people from the deadly disease, but also helping the labor force regain their confidence in the workplace, vaccine is highly necessary in a speedy manner during a pandemic.

1.7 Conclusion

The taxi labor market is severely impacted by COVID-19 pandemic. The initial shock was considerable, and the subsequent fatality risk in the workplace also has a significant influence on drivers working decision. Fortunately, after the broad delivery of COVID-19 vaccines, welfare loss from the fatality risk is almost canceled, although the market hasn't fully recovered. I investigate drivers' individual working effort decisions and find evidence that drivers make their decisions across days, which is a compliment to the heated discussed in-day decision making process. I also investigate the aggregated taxi labor market and find consistent estimates for VSL and hazard payments. Taxi drivers have a slightly lower risk premium than estimates from other cross-sectional analysis. The framework can be used to estimate VSL and risk-wage tradeoff in other industries with time varying fatality risk, eliminating industry level heterogeneity and selection bias. I also estimate the welfare loss from the fatality risk, which can be reduced if timely protection methods and vaccination can be provided to the labor force.

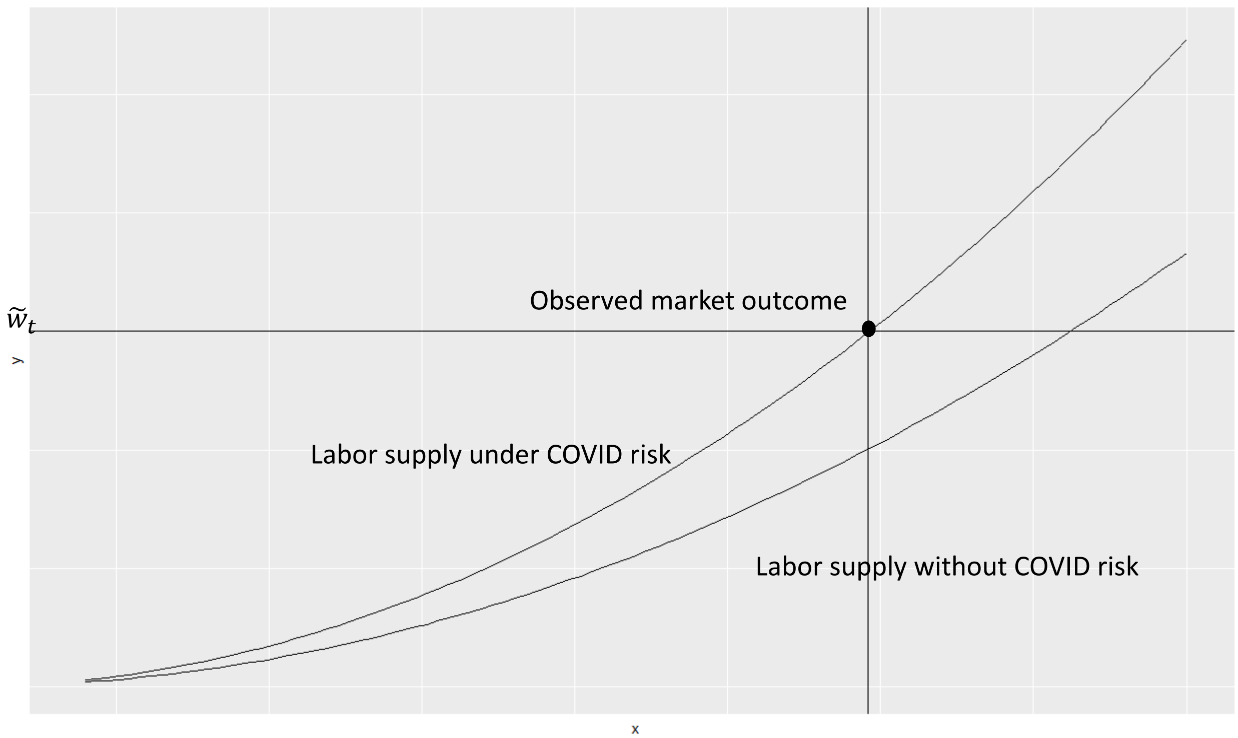


Figure 1.6: Illustration of welfare analysis

Note: the top labor supply curve is the one under COVID fatality risk, and the lower one is the labor supply without such risk. The labor demand curve lays between the perfectly elastic one at $\widetilde{w}_t$ and the perfectly inelastic one. The welfare loss is the area surrounded by the two supply curves and the demand curve.

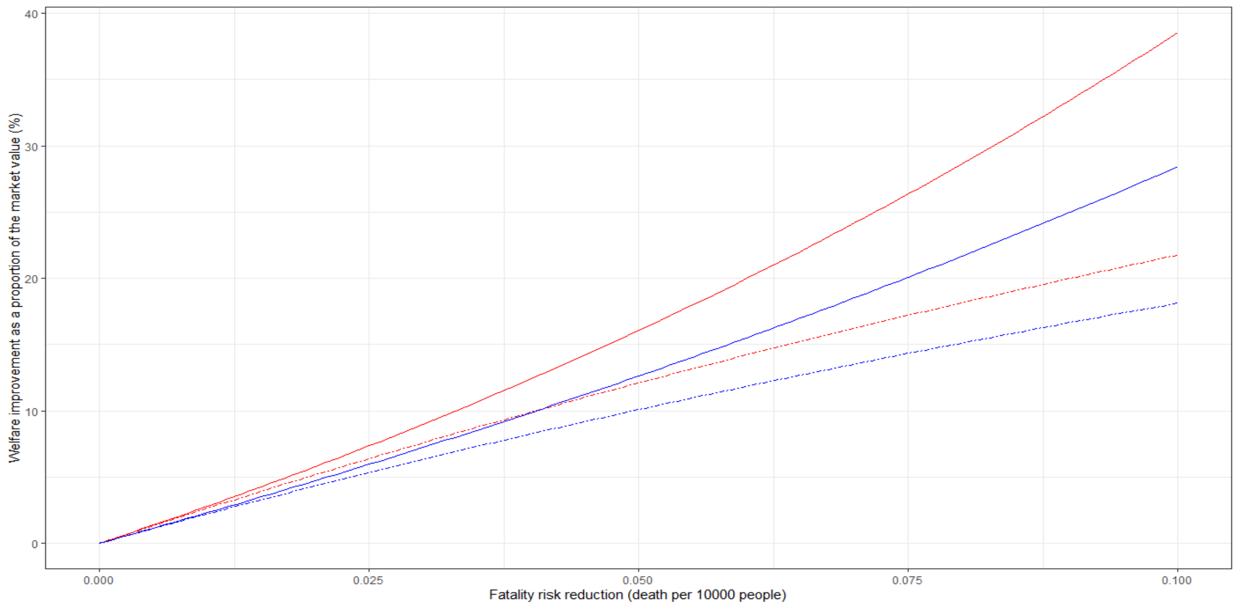


Figure 1.7: Illustration of the relationship between the improved welfare and the reduced fatality risk

Note: $\beta_1 = 1$ & $\gamma_1 = 4.5$ for blue curves, and $\beta_1 = 1.5$ & $\gamma_1 = 6.75$ for red curves. Dashed curves are lower bounds and solid curves are upper bounds.

Chapter 2

RISKY RIDES: HOW CARJACKING RATES DRIVE THE ACCESSIBILITY OF RIDE-HAILING SERVICES AND SURGE PRICING IN CHICAGO

2.1 Introduction

In the landscape of modern urban mobility, ride-hailing services such as Uber and Lyft have emerged as significant players in public transportation systems. Compared with taxis, their traditional counterparts, they offer a higher level of convenience and accessibility to passengers. They have also introduced unprecedented flexibility to the labor market by enabling drivers to register easily as service providers, bypassing the often-cumbersome process of acquiring a medallion, a requirement for taxi operations in many cities. With those advancements, the rapid growth of the ride-hailing industry is witnessed in cities worldwide.

Nevertheless, despite these advancements, the ride-hailing industry and its traditional counterpart, the taxi industry, grapple with common challenges. Foremost among these is crime, particularly carjacking incidents, which pose a significant safety threat to drivers. Given the nature of their roles, both ride-hailing and taxi drivers are often required to work at night, frequent potentially unsafe locations, and regularly engage with strangers. This unique occupational landscape amplifies their exposure to elevated risks and potential criminal activities. Consequently, the risk avoidance behaviors adopted by drivers become an essential aspect of their service provision.

Interestingly, this tradeoff between risk and service provision is further complicated in the ride-hailing industry due to the implementation of dynamic pricing models. Distinct from the taxi industry, ride-hailing companies typically employ 'dynamic pricing' systems that increase service prices in areas where driver availability is limited. In high-risk locations, these dynamic pricing systems present drivers with a tradeoff between risk and reward. While drivers may be inclined to avoid these areas to minimize their exposure to criminal activities, fewer drivers willing to serve these locations can drive up the price of services due to increased demand, thereby raising the potential returns. Hence, dynamic pricing systems can inadvertently enhance service accessibility in less desirable or higher-risk locations.

However, the potential influence of drivers' navigational choices, shaped by the risk of victimization, on service provision and pricing, is currently underrepresented in the literature. Past studies were often restricted to the traditional taxi industry, where regulatory measures typically fixed prices, leaving limited scope for quantitative analyses. Dynamic pricing in the ride-hailing industry, opens a new avenue to examine drivers' risk-avoidance behaviors. This paper intends to explore these behaviors and their potential impacts by concentrating on the ride-hailing industry in Chicago, a city that has experienced an alarming rise in carjacking incidents from 2020 to 2022. This increase has sparked significant concern among ride-hailing drivers, likely leading to adaptations in their behaviors and operational strategies. As such, the intricate relationship between urban crime, driver behavior, and the overall provision of ride-hailing services presents a critical field for investigation.

I propose the hypothesis that safety concerns may lead ride-hailing drivers to deliberately avoid accepting rides from areas with high crime rates. If this behavior is pervasive, it could result in higher service prices in these areas due to supply-demand dynamics influenced by a lack of available services. The potential implications of this scenario are significant from both social and economic perspectives. Neighborhoods with high crime rates could be systematically underserved, causing residents to bear the brunt of elevated transportation costs. By exploring these dynamics within the context of Chicago's ride-hailing market, this study aims to provide a more profound understanding of the impact of urban crime on the delivery and pricing of ride-hailing services.

2.2 Related work

The emergence of the sharing economy has sparked significant interest, given the substantial impact it's made on traditional industries. Belk (2014) draws a comparison between the sharing economy and collaborative consumption (e.g., Uber versus Zipcar), two innovative business models, and concludes that they both contribute to reducing ownership and increasing flexibility.

Such new business models enhance efficiency significantly. Sundararajan (2017) discusses how the sharing economy, beyond transportation, has altered perceptions about employment and capitalism, potentially destabilizing the workforce of traditional industries. The study emphasizes that companies such as Uber and Airbnb, which pave the way for a new form of crowd-based capitalism, differ from earlier e-commerce companies like eBay and Craigslist.

Focusing on ride-hailing services, Cramer and Krueger (2016) quantitatively evaluate the capacity utilization rates for both Uber and taxi drivers in four major cities: NY, Boston, Seattle, and LA. They conclude that Uber more effectively matches individual drivers and passengers, and it better synchronizes market supply and demand throughout the day due to its flexible labor supply model and surge pricing strategy, leading to disruptive changes in the taxi industry. Burtch, Carnahan, & Greenwood (2018) exploit a natural experiment concerning the entry of Uber X and Postmates services against local crowdfunding campaigns launched on Kickstarter. They suggest that gig economy platforms mainly diminish lower-quality entrepreneurial activity by offering viable employment for the un- and under-employed. Wallsten (2015) performed a before-and-after analysis of taxi complaints against Uber's entry into New York and Chicago. A decline in taxi complaints upon Uber's entry is observed in both cities, resulting from unsatisfied passengers opting for new alternative services. In response, taxi drivers enhance their service quality to secure their market share.

Some studies illustrate that lower cost is a key factor to the success of the sharing economy. For example, Einav, Farronato, & Levin (2016) develop a theoretical model to illustrate how peer-to-peer markets enable entry by small or flexible suppliers, concluding that this design can reduce entry cost and subsequently, market prices.

Meanwhile, several studies attribute the success of ride-hailing services to the flexible work arrangements from the driver's perspective. Lee et al. (2015) conduct a qualitative study involving interviews with ride-hailing drivers and passengers. Their study reveals that drivers highly appreciate the flexibility offered by ride-hailing platforms, and increased transparency in the trip assignment algorithm encourages driver cooperation. The passenger-driver rating system fosters basic trust and service attitudes in the ride-hailing market. Chen & Sheldon (2016) study how Uber drivers respond to dynamic pricing of trips, finding that drivers drive more during high-earning periods in contrast to income-target findings in Camerer et al. (1997). Shokoohyar (2018) analyzes 7183 online reviews from drivers and discovers that job flexibility and work-life balance are the primary reasons they are attracted to work as ride-hailing drivers. However, the job also has major drawbacks, including insufficient compensation, poor job security, bad rider behaviors, and poor customer services.

Research also focuses on the consumer side. Rayle et al. (2016) conducted a survey in San Francisco, discovering that ride-hailing services attract younger, well-educated users seeking short wait times and quick point-to-point service. Alemi et al. (2018) analyze the California Millennials Dataset to investigate factors affecting the adoption of ride-hailing services. They find that increased land-use mix, car accessibility, higher education, older millennials, pro-environmental, technology-embracing, and variety-seeking attitudes are associated with greater likelihood of adopting on-demand ride services.

In terms of connections between ridesharing industry and crimes, the focus of research has been on the impact brought by the emerging ride-sharing industry. Martin-Buck (2017) finds introduction of ridesharing service significantly reduce alcohol-related auto accidents and DUI/DWI arrests. Greenwood and Wattal (2017) also find a significant drop in alcohol related motor vehicle fatality rate using a difference-in-difference approach. Dills & Mulholland (2018) find reduced fatal vehicular crashes associated with the introduction of ride-sharing services, they also find no significant changes in the local victimization rate of robbery, assault, or drunkenness. However, they document that arrest rate for motor vehicle thefts rises with the entry of ride-sharing services. Weber (2019) further finds that the significant crime reduction associated with ride-sharing services is almost entirely composed of alcohol-related assaults. Park et al. (2016) find that Uber availability reduces the likelihood of sexual assaults, mostly in taxi-sparse areas. On the contrary, there are also studies finding that ridesharing services can be associated with increased risk of crime. For example, Mehranbod et al. (2022) find that ridesharing is positively correlated with nighttime assaults at bars. Despite the thick literature on the impact of ridesharing services on crime, the risk of crime faced by ridesharing drivers is largely overlooked. Almoqbel and Wohn

(2019) conduct a qualitative study on drivers' concerns on unsafety working environments using in-depth interviews with drivers. They find that drivers select working times and locations strategically to avoid unsafe areas or situations as one of their strategies to reduce the risk of victimization. This paper contributes to filling the gap of quantitative analyses on ridesharing drivers' strategical behaviors reacting to the risk of crime.

Spatial segregation is another significant topic in ride-hailing industry research. Hughes & MacKenzie (2016) use a spatial error regression model to estimate the effects of different socioeconomic indicators on the expected wait time of ride-hailing services in the Greater Seattle area. They find that geography plays an essential role in ride-hailing service markets and creates systemic biases. Thebault-Spieker et al. (2017) also conclude that the sharing economy is more effective in dense, high socioeconomic status areas than in low socioeconomic status areas and the suburbs.

This article provides the following novel insights into the literature on ride-hailing services:

1) It focuses on the specific type of crime that affects ride-hailing drivers' work provision.
2) It introduces evidence that ride-hailing drivers consciously reduce pickups in areas with elevated carjacking rates.
3) It quantitatively assesses the influence of drivers' risk-avoidance behavior on surge pricing.

The rest of this paper is organized as follows: Section III introduces the data used in this study, offering an overview of the prevailing circumstances; Section IV delves into the empirical analysis, discussing the methodology and resultant findings; finally, Section V provides a comprehensive conclusion to the study.

2.3 Data

The data used in this study is mainly from Chicago Data Portal, which is an open data portal released by the city of Chicago. There are two major components: crime records and ride-hailing trips.

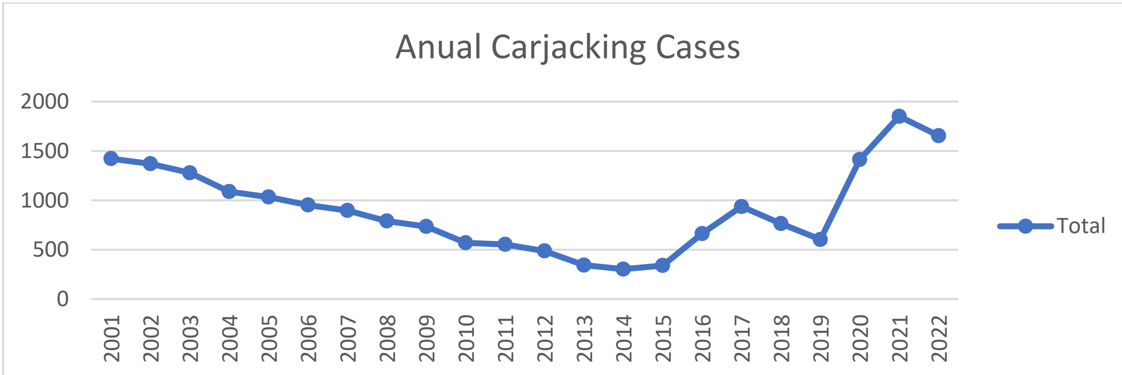


Figure 2.1: Annual Carjacking Cases from 2001 to 2022.

2.3.1 Crime records

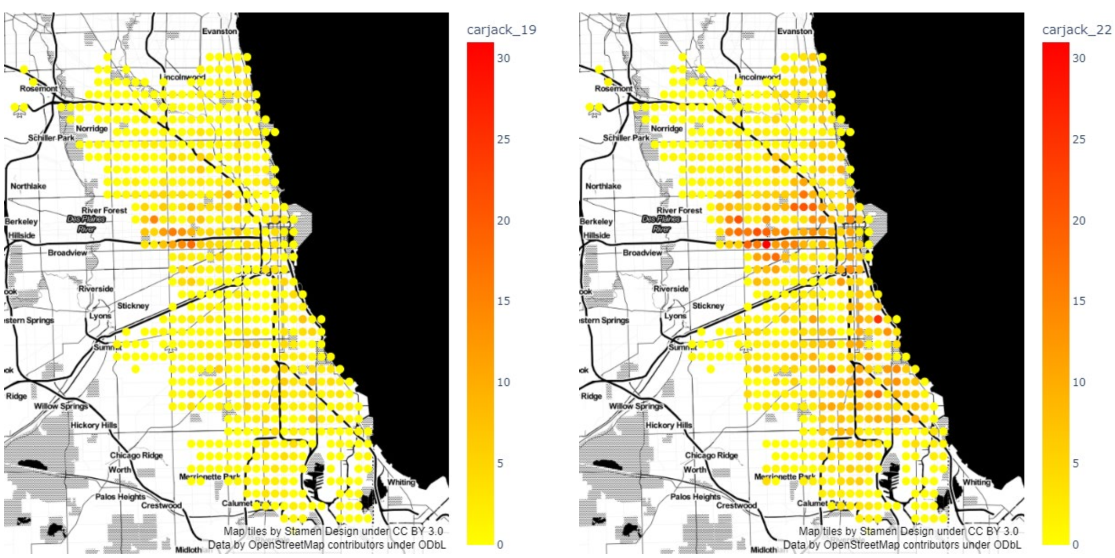
The City of Chicago has been documenting crime incidents since 2001. Each record includes specifics such as the time and location of the incident, along with the FBI classification code and descriptions per the Illinois Uniform Crime Reporting (IUCR) code. This study focuses on the specific classification of carjacking (IUCR: 0325, ROBBERY: VEHICULAR HIJACKING) as the principal measure of victimization risk to ride-hailing drivers.

As shown in Figure 2.1, the city experienced a dramatic surge in carjacking cases in 2020, with figures more than doubling compared to the previous year. This alarming trend continued into 2021, reaching the highest recorded levels of carjacking incidents in two decades. Amidst this escalating trend specifically targeting vehicles and drivers, ride-hailing drivers are confronted with substantial risks when providing services to strangers. Despite their appeals for more effective countermeasures from both the government and ride-hailing platforms, there have been minimal improvements. Consequently, drivers are unable to overlook the risk of falling victim to carjacking incidents and may have recalibrated their working strategies to navigate this challenge.

The data is prepared for the analysis as follows: All records are aggregated on geographic locations and incident time slots. The locations of crime incidents are regrouped based on GPS latitude and longitude up to two decimal places in degrees, each representing roughly an area of 0.92 km² (0.83 km x 1.11 km). To avoid unnecessary data sparsity in this study, incident times are consolidated into four-hour blocks: 7 to 10 are categorized as morning, 11 to 14 as noon, 15 to 18 as afternoon, 19 to 22 as evening, 23 to 2 as midnight, and 3 to 6 as predawn. Workdays and holidays are grouped separately, with holidays encompassing all weekends and public holidays in Chicago. The study focuses on two periods: 2018 to 2019 representing the pre-period prior to the surge in carjacking cases, and 2021 to 2022 as the post-period.

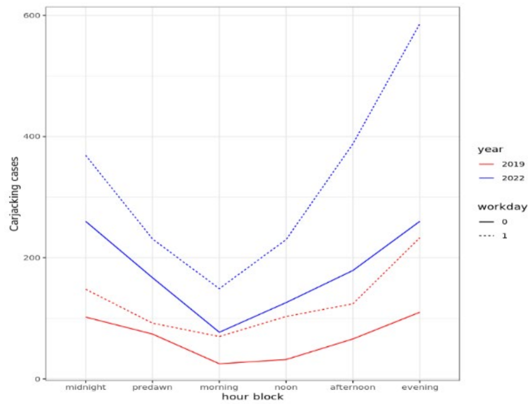
Figure 2.2 maps out the distribution of carjacking incidents during the pre- and post-periods. West Garfield Park, a hotspot for carjacking incidents in the pre-period, remains a significant area for such incidents in the post-period. However, it's not the only area affected. Other regions, such as the University of Chicago and downtown Chicago, have also seen a noticeable increase in carjacking incidents. This figure highlights the changing landscape of carjacking incidents across the city over time, underscoring the widespread nature of this issue.

Figure 2.3 illustrates the intra-day seasonality. Total carjacking incidents during the post-period have approximately doubled, while the total incidents of other crimes have remained at a similar level as during the pre-period. The diurnal pattern for both carjacking and other crimes is analogous; most incidents occur during the evening and midnight, with significantly fewer cases in the morning or noon.

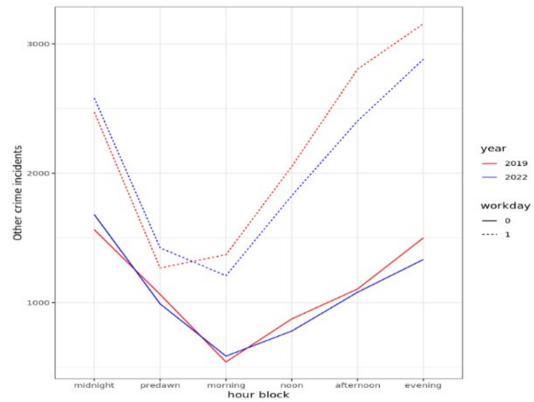


(a) 2018-2019 (b) 2021-2022

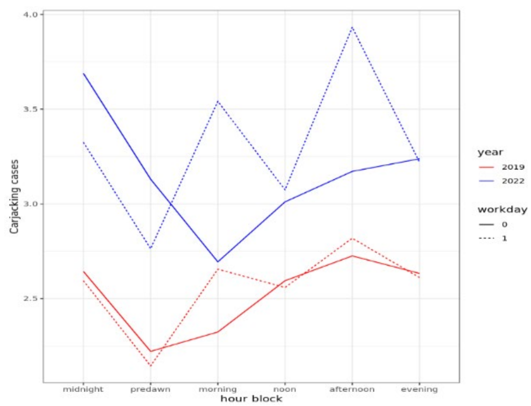
Figure 2.2: Spatial Distribution of Carjacking Cases



(a) Intra-day Carjacking Cases



(b) Intra-day Other Crime Cases



(c) Intra-day Average fare per mile

Figure 2.3: Intra-day Seasonality of Carjacking and other Crime Cases, and Average Fare per Mile

2.3.2 Ride-hailing trips

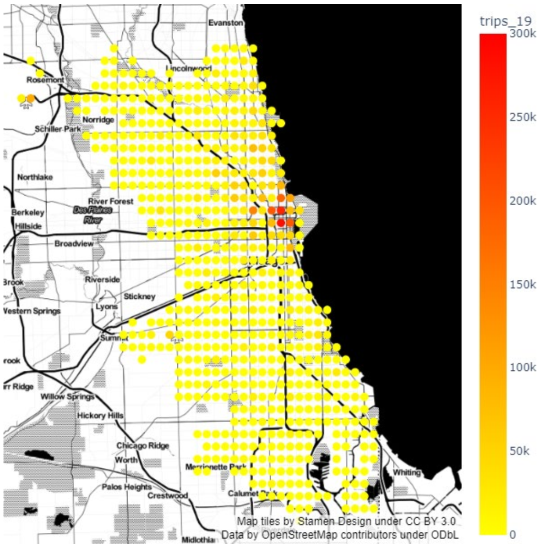
The City of Chicago has been collecting data on ride-hailing services, also known as Transportation Network Providers, since late 2018. This dataset captures various key attributes of each individual trip within the city, such as pickup and drop-off locations and times, trip fares and fees, trip distances, and whether the trip was shared or solo. In this research, we specifically concentrate on the dynamic pricing element of ride-hailing services. To maintain comparability across observations, shared trips are excluded from this study as their distributed fares among multiple unidentified pooled trips make them incomparable to solo trips.

The time periods considered for this analysis are October 2019 (pre-period) and October 2022 (post-period). The year 2019 was selected as the pre-period as it represents the latest available data prior to the onset of the COVID-19 pandemic in the U.S., while October 2022 marks a period when the pandemic had largely transitioned into an endemic phase. The holiday months of November and December are purposely excluded to avoid potential distortions from holiday seasons. The association of crime records with ride-hailing trip data has been executed in a manner that matches aggregated crime data from 2018-2019 with the October 2019 trip data, and similarly, aggregated 2021-2022 crime data is associated with the October 2022 trip data based on corresponding locations and time blocks.

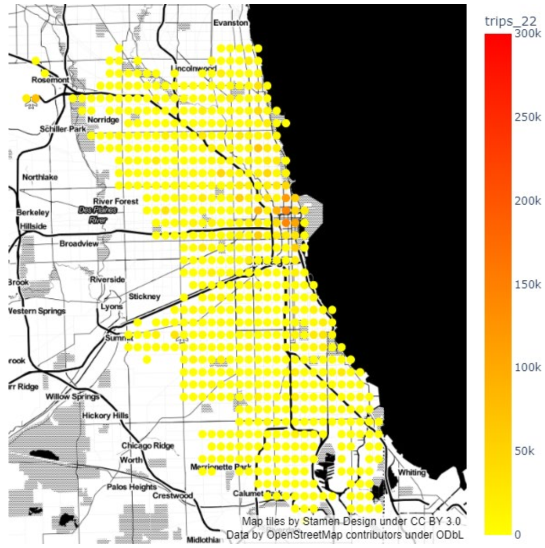
As depicted in Figure 2.3(c), there was a significant rise in the average fare per mile in the post-period. This increase can be attributed largely to a reduced labor supply due to the pandemic, with the overall driver labor force falling by approximately 50%. With the resumption of normal daily activities, the diminished driver labor force was insufficient to meet the resurging demand, prompting ride-hailing platforms to increase service prices via their dynamic pricing mechanisms.

Furthermore, in 2022, Uber introduced a feature allowing drivers to view multiple trip requests simultaneously, a change that was soon adopted by Lyft as well. This adjustment provides drivers with a degree of discretion in choosing among different trips. Given this newfound flexibility and the existing labor shortage, drivers are presented with substantial opportunities to employ strategic behaviors to enhance their personal welfare.

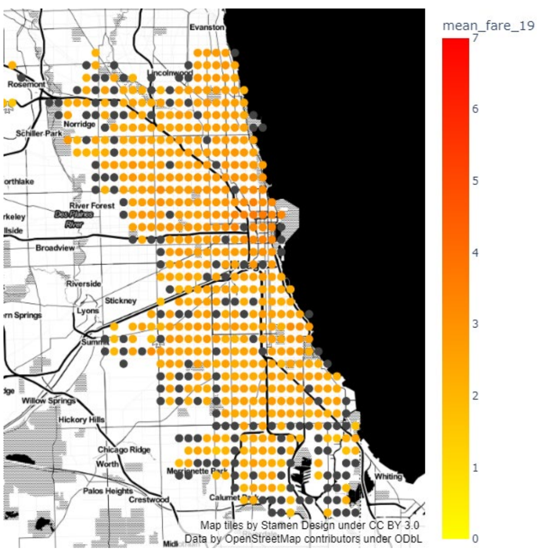
Figure 2.4 visually depicts the geographic distribution of ride-hailing trip origins. As expected, the majority of trips originate in downtown Chicago. However, following the pandemic, there is a considerable decrease in this volume. It also illustrates the shift in average fare per mile between 2019 and 2022, highlighting a notable price increase of ride-hailing services in two specific areas. Firstly, the downtown area, which represents a key nucleus for ride-hailing pickups, and secondly, the region around the University of Chicago. These areas of price surge coincide with a significant rise in carjacking incidents during the same period, thus revealing a likely correlation. Conversely, the area of West Garfield Park, while hosting the highest total volume of carjacking incidents throughout the period examined, did not experience a comparable escalation in ride-hailing prices.



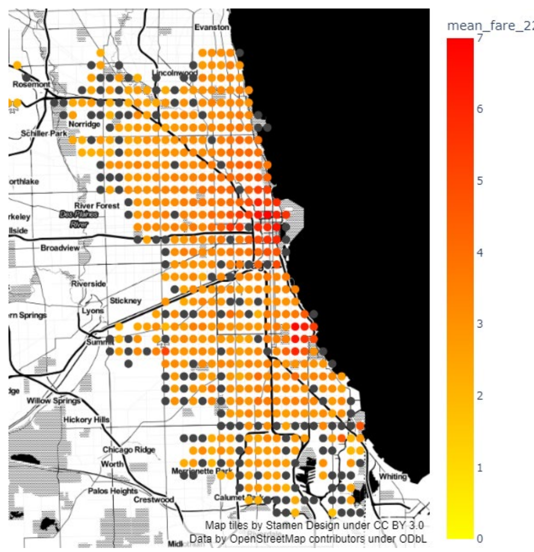
(a) Oct 2019 Ride-hailing Trips



(b) Oct 2022 Ride-hailing Trips



(c) Oct 2019 Average fare-per-mile



(d) Oct 2022 Average fare-per-mile

Figure 2.4: Spatial Distribution of Ride-hailing Trips and Average fare-per-mile

A potential explanation for this discrepancy could lie in the pre-existing reputation of West Garfield Park. Prior to the 2020 carjacking wave, the area was already considered one of the most hazardous to drive in, leading to it being significantly underserved by ride-hailing services. The surge in carjackings in 2020 impacted the safety perceptions of other areas, turning them from safe to less safe.

However, it had a lesser impact on West Garfield Park as its reputation as a high-risk area was already firmly established. Hence, alterations in the strategic behaviors of ride-hailing drivers are most apparent in areas that have transitioned from being traditionally perceived as safe to now being considered less so.

Table 2.1: Summary statistics of level variables

	Workday				Non-workday			
	Pre-period (2019)		Post-period (2022)		Pre-period (2019)		Post-period (2022)	
Variables	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev
avg_fare_per_mile	2.569	0.434	3.317	1.048	2.534	0.450	3.157	1.079
avg_trip_fare	11.323	2.455	14.529	3.606	11.281	2.469	14.036	3.178
# of trips	1773.084	5438.388	1128.812	3070.531	886.329	2396.182	812.163	2148.023
avg_miles	5.738	1.972	6.069	2.148	5.850	2.053	6.225	2.242
# of carjackings	0.283	0.628	0.718	1.197	0.153	0.447	0.400	0.775
# of other crimes	4.824	5.801	4.530	5.271	2.488	3.714	2.414	3.276

Table 2.2: Summary statistics of first difference variables

	Workday		Non-workday	
Variables	Mean	Std.dev	Mean	Std.dev
delta_avg_fare_per_mile	0.229	0.220	0.190	0.239
delta_fare	0.244	0.199	0.218	0.215
delta_trips	-0.460	0.629	-0.184	0.648
delta_miles	0.052	0.244	0.058	0.277
delta_carjack	0.207	0.525	0.138	0.443
delta_other_crime	-0.025	0.628	0.012	0.651

Note: all first differences are calculated on logarithm
i.e., delta = log(2022 value) - log(2019 value)

2.4 Methodology

2.4.1 First Difference Structure

A conventional spatial model often applied in relevant literature employs demographic variation to discern the provision of ride-hailing services. It can be represented as:

$$Y = \beta_0 + \beta_1 X + \varepsilon \tag{2.1}$$

where Y represents the variable of interest (for instance, fare, wait time, number of trips, etc.), while X represents a range of demographic characteristics. These typically include factors such as population density, minority representation, age demographics, gender ratios, education levels, income brackets, and crime rates, among others.

However, the application of this model presents certain challenges. Firstly, a trade-off invariably exists between spatial granularity and data completeness. This challenge arises due to privacy protections and resource limitations inherent to the census or survey processes which generate demographic data. Secondly, the spatial distributions of the aforementioned demographic characteristics often exhibit unevenness or incomparability across different areas, which may compromise the model's validity. Thirdly, the model might contend with unobserved fixed effects when deployed in cross-sectional spatial analyses, introducing potential bias to the results. Lastly, these demographic characteristics often exhibit limited temporal variation or are lagged in their updates. This can make it challenging to accurately identify time-sensitive effects, particularly those requiring high-frequency analysis.

In order to address unobserved fixed effects and detect time-sensitive variations, this study constructs a panel data framework and implements a first difference model. This model is expressed as:

$$\Delta y_i = \beta_0 + \beta_1 \Delta x_i + \beta_2 \Delta q_i + \gamma M_i + H + \varepsilon_i \tag{2.2}$$

where Δy_i is the metric that represents the pricing of ride-hailing services. This study concentrates on average fare per mile as the measurement of price. i indicates the location. Δx_i is the change in local carjacking cases, Δq_i is the change in the number of local ride-hailing trips, and M_i is a set of control variables. It contains a set of 2019-level control variables, including 2019 levels of: average fare per mile, average trip fares, number of trips, average trip miles, cases of carjacking. It also includes the change of average trip miles as an important component in the service price. H is a set of time block fixed effects. It's anticipated that spatial fixed effects would be eliminated during the first-difference process. To mitigate the issue of scale, all variables, apart from fixed effects, are expressed in logarithmic form. As a result, first difference variables measure percentage changes in this model, and β_2 is correspondingly the inverse of the price elasticity of the ride-hailing supply.

2.4.2 Instrumental variables

There are potential omitted effects or measurement errors in the estimation, which lead to bias towards the estimates on quantity supplied. It's noticeable that the model is essentially a reversed supply function, and it's a common issue in the estimation of demand or supply functions that quantity and price are considered endogenous variables.

Besides quantity supplied, there are several potential sources of endogeneity on the crime variables as well. There can be some unobserved factors that affect Δx_{it} and Δy_{it} simultaneously. For example, carjackings may happen more frequently in low-income areas, where the demand for ride-hailing is lower. So, this unobservable may be negatively correlated with the change in price, and thus underestimates the effect of carjackings on ride-hailing price. On the other hand, carjackings can be more frequent in busy areas where the market is thicker with more passengers and drivers, making it easier to find a victim there. Then, when the service price is higher in the busy area, this unobservable may overestimate the effect. To consistently estimate the supply function and the effects from carjacking, instrumental variables are needed. The designated instrumental variables should on one hand serve as demand shifters so that the supply function can be identified, on the other hand be correlated with carjacking so that it can predict carjacking properly.

I introduce two demand shifters into the model: change in other crimes, and 2019 level of other crimes. The choice of instrumental variables is based on the following rationale. Firstly, it is documented that people are more likely to choose ride-hailing or taxies services when the location or time slot is associated with higher crime rate, which makes the variables good demand shifters. Secondly, the spatial distributions of carjacking and other crimes are correlated. Areas with a higher rate of other types of crimes are usually more dangerous and have more carjacking cases. Finally, crimes other than carjacking are less likely to affect the labor supply in ride-hailing services. Drivers usually run their vehicles around and do not necessarily stay at the pickup location for a long time. They also spend most of their time in their vehicles, isolated from many kinds of criminal activities. The victimization risk of carjacking is much more concerning than other types of crimes to those drivers. Given those rationales, the selected instruments should handle the discussed endogeneity.

2.4.3 Spatial dependency

Another common concern in the analysis of spatial data is spatial dependency. It is widely observed in various disciplines ranging from environmental science to economics. Spatial dependency refers to the phenomenon where the activity in one location is affected by the activities in its neighboring or nearby locations. This concept is rooted in the idea that spatial proximity often breeds resemblance and influence, leading to clusters or patterns within the spatial data that are not random but inherently structured. The existence of spatial dependency usually violates the independence assumption in many classical empirical models as nearby units behaving in non-random arrangements. Therefore, analyzing spatial data without considering spatial dependency can lead to erroneous inferences.

Spatial dependency can manifest in two principal ways. One is Spatial Lag, which is also known as Spatial Autocorrelation. This occurs when a variable at one location depends on the values of that variable at neighboring or nearby locations. For example, in the case of ride sharing market, service price in one location may be influenced by the prices of nearby locations due to some spill over effects. The other one is Spatial Error (Autocorrelation), which arises when there are unobserved variables that have a spatial structure and influence the dependent variable, but are independent of other explanatory variables. For example, if local income level, which is unobserved in this study, affects the ride share price, and high/low-income neighborhoods are clustered (spatially correlated), then ride share price may exhibit spatial error correlation. Ignoring either dependency while they exist yields different issues in empirical study. Ignoring Spatial Lag dependency would lead to an omitted-variable-bias situation where the final estimates are inconsistent and biased. Ignoring Spatial Error, on the other hand, would not result in inconsistent or biased estimates, but the statistical power would suffer because the estimation is inefficient.

I use a combined spatial lag and error model to address the spatial dependency once detected. More specifically, the spatial model is as follows:

$$\Delta y_i = \beta_0 + \lambda W_i \Delta y_i + \beta_1 \Delta x_i + \beta_2 \Delta q_i + \gamma M_i + u_i \tag{2.3}$$

$$u_i = \rho W_i u_i + \varepsilon_i \tag{2.4}$$

It's different from the specification in Equation (2.2) on two parts. Firstly, it includes a spatial lag component, $\lambda W_i \Delta y_i$ to accommodate the spatial autocorrelation on Δy_i. W_i is the spatial weight matrix formulating the correlation between nearby locations. Secondly, this specification allows for spatial error autocorrelation by including u_i and $\rho W_i u_i$. Additionally, this specification is estimated separately for time blocks to avoid misassignment of neighbors.

The spatial weight matrix, W_i, is built from a frequently used spatial structure, queen contiguity. It's akin to the movement of a queen in chess, which can move both along ranks and files and diagonally. Two regions are defined as neighbors if they share a common edge or a common vertex (corner). Figure 2.5 illustrates an example of queen contiguity. Area E in the figure has 8 neighbors according to queen contiguity. A spatial weight matrix W can define the elements in the weigh matrix to be 1/8 for those 8 neighbors and 0 for other areas. Thus Wz returns the average of z from 8 neighbors. It's similar to the concept of temporal proximity in the lag term from time series analysis, where a lag term represents a previous time period adjacent to the current one. So, it's referred to as a spatial lag term to account for geographical proximity.

A multi-step GMM/IV estimation procedure developed by Arraiz, Irani, et al. (2010) is employed for the estimation of the combined spatial lag and error model. It has the advantage of accommodating both spatial lag and spatial error structures, as well as heteroskedasticity. To consistently estimate the parameters, the instruments for endogenous explanatory variables are also used in this procedure, and spatial lag terms for all independent variables are used as additional instruments for $W_i \Delta y_i$, the spatial lag term of the dependent variable.

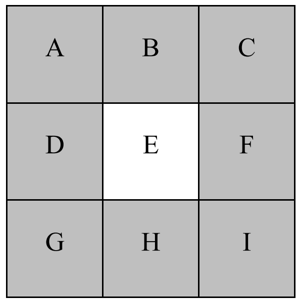


Figure 2.5: Illustration of Queen Contiguity

Queen contiguity defines neighbors as areas sharing a common edge or a common vertex (corner). In the graphed spatial structure, area A, B, C, D, F, G, H, and I are all neighbors to area E. The spatial matrix assigns a positive value on those neighbors and 0 on other areas.

2.5 Results

2.5.1 OLS and IV regression results

In this section, I present the estimation of the effects of carjacking upon ride-hailing service prices. Table 2.3 shows the results of estimating Equation (2.2) on both workday and non-workday observations. The first three columns show the OLS estimates on average fare per mile on workday. Column 3 is the full model, while column 2 removes all time block fixed effects, and column 1 removes all controls on 2019 levels in addition. Column 4 to 6 replicate the same specifications as column 1 to 3, but on the non-workday observations. The directions of estimates on Δx_{it} are consistent with the hypotheses, where Δx_{it} has a positive effect on the price since it scares drivers away from the area with increased carjacking cases. The sign of estimates on quantities, *delta_trips*, is also consistent with the hypotheses that the labor supply curve in the ride share market has a positive elasticity.

Table 2.3: OLS Regression Results on Ride-hailing Service Prices on both Workday and Non-workday

	Dependent variable					
	Workday delta_avg_fare_per_mile OLS			Non-workday delta_avg_fare_per_mile OLS		
	(1)	(2)	(3)	(4)	(5)	(6)
delta_carjack	0.027***	0.009	0.013**	0.042***	0.036**	0.018**
	(0.007)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)
delta_trips	0.012**	0.027***	0.027***	0.039***	0.037***	0.049***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
delta_miles	-0.492***	-0.522***	-0.559***	-0.520***	-0.579**	-0.569***
	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.012)
log_avg_fare_per_mile_19		-0.399***	-0.379***		-0.437***	-0.428***
		(0.038)	(0.037)		(0.036)	(0.034)
log_carjack_19		-0.002	0.014		0.032***	-0.009
		(0.010)	(0.009)		(0.014)	(0.013)
log_fare_19		0.221***	-0.236***		-0.360***	-0.135***
		(0.049)	(0.054)		(0.050)	(0.050)
log_miles_19		-0.369***	-0.143***		-0.088***	-0.213***
		(0.041)	(0.041)		(0.040)	(0.039)
log_trips_19		0.037***	0.034***		0.027***	0.031***
		(0.002)	(0.002)		(0.002)	(0.002)
hour_predawn			0.148***			0.037***
			(0.012)			(0.011)
hour_morning			0.126***			-0.144***
			(0.013)			(0.011)
hour_noon			0.017			-0.125***
			(0.012)			(0.011)
hour_afternoon			0.147***			-0.107***
			(0.012)			(0.011)
hour_evening			-0.016			-0.095***
			(0.011)			(0.010)
const	0.254***	0.532***	1.170***	0.221***	1.508***	1.237***
	(0.005)	(0.075)	(0.083)	(0.004)	(0.074)	(0.075)

p-value: 0.1 > * > 0.05 > ** > 0.001 > ***

Table 2.4: IV Regression Results on Ride-hailing Service Prices on both Workday and Non-workday

	Dependent variable					
	Workday delta_avg_fare_per_mile IV			Non-workday delta_avg_fare_per_mile IV		
	(1)	(2)	(3)	(4)	(5)	(6)
delta_carjack	0.102***	0.050***	0.047**	0.079**	0.035	0.014
	(0.024)	(0.016)	(0.021)	(0.035)	(0.042)	(0.035)
delta_trips	0.252***	0.067	0.104	0.342***	0.246**	0.154**
	(0.095)	(0.082)	(0.096)	(0.070)	(0.103)	(0.073)
delta_miles	-0.474***	-0.515***	-0.555***	-0.504***	-0.563***	-0.559***
	(0.026)	(0.021)	(0.021)	(0.031)	(0.027)	(0.024)
log_avg_fare_per_mile_19		-0.425***	-0.396***		-0.418***	-0.408***
		(0.068)	(0.051)		(0.059)	(0.053)
log_carjack_19		0.018	0.025		-0.011	-0.036
		(0.018)	(0.026)		(0.052)	(0.044)
log_fare_19		0.281***	-0.096		-0.361***	-0.082
		(0.075)	(0.166)		(0.078)	(0.085)
log_miles_19		-0.408***	-0.237*		-0.070	-0.232***
		(0.078)	(0.125)		(0.063)	(0.059)
log_trips_19		0.037***	0.036***		0.029***	0.032***
		(0.003)	(0.004)		(0.003)	(0.003)
hour_predawn			0.146***			0.039**
			(0.016)			(0.016)
hour_morning			0.107***			-0.154***
			(0.038)			(0.017)
hour_noon			-0.013			-0.155***
			(0.046)			(0.027)
hour_afternoon			0.113**			-0.129***
			(0.045)			(0.022)
hour_evening			-0.031*			-0.111***
			(0.018)			(0.017)
const	0.348***	0.481***	1.045***	0.271***	1.495***	1.156***
	(0.046)	(0.107)	(0.148)	(0.017)	(0.124)	(0.129)

p-value: 0.1 > * > 0.05 > ** > 0.001 > ***

The results also show the expected effect from delta trip miles. With an increase in trip miles, the average fare per mile decreases, potentially due to the dilution of the fare. The total number of ride-hailing trips in the area also has consistent positive effects on both price variables, potentially caused by excessive demand.

The in-day seasonality shows the most distinguishing pattern between workday and non-workday. On workdays, the price of ride-hailing services is topped during the predawn hours, followed by afternoon and morning rush hours. It is likely the result of extremely low labor supply during the predawn hours, and the ride-hailing platforms need to increase the price dramatically to encourage drivers to pick up passengers. However, on non-workdays without rush hours, the price of ride-hailing services is lower in the morning and the afternoon. It charges more to use ride-hailing services during midnight and predawn hours on non-workdays.

The results of the instrumental variable regression are delineated in Table 2.4. The IV regressions are conducted with identical specifications as in the previous table for workdays and non-workdays. The estimates on quantity supplied are still consistent with the hypotheses, however, the significance drops slightly. Considering the flexibility of becoming a ride-hailing driver, it is not surprising that the labor supply is highly elastic, which may help explain the insignificant estimates on the quantity supplied.

Once endogeneity is addressed, the influence of carjacking remains consistently robust on workdays, revealing a continuous positive effect on service pricing during workdays. This supports the study's hypothesis that ride-hailing drivers consciously steer clear of areas with a heightened risk of carjacking, leading to passengers in these areas having to pay a premium to avail services. The magnitude of this effect, however, is relatively minor. For instance, when there's a doubling in the number of carjacking cases (approximately a 0.3 increase in the delta log of carjacking cases), the per-mile service cost rises by approximately 1.4%.

2.5.2 Spatial regression results

As discussed in the previous section, the existence of spatial dependency can lead to erroneous inferences. To exam the existence of spatial dependency, I rerun the full specification IV regression for each hour block individually and apply Anselin-Kelejian test to regression residuals. Anselin-Kelejian test is a generalized version of Moran's I test, a widely used residual spatial autocorrelation test, on residuals from IV regressions. If the null hypothesis is rejected, there likely exists spatial autocorrelation in regression residuals. It suggests that the model may not be properly specified and there may be omitted spatial effects that have not been fully addressed. Table 2.5 shows test results. The null hypothesis is rejected for afternoon and evening on both workdays and non-workdays, but for midnight and morning, it's not rejected either on workdays or non-workdays. It's a bit ambiguous, but more likely to reject the null hypothesis for predawn and noon. In general, it provides evidence that there may be unaddressed spatial effects in the model. The spatial autoregressive model is necessary to address this issue.

Lagrange-Multiplier (LM) test is another usually employed in spatial analysis to test for a particular type of misspecification in a model. There are generally two types of LM tests used for diagnosing spatial models: LM-Lag test checks whether there is a missing spatially lagged dependent variable that should be included in the model; and LM-Error test assesses whether the residuals from the model are spatially autocorrelated, which would suggest that there is a missing spatial component in the error term. Although both tests are robust examining the existence of misspecification in the model, they are not robust on specifying the exact form of spatial dependency. A missing spatially lagged dependent variable can reject a LM-Error test, and vice versa. The robust forms of the tests make asymptotic adjustments to improve the performance on exact specifications (Anselin et al 1996). Table 2.6 shows the LM tests on hour block level IV regressions. All null hypotheses are rejected, suggesting the existence of unaddressed spatial dependency in the IV model, but the exact specification is undetermined. The combined spatial lag and error model is a proper specification in this case.

Table 2.7 shows the spatial regression results on workdays and Table 2.8 shows the results on non-workdays. Firstly, the two spatial autoregressive parameters, λ for spatial lag of dependent variable and ρ for residual spatial autocorrelation, both show significant estimates to some extent. ρ shows more consistent and robust estimates across all subsamples. This suggests that residual spatial autocorrelation addresses most of the spatial dependency in this model, despite that Lagrange-Multiplier test suggests no preference on either specification.

Table 2.5: Anselin-Kelejian Test on Hour Block Level IV Regressions

	Delta_avg_fare_per_mile			
	Workdays		Non-workdays	
Hour block	AK-test	p-value	AK-test	p-value
midnight	1.281	0.258	0.408	0.523
predawn	2.722	0.099	10.358	0.001
morning	1.630	0.202	0.838	0.360
noon	0.012	0.914	3.482	0.062
afternoon	16.934	0.000	35.478	0.000
evening	41.718	0.000	3.924	0.048

Table 2.6: Lagrange-Multiplier Test on Hour Block Level IV regressions

		LM-Lag	p-value	LM-Error	p-value	Robust LM-Lag	p-value	Robust LM-Error	p-value
workday	midnight	55.518	0.000	24.152	0.000	64.750	0.000	58.058	0.000
	predawn	77.119	0.000	68.869	0.000	90.978	0.000	90.009	0.000
	morning	54.008	0.000	68.126	0.000	84.764	0.000	62.122	0.000
	noon	9.546	0.002	80.163	0.000	12.782	0.000	11.342	0.001
	afternoon	116.942	0.000	76.622	0.000	195.915	0.000	137.328	0.000
	evening	98.496	0.000	84.561	0.000	146.901	0.000	118.736	0.000
Non-workday	midnight	45.456	0.000	46.309	0.000	71.196	0.000	49.904	0.000
	predawn	31.932	0.000	61.578	0.000	38.245	0.000	37.337	0.000
	morning	38.150	0.000	84.812	0.000	48.891	0.000	46.346	0.000
	noon	9.879	0.002	44.583	0.000	15.611	0.000	10.774	0.001
	afternoon	54.292	0.000	45.225	0.000	66.812	0.000	59.166	0.000
	evening	53.691	0.000	35.931	0.000	60.057	0.000	57.337	0.000

Table 2.7: Spatial Regression Results on Ride-hailing Service Prices on Workday

	Dependent variable					
	delta_avg_fare_per_mile Combined Spatial Lag and Error					
	midnight	predawn	morning	noon	afternoon	evening
	(1)	(2)	(3)	(4)	(5)	(6)
delta_carjack	0.086**	0.050	-0.083**	0.024	0.031**	0.011
	(0.042)	(0.034)	(0.031)	(0.032)	(0.016)	(0.018)
delta_trips	-0.117	-0.074**	0.042	-0.051	-0.046	0.159***
	(0.086)	(0.036)	(0.027)	(0.048)	(0.032)	(0.038)
delta_miles	-0.577***	-0.375***	-0.526***	-0.533***	-0.554***	-0.544***
	(0.049)	(0.054)	(0.079)	(0.059)	(0.045)	(0.035)
ρ	0.111	0.081	-0.48***	0.097	-0.069	0.155*
	(0.118)	(0.137)	(0.125)	(0.143)	(0.087)	(0.088)
λ	0.317**	0.302*	0.686***	0.362**	0.570***	0.468***
	(0.099)	(0.163)	(0.072)	(0.113)	(0.052)	(0.062)
log_avg_fare_per_mile_19	-0.786***	-0.022	-0.203**	-0.355**	-0.415***	-0.675***
	(0.150)	(0.118)	(0.089)	(0.167)	(0.113)	(0.100)
log_carjack_19	0.113**	0.010	-0.070**	0.045	0.012	-0.025
	(0.041)	(0.031)	(0.035)	(0.031)	(0.017)	(0.024)
log_fare_19	0.041	-0.715***	0.043	-0.487**	0.008	0.328
	(0.267)	(0.168)	(0.112)	(0.193)	(0.131)	(0.225)
log_miles_19	-0.537**	0.485**	-0.237**	0.090	-0.256**	-0.53***
	(0.193)	(0.148)	(0.108)	(0.188)	(0.120)	(0.155)
log_trips_19	0.025***	0.009*	0.022***	0.009*	0.022***	0.017**
	(0.007)	(0.005)	(0.004)	(0.005)	(0.004)	(0.006)
const	1.397***	0.995***	0.511**	1.42***	0.815***	0.844**
	(0.386)	(0.205)	(0.175)	(0.324)	(0.204)	(0.282)

p-value: 0.1 > * > 0.05 > ** > 0.001 > ***

Table 2.8: Spatial Regression Results on Ride-hailing Service Prices on Non-workday

	Dependent variable					
	delta_avg_fare_per_mile Combined Spatial Lag and Error					
	midnight	predawn	morning	noon	afternoon	evening
	(1)	(2)	(3)	(4)	(5)	(6)
delta_carjack	0.040	0.080*	-0.015	-0.038	0.030	0.000
	(0.054)	(0.048)	(0.043)	(0.033)	(0.027)	(0.024)
delta_trips	0.015	0.181**	0.075**	0.108**	-0.033	0.071**
	(0.073)	(0.081)	(0.028)	(0.036)	(0.035)	(0.035)
delta_miles	-0.550***	-0.657***	-0.389***	-0.483***	-0.477***	-0.592***
	(0.044)	(0.060)	(0.045)	(0.038)	(0.051)	(0.062)
ρ	-0.209*	0.262**	-0.063	-0.015	0.208	0.124
	(0.127)	(0.114)	(0.127)	(0.142)	(0.143)	(0.110)
λ	0.453***	0.196	0.458***	0.435**	0.315**	0.264***
	(0.064)	(0.120)	(0.120)	(0.136)	(0.101)	(0.079)
log_avg_fare_per_mile_19	-0.618***	-0.336**	-0.529***	-0.412***	-0.460***	-0.458**
	(0.128)	(0.151)	(0.093)	(0.082)	(0.112)	(0.149)
log_carjack_19	-0.002	0.015	0.019	-0.038	0.003	-0.002
	(0.054)	(0.055)	(0.057)	(0.048)	(0.029)	(0.028)
log_fare_19	-0.087	-0.271	-0.063	-0.012	-0.148	-0.031
	(0.165)	(0.244)	(0.154)	(0.149)	(0.174)	(0.197)
log_miles_19	-0.334**	-0.153	-0.252*	-0.222**	-0.180	-0.214
	(0.125)	(0.200)	(0.136)	(0.103)	(0.135)	(0.137)
log_trips_19	0.030***	0.020**	0.023***	0.021***	0.022***	0.016**
	(0.006)	(0.007)	(0.004)	(0.005)	(0.004)	(0.006)
const	1.364***	1.439***	1.046***	0.813***	1.120***	0.958**
	(0.290)	(0.310)	(0.189)	(0.230)	(0.293)	(0.313)

p-value: 0.1 > * > 0.05 > ** > 0.001 > ***

After controlling for both endogeneity and spatial dependency, the estimates on quantity supplied and carjack effect become less significant. The quantity supplied becomes mostly insignificant on workdays, and even shows a negative sign during predawn hours. It's still consistently positive on non-workdays. The insignificant labor supply elasticity on workdays may still be due to the flexible nature of the job, but needs further study to confirm. The estimates on the carjack effect also suffer from a declined significance level. The estimates only meet the hypothesis during midnight and afternoon hours on workdays and predawn hours on non-workdays, and the magnitude of the estimates are still close to the IV estimates. But it shows no significance in other hours besides a contradictory effect during morning hours on workdays. One possible explanation is that drivers are less repelled by carjack cases during daytime hours, but become more concerned when it gets dark. The results still provide evidence that drivers have avoidance behaviors towards carjack risks, although it may vary across the day and differ between workdays and non-workdays.

2.6 Conclusion

By analyzing the ride-hailing market in the city of Chicago, this research has provided empirical evidence supporting the hypothesis that ride-hailing drivers intentionally avoid areas with higher rates of specific types of crime, such as carjacking. This finding supplements the conclusions from other literature that ride-hailing service is more accessible in areas with higher crime rate. Not all types of crimes repel drivers, but carjacking does. Drivers' avoidance behavior influences the dynamic pricing of ride-hailing services, leading to a surge in prices in these high-risk areas. Notably, the observed effect was relatively minor, indicating that a significant increase in carjacking incidents would translate into only a moderate surge in per-mile cost for passengers. And this effect seems to be mostly during night hours. These findings not only contribute to our understanding of drivers' behavior and decision-making processes in the ride-hailing industry, but they also highlight the social implications of crime rates on market dynamics in the sharing economy. Future research could delve deeper into exploring other potential factors impacting ride-hailing driver behavior and pricing algorithms, and how these factors interact with one another in shaping the landscape of the sharing economy. This study lays the foundation for such further exploration, encouraging us to consider the complex interplay of safety, service pricing, and labor supply in ride-hailing services.

Chapter 3

MARKET IMPACTS OF A TOXIC ALGAE EVENT: THE CASE OF CALIFORNIA DUNGENESS CRAB[1]

3.1 Introduction

Climate change is predicted to impact seafood safety through increased incidence of chemical and biological contamination in seafood (e.g. toxic metals) and biological contamination events, e.g. harmful algal blooms or HABs (Moore et al. 2008; Berdalet et al. 2016; Marques et al. 2010). In addition to the direct economic impacts of contamination events, i.e. contaminated product cannot be brought to market, there are indirect impacts when consumers have imperfect information in the market for the contaminated product. Specifically, imperfect information regarding the health risks of contamination events can lead to avoidable welfare losses known as avoidance costs (Swartz and Strand 1981).[2]

Previous research has demonstrated that contamination events may negatively impact demand even after the product is no longer contaminated (Smith, Van Ravenswaay, and Thompson 1988; Wessells, Miller, and Brooks 1995) and a decline the in consumer demand can occur for similar products from uncontaminated areas (Swartz and Strand 1981; Wessells and Wilen 1994). Therefore, with avoidance cost, imperfect information related to the spatial and temporal scale of a contamination event can generate a behavioral response that is misaligned with health risks, unnecessarily reducing consumer surplus.

Furthermore, contamination events can enable other indirect economic impacts in the form of welfare transfers between various stages in the supply chain. For example, if a contamination event

[1] This Chapter is a joint work with Sunny Jardine from the University of Washington. It is based in full on the previously published article listed below: Mao, Junwei, and Sunny L. Jardine. "Market impacts of a toxic algae event: The case of california dungeness crab." Marine Resource Economics 35.1 (2020): 1-20. https://doi.org/10.1086/707643

[2] Note we use the definition of avoidance cost from Swartz and Strand (1981), which differs from the definition used by Shuldstad and Stoevener (1978) in that only avoidable welfare losses are considered.

altered the bargaining power between primary producers and processors, a welfare transfer may result. Price negotiations could be affected if there is either processor market power or the potential for opportunistic behavior on the part of processors, as described by Klein, Crawford, and Alchian (1978). While, to our knowledge, welfare transfers following contamination events have not previously been discussed in the literature, these equity implications of contamination events are important to consider.

Here we explore the empirical evidence for avoidance costs *or* welfare transfers generated by the 2015 harmful algal bloom event on the west coast of North America. During this event, harmful algae produced high levels of domoic acid (a neurotoxin) resulting in fishery closures for commercial Dungeness crab in California, beginning in November of 2015 and ending in March of 2016, lasting roughly 4.5 months. During this time there was both local and national media coverage of the fisheries closures, suggesting the possibility of a consumer response to the event.

As a first step, we empirically estimate the impacts of the 2015 HAB event on ex-vessel and retail prices using a differences-in-differences approach with Oregon and Washington as control groups. The approach is similar to other contamination impact assessments, which compare pre- and post-event outcomes, or conduct “before/after” analyses (Foster and Just 1989; Mazzocchi 2006; Jin, Thunberg, and Hoagland 2008). The difference is that including control groups allows us to control for changes in market conditions after the event that are unrelated to the event and may be confounded with the event. Because Oregon and Washington also experienced elevated levels of domoic acid, and relatively short fisheries closures, our estimates provide the impacts from a contamination event with a prolonged period of contamination and media coverage relative to the impacts from a short-lived contamination event.

Next, acknowledging the fact that Oregon and Washington were potentially affected by the fisheries closure in California through linked seafood markets, we estimate a non-parametric predictive price model to assess whether the control groups were affected by the California event. We then use the predictive price model to inform a refinement of our differences-in-differences estimate, by selecting control-group observations for which prices are not systematically different from our forecasts.

Our work makes three important contributions to the literature. First, we contribute to the revealed-preference literature on avoidance costs, which has found evidence of avoidable welfare losses in markets for dairy products (Smith, Van Ravenswaay, and Thompson 1988; Foster and Just 1989), meat (Mazzocchi 2006), and seafood (Wessells, Miller, and Brooks 1995). Second, we explore empirical evidence for multiple indirect economic impacts of a contamination event (avoidance costs and welfare transfers) by analyzing data at two stages in the supply chain. Therefore, we explore a broader set of economic impacts from contamination events than has been considered thus far. Third, we use non-parametric forecasting models to evaluate the stable unit treatment value assumption (SUTVA) required in difference-in-differences analyses (see Holland 1986) and we refine our estimate of the impacts using forecasting results. The approach can be useful in other settings where SUTVA is potentially violated.

Summarizing our results, we find ex-vessel prices fell by at least 9.6% but consumer prices were not impacted. We put forth three competing theories to explain this apparent disconnect and discuss the efficiency and distributional implications of each alternative.

3.2 Background and Data

The Dungeness crab fishery is the most valuable fishery on the West Coast of the U.S.A. In the 2012/2013-2014/2015 fishing seasons, average seasonal coast-wide landings of Dungeness crab totaled roughly 53.5 million pounds generating roughly $197.7 million in ex-vessel revenue (Table 3.1). California, Oregon, and Washington respectively landed an average of roughly 36%, 25%, and 40% of the coast-wide volume of Dungeness crab.

Dungeness crab fisheries in all three states are managed with size and sex restrictions and season length. Specifically, only male crabs larger than 159 mm in carapace width can be harvested, and the season is closed in the late summer and fall to protect vulnerable and unmarketable molting crabs (Didier, 2002). Each state implemented some form of limited entry in the mid 1990s although license eligibility requirements, fees, and rules regarding license transfers vary across the states (Didier, 2002). More recently, each state has adopted restrictions on the number of crab pots that a permitted vessel can use (Porzio 2015). Despite input controls, the Dungeness crab fisheries in all three states resemble derby fisheries where the “race to fish” creates a compressed season length. Although the season is open in the summer months, the majority of landings are taken before the end of January in each state (Figure 3.1).

In addition to the fisheries in each state being managed separately, management of the California Dungeness crab fishery is further divided into northern and central management districts at the Mendocino-Sonoma County border. The central California season runs from November 15th to June 30 and the northern California season runs from December 1st through July 15. Oregon and Washington fisheries typically begin on December 1st. Barring any delays in season start dates, California landings peak just before peak landings in Oregon and Washington (Figure 3.1).

However, delays in season start dates are not uncommon. The season opening date in California’s northern district and in Oregon and Washington are subject to delay if meat content is below established thresholds, and season open dates in all three states are subject to delay if domoic acid levels are higher than established thresholds for the toxin, i.e. >30 ppm in crab viscera or > 20 ppm in the crab meat. Season opening dates have also been delayed due to weather and extended price negotiations between fishermen and processors. Examples include a delay in the central California district in the 2011/2012 season due to price negotiations (Kauffman 2011) and delays in the 2012/2013 season in all three states due insufficient meat content after which weather conditions lengthened the delay for California vessels (Dillman 2013).

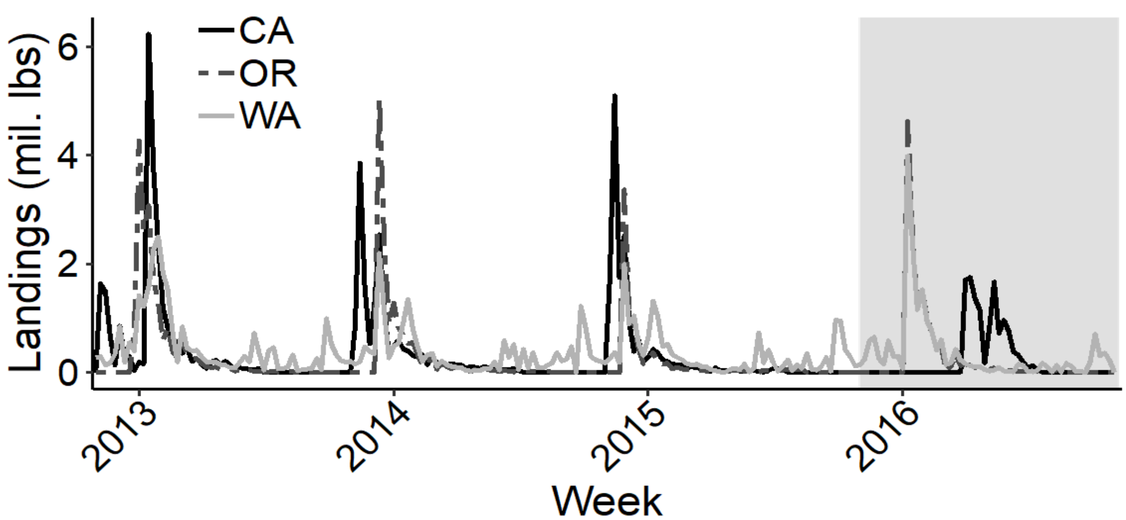


Figure 3.1: Crab landings by state

While not uncommon, delays in the Dungeness crab season start dates are typically short in duration, but the contamination event we examine here led to an unprecedented delay. Specifically, in the spring of 2015, a massive harmful algal bloom (*Pseudo-nitzschia*) led to elevated levels of the neurotoxin domoic acid, resulting in fisheries closures along the entire West Coast for many species including Dungeness crab (McCabe et al. 2016). On November 6th the California Department of Fish and Wildlife issued an emergency rule delaying the California Dungeness fishery, previously planned to open on November 15th. Following the announcement, the Center for Food Safety of Hong Kong announced a ban on Dungeness and Rock crabs harvested from California waters (Hong Kong Special Administrative Region 2015).

The closure in California lasted roughly 4.5 months in the central district and roughly 5 months in the northern district. The event was unmatched in its duration and impacts. In January of 2016, the State of California declared 15 California counties affected by the crab closures a disaster area and requested disaster relief funds for impacted businesses.

During this time, the event received extensive local and national media coverage in outlets such as the Wall Street Journal online edition (Elinson 2016) and National Public Radio (Sommer 2015). Oregon and Washington, both with scheduled opening dates December 1st, 2015, also experienced elevated domoic acid levels. However, the season delays in these two states were relatively short-lived, with crab in both states declared safe for consumption on January 4th, 2016. [3]

While public awareness of the event is difficult to observe, we examine data from NewsBank Inc. to understand the level and duration of media coverage in each state. Because the HAB event also affected recreational razor clam fisheries in Oregon and Washington, we focus on articles that strictly cover Dungeness crab. To that end, we conduct a search on unique newspaper articles including the terms "domoic acid" and "Dungeness", but not including the word "clam", published from 11/01/2015 to 7/31/2016. We find a total of 494, 45, and 37 newspaper articles fitting these

[3] Dungeness crab is also harvested in Alaska and in British Columbia, Canada. A reviewer notes that product from both of these fisheries may have been on the market during the time of the HAB event.

criteria in California, Oregon, and Washington respectively.[4] Figure 3.2 shows that weekly news publications fitting our search criteria peak during the week of Nov 1st,2015, the week when the delay in the California fishery was announced, and remain at high levels in California until the end of March 2016. News coverage in Oregon and Washington was noticeably lesser in magnitude and duration than in California.

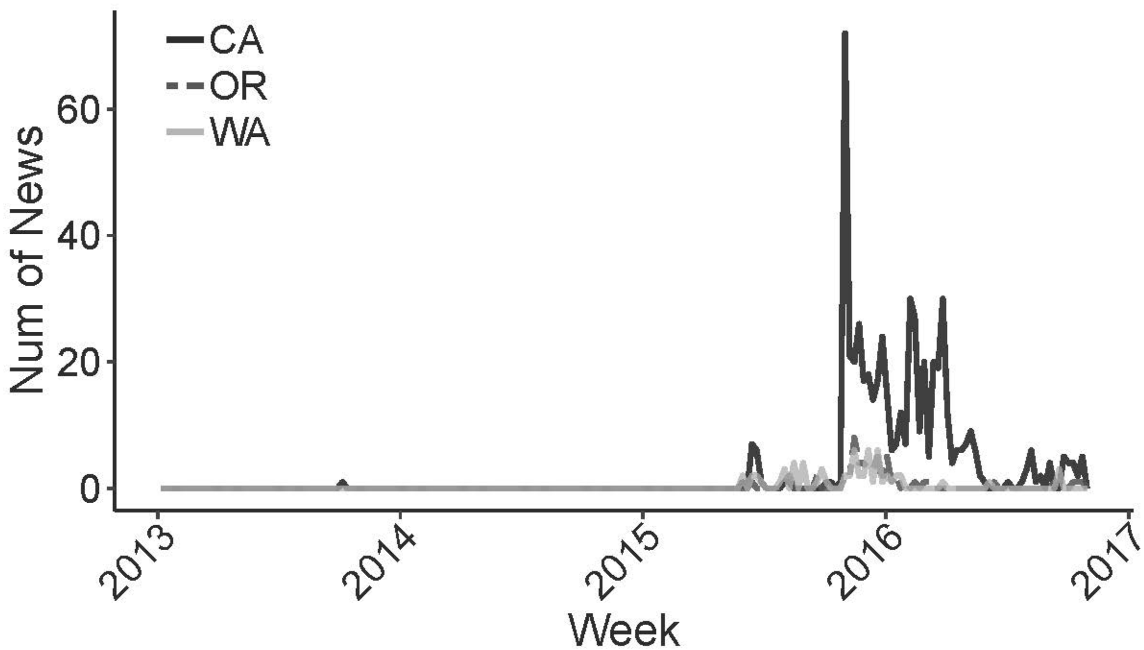


Figure 3.2: Number of unique news articles including the terms “domoic acid”, “Dungeness”, and not including the word “clam”

[4] Newspaper articles with identical title in one state is counted as one article. Duplicated articles are usually published by different local publishers to cover local population. Removal of duplicated articles helps adjust for population differences between states.

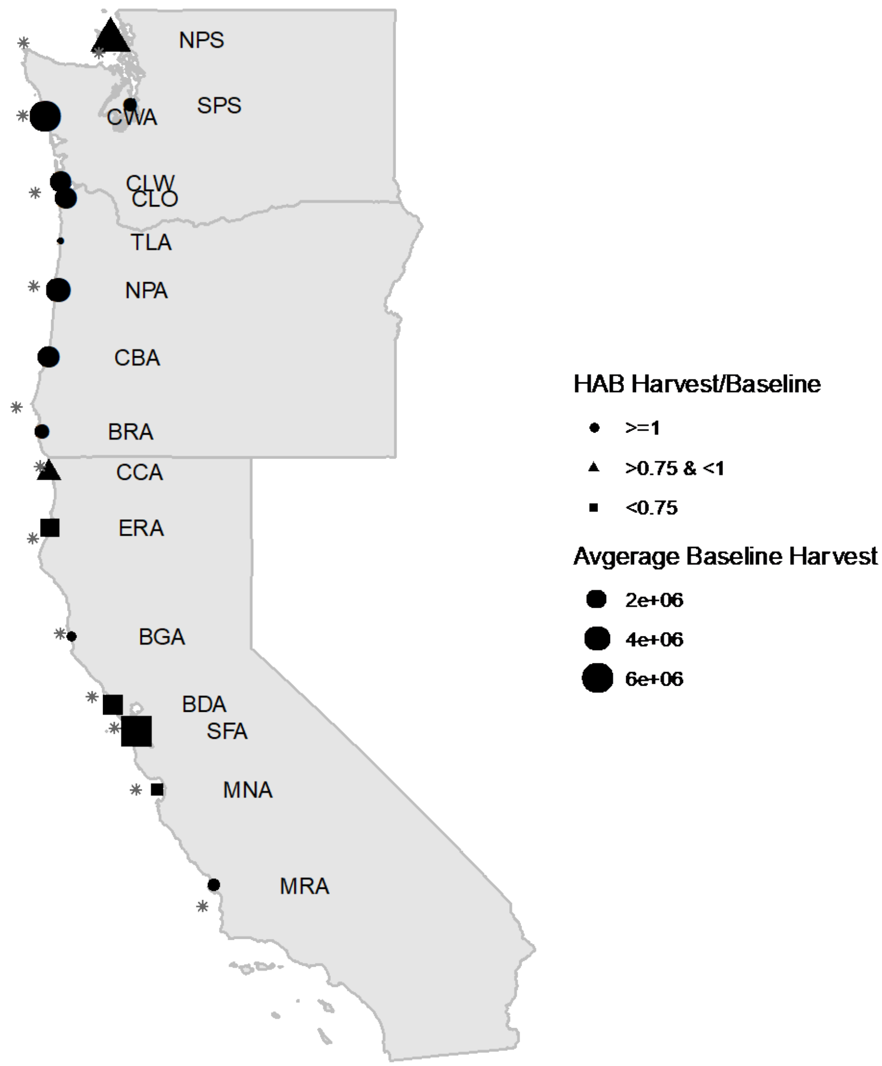
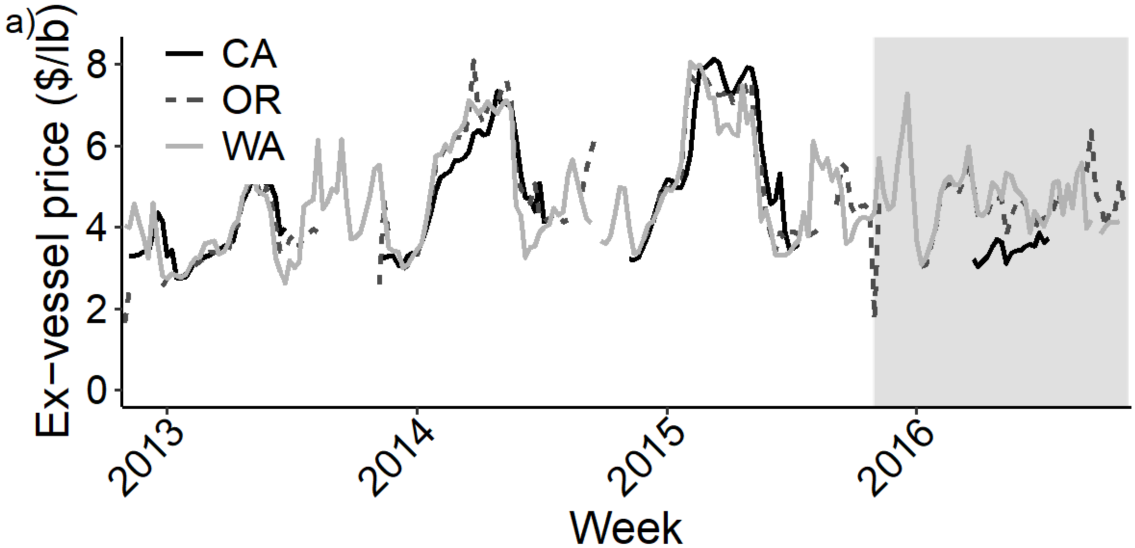


Figure 3.3**:** Map of port groups on the West Coast

The size of the markers indicates the port group's average landing quantity in pounds from 2013/2014 and 2014/2015 seasons. The shape of the circles indicates the ratio of the landing quantity in 2015/2016 season over the previous two years' average. The stars beside port groups indicate the location of climate stations.



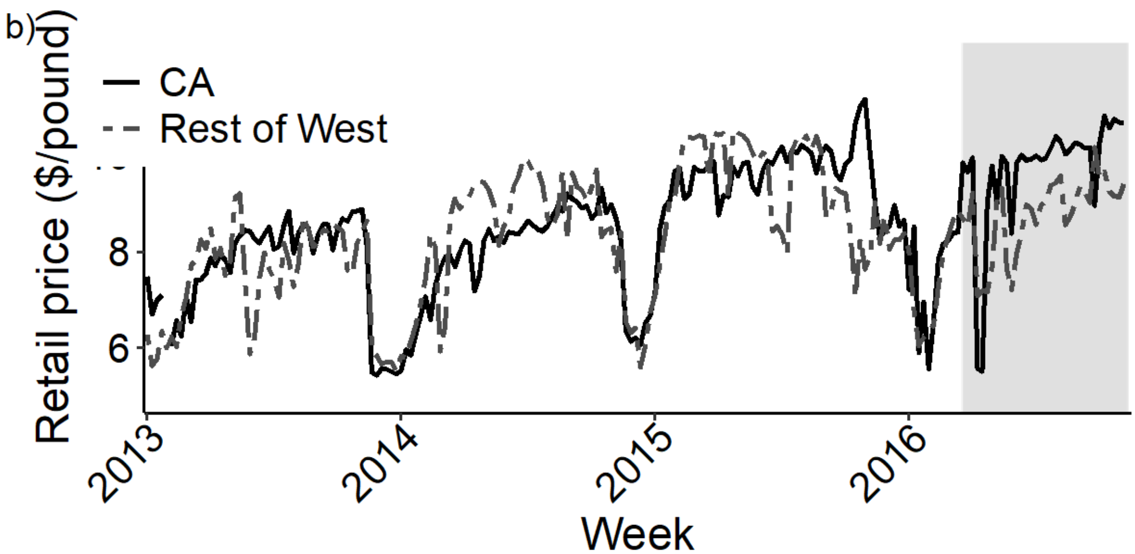


Figure 3.4: Average prices in the ex-vessel and retail markets

The top panel shows the average ex-vessel price in three states. The bottom panel shows the average retail price in California and the rest of the West. In panel (a) the gray shaded area is the period of the 2015/2016 season in California. In panel (b) the gray shaded area is the period of the 2015/2016 season when California fisheries were open.

To investigate the impacts of the event empirically, we utilize datasets at both the upstream (primary producer) level and downstream (retail) level. Upstream market data are taken from the Pacific Fisheries Information Network (PacFIN). Weekly Dungeness crab landing quantity in pounds and ex-vessel revenue are reported at port-group level and used to calculate weekly port-group average ex-vessel prices. Our data covers the period from November 15, 2012 to November 13, 2016, or 4 fishing seasons, and 18 port-groups (Figure 3.3).[5] Due to confidentiality restrictions, our data only include weekly observations where there are more than 3 active vessels and 3 active dealers. All prices are adjusted for inflation and reported in 2017 USD.[6] We identify five outliers in the data, with prices below $1 per pound, which we remove.

Figure 3.4, panel a, shows average ex-vessel prices by state. Noticeable from the figure is that there is a strong seasonal component to ex-vessel prices, driven by both seasonality in crab volume the market share of the various product types. A market assessment conducted in the early 2000s revealed that at the beginning of the season large volumes of crab are delivered in a narrow window of time, therefore, the majority of landed crabs are frozen to facilitate quick processing (Hackett et al. 2003). Towards the end of the season, if target inventory levels are met, processors shift to producing fresh and live products (Hackett et al. 2003). Additionally, over the last decade there has been rapid growth in Chinese demand for live Dungeness crabs, restricting the market share of traditional products (Talley 2011) and potentially dominating the market late in the season (Tobias 2011). Thus, a visual examination of our price data suggests the importance of controlling for the impact of quantity and seasonality on ex-vessel prices.

In addition to ex-vessel prices, we also analyze the downstream consumer markets using IRI scanner data from January 1, 2013 to November 13, 2016.[7] Figure 3.4, panel b, shows retail prices in California and the rest of the Western U.S. [8] The data include weekly sales value and pounds of

---

[5] Port group codes and definitions are provided on the PacFin website https://pacfin.psmfc.org/pacfin_pub/codes.php. We exclude 3 port groups: WA5, OR1, and CA2 that include transactions from "other or unknown ports" in each state.
[6] We adjust for inflation using the Bureau of Labor Statistics Series ID CUUR0400SA0.
[7] Due to data limitations we only include part of the 2012/2013 fishing season in the retail price analysis.
[8] The region includes Arizona, Colorado, Idaho, Montana, New Mexico, Nevada, Oregon, Washington, Utah, and Wyoming. Some grocery chains included in the IRI universe may not exist in the Western U.S.

round weight Dungeness crab, or crab sold by the pound at the seafood counter in the grocery store, from over 100 grocery store chains, are aggregated to regional level and used to calculate average prices. The data represent only a small fraction of total landings (less than 10%) but allow for an opportunity to further investigate the evidence of a consumer response to the HAB event. As with the ex-vessel price data, there is a strong seasonal component to retail Dungeness crab prices likely also driven by seasonality in Dungeness supply suggesting the importance of controlling for the quantity supplied and seasonality.

Table 3.1 reports total sales volumes and values at both the ex-vessel and retail levels, by season and state. These data show that Washington is the largest producer of Dungeness crab followed by California and then Oregon. Table 3.1 also shows that, in contrast to California, landings were stable in Oregon and Washington during the HAB event (the 2015/2016 season) meaning that the season delays in Oregon and Washington did not lead to detectable reductions in total landings during the season. Data on the retail market show sales volumes and values vary widely over the time period of our analysis and that there is no noticeable reduction in the volume of crab sold at the retail level during the 2015/2016 season.

We also summarize our data in relation to the 2015 harmful algal bloom event. Table 3.2 reports summary statistics of the weekly ex-vessel dataset before, during, and after the event. The data show that mean ex-vessel prices after the event are much lower in California, but remain relatively stable in Oregon and Washington. Table 3.3 shows summary statistics of the retail price data before and after the event. Interestingly, in the post-treatment period, we observe an increase in mean prices in California (of $1.42), which is relatively large compared to the increase observed in the Rest of the West (an increase of $0.51).

Because weather can drive the quantity of crab supplied to the market, and thus the price of crab, we draw on data from the National Oceanic and Atmospheric Administration's (NOAA's) National Data Buoy Center. The NOAA climate data include significant wave height (m) and wind speed (m/s) are available on hourly bases depending on the climate station. To match the temporal frequency in weather variables to that of the ex-vessel price data, we first form a daily average within the calendar day, and then a weekly average is calculated from the daily averages. Climate data is then matched to ex-vessel data manually using the geographic locations of the climate station and port group as shown in Figure 3.3. The small stars beside each port group represent the relevant climate stations.[9]

Table 3.1: Total Seasonal Landings, Revenue, Retail Volumes, and Sales

	Ex-vessel market					
	California		Oregon		Washington	
State season total	Landings (mil. lbs.)	Revenue (mil. USD)	Landings (mil. lbs.)	Revenue (mil. USD)	Landings (mil. lbs.)	Revenue (mil. USD)
2012/2013	24.051	74.810	18.075	52.867	26.067	88.086
2013/2014	17.210	64.280	14.390	53.754	18.897	80.001
2014/2015	16.243	62.839	8.203	35.479	17.424	81.091
2015/2016	12.245	40.561	14.205	53.213	19.504	82.037

	Retail market			
	California		Rest of the West	
Season	Volume (mil. lbs.)	Sales (mil. USD)	Volume (mil. lbs.)	Sales (mil. USD)
2012/2013	2.560	15.135	1.674	10.256
2013/2014	1.813	11.527	1.302	8.491
2014/2015	0.473	4.043	0.493	4.130
2015/2016	1.147	7.818	1.172	8.309

[9] Note that some port groups share a climate station.

Table 3.2: Summary Statistics of Ex-vessel Market Data

		California	Oregon	Washington
Landing quantity (thousand pounds)	Before-event	95.048	77.908	115.108
		(316.595)	(202.263)	(226.763)
	During-event		232.077	233.386
			(337.025)	(401.576)
	After-event	137.587	10.508	39.057
		(222.995)	(11.390)	(81.412)
Ex-vessel revenue (thousand 2017 USD)	Before-event	333.767	272.222	459.738
		(957.455)	(617.788)	(785.820)
	During-event		854.206	967.702
			(1030.673)	(1300.352)
	After-event	455.744	46.757	173.745
		(710.261)	(49.852)	(335.330)
Average price (2017 USD)	Before-event	4.854	4.814	4.763
		(1.479)	(1.530)	(1.371)
	During-event		4.447	4.876
			(0.893)	(1.055)
	After-event	3.493	4.554	4.632
		(0.281)	(0.587)	(0.668)

Note: Data are at the week and port-group level from November 15, 2012 to November 13, 2016. Standard deviations appear in parentheses.

Table 3.3: Summary Statistics of Retail Market Data

		California	Rest of West
Average price (2017 USD)	Pre-treatment	8.31	8.27
		(1.30)	(1.37)
	Post-treatment	9.73	8.78
		(1.22)	(0.82)
Sales quantity (pound)	Pre-treatment	32757.48	25504.85
		(66428.25)	(45091.09)
	Post-treatment	14314.15	10311.00
		(36979.00)	(6175.44)

Note: Data are at the weekly level from January 1, 2013 to November 13, 2016. Standard deviations appear in parentheses.

3.3 Identification Strategy

We quantitatively estimate the indirect economic impacts of the 2015 HAB event with a difference-in-differences (DiD) framework. The method is a counterfactual analysis estimating how Dungeness crab markets in California would have behaved in 2015/2016 season in the absence of the HAB event. The DiD framework allows us to control for unobserved factors in the post-contamination period that affect California and our control groups equally, e.g. changes in the preferences for Dungeness crab, that would otherwise be confounded with the impact of the contamination event.

We use Dungeness crab markets in Washington and Oregon as control groups. Our hypotheses are as follows: the prolonged contamination event in California, during the 2015/2016 fishing season, led to an avoidance cost that will be detectable in both ex-vessel and consumer markets for Dungeness crab. If the SUTVA assumption holds, the results from our DiD model will be relative to the impacts of a short-lived contamination event (as experienced by the control groups).

The literature has shown that, in contrast to short-lived contamination events, which tend to have small and temporary impacts on consumer behavior, prolonged events with longer periods of media coverage can result in large and lasting behavioral responses, although behavioral responses are not guaranteed (Kalaitzandonakes, Marks, and Vickner 2004; Wessells, Miller, and Brooks 1995; Lloyd et al. 2001; Dahlgran and Fairchild 2002). Therefore, we expect to find evidence of a relative impact in the California markets (both upstream and downstream) and our DiD impact estimates will be conservative estimates if Oregon and Washington were negatively affected by the 2015 event.[10]

To ensure the DiD impact estimates are unbiased, certain conditions are required. The essential assumption is the common trends assumption, i.e. outcomes in the treatment and control groups are driven by the same market forces and follow the same trends. Figure 3.4 provides evidence in support of this assumption. We see that in the ex-vessel market, prices in three states track each

---

10 In a before/after analysis on ex-vessel prices in Oregon and Washington we do not see a statistically significant difference in prices after the event at standard levels of significance. Estimates are available on request from the authors.

other very closely before the event. A similar pattern is found in the consumer market data, where price in California and the rest of West US follow similar trends before the event. Given that these three states are geographically close to each other the product is nearly homogeneous, the result is not surprising and the assumption of parallel trends seems reasonable.

Another important assumption in the DiD framework is that the treatment, i.e. a prolonged contamination event in California, must not impact the control groups (also known as SUTVA). It is possible that market linkages between the three states could lead to a violation of the assumption. For example, the closure of the California crab fishery led to a shortage of Dungeness crab during the holiday season, which may have driven up demand for crabs from Oregon and Washington. Additionally, if consumers were avoiding California crabs due to safety concerns, the demand for Oregon and Washington crabs could also increase. Another possibility is that there are spillover effects where extended media attention during the California event led to consumer avoidance of Dungeness crab from all three states. Similar spillover effects have been documented in other markets (e.g. Swartz and Strand 1981; Wessells, Miller, and Brooks 1995). With any of these scenarios the SUTVA would be violated and the DiD framework would produce a biased estimate of the impact of the HAB on outcomes in California.

Therefore, in order to assess the validity of the SUTVA assumption, we introduce a machine-learning (LASSO) model to predict the ex-vessel price for each port group based on own-port landings and landing quantities from all other port groups. The model is estimated on the pre-treatment data (before the HAB event) and used to predict ex-vessel prices during and after the California HAB-related closures. We examine the prediction errors to look for systematic differences in the predicted and actual prices as evidence that SUTVA is not satisfied.

3.4 Models

In this section, we discuss the econometric models used to estimate the impact of the 2015 HAB event in both the ex-vessel and consumer markets. The DiD estimator indicates the impact of the treatment event on the outcome variable after controlling for other factors.

3.4.1 Ex-vessel Market

The DiD model for the ex-vessel market is as follows:

$$y_{i,t} = \beta_0 + \beta_1 T_i + \beta_2 P_t + \beta_3 TP_{i,t} + \beta_4 x_{i,t} + \rho_i + Z_t + \epsilon_{it}, \tag{3.1}$$

where i indexes a port group and t indexes a week; $y_{i,t}$ is the natural log of the per-pound ex-vessel price for Dungeness crab in port group i and week t; T_i is a dummy variable that equals one for observations in port groups that are in California; P_t is a dummy variable that equals one for observations after the first week of November 2015; $TP_{i,t}$ is the interaction of T_i and P_t, and is the variable of our interest. If the coefficient on $TP_{i,t}$ (β_3) is negative, it suggests that the treatment state (California) suffered from excessive price drop due to the event.

Our model controls for $x_{i,t}$, the natural log of Dungeness crab harvest (in pounds) in port group i and week t. Quantity is an important control variable because the product type that crab is processed into, and therefore the value of marginal product for Dungeness crab, are a function of the volume of crab landings (Hackett et al. 2003). Additionally, U.S. West Coast Dungeness crab fisheries are the largest producers of Dungeness crab and processors likely face a downwards-sloping demand curve, making it important to control for the impact of quantity on prices.

We also include a rich set of fixed effects: ρ_i, port group fixed effects which capture any time-invariant differences in outcomes across port groups; and Z_t, a vector of time fixed effects, including week fixed effects, year fixed effects, and year-week fixed effects, capturing time-varying drivers of outcomes that are constant across port-group. The port group fixed effects control for all factors that drive differences in prices between port groups that do not change over time, e.g. they can control for the possibility that processors in different areas have different production functions, or market access, and therefore output prices. The week fixed effects capture the price effects of seasonality in demand, seasonality in crab quality due to changing meat during the season, and holidays that occur in the same week of each year, e.g. Christmas. The year-week fixed effects can capture the important impacts of holidays that fall in different weeks of the year depending on the year, e.g. the Thanksgiving holiday. Finally, the model includes an error term, ϵ_{it}, assumed to be normal and identically and independently distributed.

Harvest quantity is an essential control in the model but potentially endogenous to the ex-vessel price if the supply of Dungeness crabs is not exogenous. To address the potential endogeneity of crab harvest we use wave height and wind speed as instrumental variables for quantity. We conduct Wu-Hausman and Weak IV tests to assess the need for and strength of our instruments respectively.

To examine the validity of the SUTVA assumption, or that control group may also have been affected by the treatment, we introduce an alternative LASSO model to predict ex-vessel prices. The solution to our predictive model is obtained by solving the following minimization problem:

$$\min_{\hat{\beta}_i, \hat{\gamma}_j} L(\hat{\beta}_i, \hat{\gamma}_j, \lambda) = \sum_{t=0}^{T} \left(y_{i,t} - \hat{\beta}_i x_{i,t} - \sum_{j=1}^{m} \hat{\gamma}_j z_{j,t} \right)^2 + \lambda \sum_{j=1}^{m} |\hat{\gamma}_j|, \tag{3.2}$$

where i indexes a port group and t indexes a week. $y_{i,t}$ is the natural log of per-pound ex-vessel price for Dungeness crab in port group i on week t; $x_{i,t}$ is the natural log of Dungeness crab harvest (in pounds) in port group i on week t; $z_{j,t}$ is a set of control variables, which includes the natural log of Dungeness crab harvest in every other port group, in each state, and West Coast as a whole. It also contains the natural log of cumulative Dungeness crab harvest in previous 4, 8, and 12 weeks, for each port group, in each state, and the West Coast as a whole. The LASSO penalty term is λ. Without the penalty, the model is the same as an OLS. The LASSO penalty helps to select the most relevant features from the set of control variables and reduces coefficients of irrelevant control variables to zero.

The forecasting model captures the fact that the price on the ex-vessel market may depend not only on supply in one port group, but also on supply in all port groups. For example, large harvest levels in one port group may lower prices in nearby ports. Therefore, our control variables include a comprehensive set of possible substitution effects in the market. A problem with including such a large number of control variables in OLS estimation is the resulting high variance and the poor accuracy in prediction. Therefore, the LASSO model is used to reduce the dimensionality by feature selection, improving predictive power relative to OLS.

We estimate the model with pre-treatment data and get $\hat{\beta}_i, \hat{\gamma}_j, \lambda$ for each port group, and predict the price for post-treatment periods using $\hat{\beta}_i, \hat{\gamma}_j, \lambda$ and post-treatment landing quantities. Prediction errors should contain demand shocks driven by the extended closure in the California Dungeness crab fishery, because they are the difference between the expected and observed price. Thus, we can use the prediction errors to determine whether the SUTVA assumption holds in our setting.

3.4.2 Retail Market

For the downstream or consumer market, we set up a similar econometric model of inverse demand for Dungeness crab:

$$y_{i,t} = \beta_0 + \beta_1 T_i + \beta_2 P_t + \beta_3 TP_{i,t} + \beta_4 x_{i,t} + Z_t + \epsilon_{it}, \tag{3.3}$$

where i indexes whether the observation is from California or in the rest of West US, and t indexes a week. The dependent variable $y_{i,t}$ is the natural log of per-pound retail price for Dungeness crab in group i on week t; T_i is a dummy variable that equals one for observations in California, and zero for those in the rest of West US; P_t is a dummy variable that equals one for observations after the last week of March 2016 when California Dungeness crab opened; $TP_{i,t}$ is the interaction term of interest; $x_{i,t}$ is the natural log of Dungeness crab sales (in pounds) in group i on week t; and Z_t are time fixed-effects controlling for time-varying factors affecting consumer demand such as prices for substitute goods.[11]

Similar to the ex-vessel market model, retail quantity is an important but endogenous control variable here. We introduce Hausman (1996) style instruments to control for the endogeneity. Specifically, retail prices in North Central U.S. and Northeast U.S. are used as IV for retail quantity. As Dungeness crab is mainly supplied by the west coast of the US, these instruments will capture the common supply shocks affecting all markets.

[11] Because we use aggregate data for California and the rest of the U.S. West Coast, we do not include group fixed effects in this model.

3.5 Results

Table 3.4 shows the results of DiD model on the ex-vessel market. The left panel shows OLS estimations, and the right panel shows instrumental variable estimations. We estimate the model with 3 specifications using all observations in the control group: in Specification 1, we only control for quantity; in Specification 2, we include the quantity control and port-group specific fixed effects; and in Specification 3, we have full controls including quantity, port group fixed effects, and period fixed effects.

DID estimates from OLS lie between -0.22 and -0.23 (results columns 1-3 of Table 3.4), suggesting that the prolonged HAB event resulted in a 22% - 23% reduction in California ex-vessel prices. Instrumental variable estimations are less in magnitude, ranging from -0.12 to -0.21 (results columns 5-7 of Table 3.4). The specification with full controls, or Specification 3, shows the lowest impact estimate, suggesting the price was 12.3% lower in California due to the HAB event.

Results from Wu-Hausman tests, which examine the endogeneity in the model, show overall support for an instrumental variables approach. Specifically, we reject the null hypothesis, that IV and OLS are both consistent, for specifications 2 and 3. First stage F-tests (or weak IV tests) for all three specifications show significance at 1% level, suggesting the climate data we use are relevant instruments.

Figure 3.5 shows the residuals from the LASSO model. In the pre-treatment period, prediction errors (observed prices minus predicted prices) are small and centered around zero for all 3 states, which is not surprising. During the event (the California closure) there appear to be systematic patterns in the residuals for both Oregon and Washington, although Oregon appears to be more greatly affected by the California event. Specifically, during the closure, our forecasting model tends to underpredict prices for both states with negative model residuals in only 7.7% and 25% weeks during the closure for Oregon and Washington respectively. After the California fishery was opened, the model overestimates prices for California (as expected) and Oregon, while model residuals are close to zero for Washington (median residual of -0.207).

Table 3.4: DiD results on ex-vessel data (2017 USD)

	Dependent variable							
	log(price)							
	OLS				Instrumental variable			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(pound)	−0.055***	−0.065***	−0.018***	−0.021***	−0.059***	−0.041***	−0.057***	−0.076***
	(0.003)	(0.004)	(0.002)	(0.003)	(0.013)	(0.011)	(0.015)	(0.020)
treat	0.013	0.034	−0.022	−0.071***	−0.013	0.019	0.029	−0.064***
	(0.013)	(0.034)	(0.015)	(0.014)	(0.014)	(0.037)	(0.019)	(0.016)
post	−0.010	−0.008	0.124	-0.115	0.013	0.009	0.016	-0.137
	(0.016)	(0.016)	(0.086)	(0.102)	(0.018)	(0.017)	(0.098)	(0.110)
cross	−0.228***	−0.215***	−0.229***	−0.234***	−0.204***	−0.205***	−0.123***	−0.096*
	(0.034)	(0.033)	(0.018)	(0.022)	(0.036)	(0.035)	(0.039)	(0.050)
group fixed eff		X	X	X		X	X	X
period fixed eff			X	X			X	X
Observations	2,095	2,095	2,095	1,335	1,814	1,814	1,814	1,165
R^2	0.143	0.188	0.858	0.867	0.147	0.180	0.845	0.837
Adjusted R^2	0.142	0.180	0.841	0.843	0.145	0.172	0.822	0.803
Weak IV					79.335***	123.863***	25.519***	17.863***
Wu-Hausman					0.087	5.608**	7.404***	8.893***

Significance key: *p<0.1; **p<0.05; ***p<0.01

Table 3.5: DiD results on retail data (2017 USD)

	Dependent variable			
	log(price)			
	OLS		Instrumental variable	
	(1)	(2)	(1)	(2)
log(pound)	−0.108***	−0.112***	−0.117***	−0.124***
	(0.003)	(0.005)	(0.017)	(0.036)
treat	−0.008	−0.008	−0.009	−0.010
	(0.009)	(0.007)	(0.010)	(0.008)
post	0.062***	0.011	0.062***	0.013
	(0.016)	(0.019)	(0.017)	(0.020)
cross	0.030	0.028	0.025	0.022
	(0.023)	(0.017)	(0.025)	(0.026)
period fixed eff		X		X
Observations	402	402	402	402
R2	0.755	0.886	0.751	0.884
Adjusted R2	0.753	0.867	0.748	0.864
Week IV			7.624***	3.543**
Wu-Hausman			0.297	0.115

Significance key: *p<0.1; **p<0.05; ***p<0.01

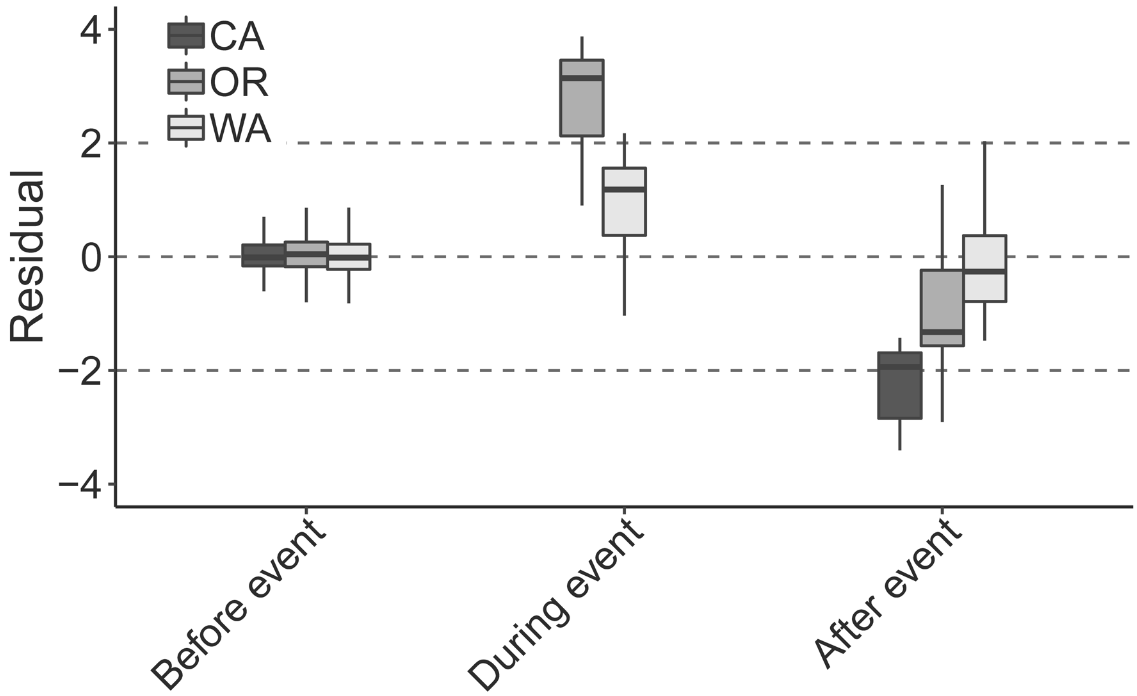


Figure 3.5: Residual from the LASSO model

Residuals are defined as actual prices minus predicted prices. The length of the whiskers of the boxplots are 1.5 times the interquartile range.

Therefore, results of the forecasting exercise suggest that the SUTVA assumption is possibly violated for both states while the California fishery was closed and for Oregon even after the California fishery reopened. Based on this result, we estimate a new specification of the model (Specification 4), which includes the same control variables as in Specification 3, but removes Oregon as a control group as well as observations from Washington when the California fishery was closed. Using this refinement of our control group we find that the OLS estimate is nearly unchanged (Table 3.4, results column 4). However, the IV estimate drops to -0.096 and the precision of the estimate decreases with a p-value of 0.056 (Table 3.4, results column 8). A test for endogeneity again suggests that the IV model is preferred, making the IV estimate of Specification 4 our preferred estimate.

Table 3.5 shows the results of the retail market. We present results of two specifications for both OLS estimation and IV estimation: in Specification 1 we only control for retail quantity; and in Specification 2 we include controls for retail quantity and period fixed effects. The OLS estimates show that the prolonged HAB event had an insignificant but positive impact on retail prices in California. The results from the IV estimates are qualitatively the same but with reduced magnitude. Wu-Hausman tests suggest that the OLS estimates are consistent and thus preferred to the IV estimates. In summary, we find no evidence of avoidance cost in the California consumer market.

3.5.1 Welfare loss

We find conflicting evidence of a consumer response and avoidance cost generated by the 2015 HAB event in the California Dungeness crab fishery. However, in what follows, we calculate the implied avoidance cost based on our results from the ex-vessel market.

Unlike traditional welfare loss estimation under linear demand system, our estimation under constant elasticity demand system cannot be done directly as the willingness to pay goes to infinity when the market quantity approaches zero. Figure 3.6 shows an example of the system. In order to calculate the welfare loss, we set a ceiling of willingness to pay at the highest observed price in our data, \$9.10. By setting a ceiling we may underestimate the welfare loss, but with this simplification we are able to avoid extrapolating the demand curve out of the range of the data as well as avoid the assumption that the willingness to pay for crab gets infinitely high as the supply approaches zero.

Next, we use the observed landing quantity as the market supply curve, i.e. we assume an inelastic fisheries supply. The assumption of inelastic fisheries supply is common in the literature due to the fact that a lack of property rights leads to overcapacity in the fishery, such that supply is constrained by regulations rather than prices (see Nielsen, Smit, and Guillen 2012 for a discussion) and has been employed in the calculation of welfare changes in fisheries settings, e.g. Jensen (2007). Figure 3.6 illustrates our assumptions on the ex-vessel market for Dungeness crab in California. If instead of a vertical supply curve, the true supply curve is upward sloping, our calculations will underestimate the true welfare loss or transfer. [12]

Our estimation of welfare loss is, therefore, determined by the real demand, estimated counterfactual demand, and the boundaries we set on price and quantity (the shaded area in Figure 3.6). Based on our most conservative impact estimate with (IV Specification 4 or results column 8 in Table 3.4), our preferred estimate of welfare loss/transfer is $4.4 million. The loss represents 10.9% of total revenue from the 2015/2016 California Dungeness crab fishing season and 6.5% of total revenue in the previous three seasons (seasons 2012/2013 to 2014/2015). Our highest impact estimates (IV Specification 2 in Table 3.3) implies a welfare loss/transfer of $9.64 million.

3.6 Discussion

Our results suggest that the prolonged HAB event, in 2015, had an impact on the California ex-vessel market. However, we observe no impacts on prices in the California consumer market. We put forth three competing theories to explain these outcomes and discuss the efficiency and distributional implications of each alternative.

One possibility is that consumers did respond to the contamination event, but we are unable to observe this response in the scanner data. The scanner data represents aggregate weekly sales of fresh Dungeness crab for California and the rest of the West Coast at grocery stores. We do not

[12] In fact, the results of our Wu-Hausman tests for endogeneity suggest that there is some elasticity to supply making our estimate of welfare loss conservative.

have information on point of origin although it is possible that information on point of origin was available to consumers at the grocery store. Therefore, if California consumers were avoiding crab from California and simply purchasing Oregon or Washington crab instead, aggregate consumption could be unchanged. Additionally, the total annual volume of fresh Dungeness crab sales in the IRI scanner data are 7.11%, 6.04%, and 5.53% of the total West Coast landings in the 2013/2014, 2014/2015, and 2015/2016 fishing seasons respectively, suggesting that our consumer data may not be representative of events in the larger consumer market for Dungeness crab.[13] We do not observe prices paid for the remainder of crab landings, e.g. crabs that were sold to restaurants, directly to consumers, or sold to export markets, although these sales comprise the majority of crab markets and these other market segments could have experienced a consumer response. For all of these reasons, it is possible that a consumer response did occur but we are unable to observe it. In this scenario, we estimate an avoidance cost of roughly $4.4 million in upstream markets. Whether the avoidance cost in upstream markets represents the total avoidance costs over the entire supply chain depends on whether changes in the ex-vessel market affect prices of other products and factors (Just, Hueth, and Schmitz 2004).

A second possibility is that there was no significant consumer response, but processors anticipated a consumer response and reduced output prices and that these expectations determined the ex-ante price negotiations. In this case, ex-vessel price was predetermined before processors could learn from consumer market. Their expectation of negative consumer response, which was not real, reduced the ex-vessel price on the contracts. The implication is that there would be no welfare loss in the Dungeness crab markets, but there would be impacts to equity within the fishing supply chain, in the form of a welfare transfer from fishermen to processors to the sum of roughly $4.4 million.

A third possibility is that there was no consumer response, but bargaining power between fishermen and processors was impacted by the 2015 event. Typically starting ex-vessel prices are

---

[13] It should be noted that our volume calculations are based on weights of whole unprocessed crab landings and processed Dungeness crab sold as a round weight product at the seafood counter. The product type is unobserved. Hackett et al. (2003) estimate that crabs yield between 25% and 87.5% of marketable product depending on product type. However, even at the lower bound of 25% product yield, the volumes suggest that our retail data do not represent a majority of landings.

set in pre-season negotiations where harvester strikes are not uncommon. Price negotiations continue throughout the season and mid-season strikes can occur (Hagenbuch 2017; Porzio 2015). Even in markets with many buyers (or processors) ex-vessel prices may not be competitive (equal to the value of marginal product) if processors engage in opportunistic behavior (as described in Klein, Crawford, and Alchian 1978). It is possible that fishermen were more sensitive to the delay and direct economic losses associated with foregone harvest, unwilling to further delay the season with a strike, reducing their willingness to accept for crab. As in the scenario of an expected but not realized consumer response, here there are no welfare losses, but instead simply a welfare transfer from fishermen to processors.

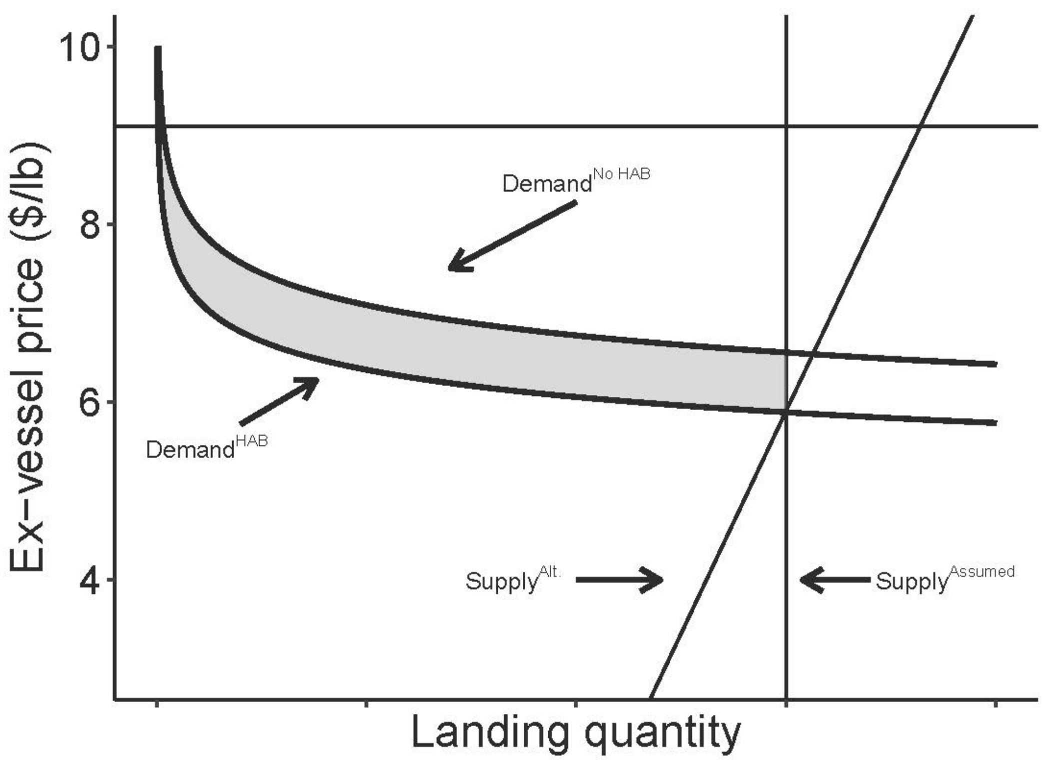


Figure 3.6: Illustration of calculation of the welfare loss

The welfare impacts from the HAB is the area between the estimated demand curve (Demand^{HAB}) and our estimate of the counterfactual demand without the HAB (Demand$^{No\ HAB}$) up to the assumed vertical supply curve (Supply^{Assumed}). If the supply curve is, in fact, upwards sloping as in Supply$^{Alt.}$, the welfare effects will be underestimated.

3.7 Conclusion

Climate change is expected to increase seafood contamination events in the future (Moore et al. 2008; Berdalet et al 2016; Marques et al. 2010). The direct impacts from these events include losses associated with the inability to bring product to market during a closure period. In the case of Dungeness crab, holiday markets are important (i.e. demand is high during Thanksgiving, Christmas, New Year, and Chinese New Year, etc.). Therefore, the direct economic impacts of future climate-driven HABs-related closures, which can span multiple holidays, may be substantial. Additionally, there is the potential for indirect economic impacts, in the form of an avoidance cost, if there is a market response to the contamination event even when crabs are safe for human consumption.

Our work quantitatively explores these indirect economic impacts of the 2015 HAB event on California Dungeness crab fishery using a counterfactual analysis. We find evidence of negative demand impacts in the ex-vessel markets but no evidence of negative demand impacts in the consumer market. Our approach allows us to estimate the total impact of the extended HAB event on outcomes in California, but does not identify the exact mechanism or mechanisms driving the impact as, e.g. a market analysis or structural econometric model potentially could.

Therefore, we provide three alternative hypotheses to explain this apparent disconnect and discuss their efficiency and equity implications respectively. Based on our results, we cannot rule out the case that there were no avoidance costs in the California Dungeness crab fishery, but simply a welfare transfer from processors to fishermen, although possibly due to an anticipated consumer response that was not realized. Developing a better understanding of current Dungeness crab markets is an important area for future work and would shed light on which of our competing hypotheses, or an alternative market mechanism, explains what we observe in the data.

Our results suggest that the indirect economic impacts of contamination events may be broader than previously discussed in the literature and include distributional consequences in addition to avoidance cost. It is important to understand this broader set of indirect economic impacts as fisheries managers work towards both economic and distributional goals in a future with environmental change.

BIBLIOGRAPHY

[1] Alemi, Farzad, Giovanni Circella, Patricia Mokhtarian, and Susan Handy. "Exploring the latent constructs behind the use of ridehailing in California." *Journal of choice modelling* 29 (2018): 47-62.

[2] Almoqbel, Mashael Yousef, and Donghee Yvette Wohn. "Individual and collaborative behaviors of rideshare drivers in protecting their safety." *Proceedings of the ACM on Human-Computer Interaction* 3, no. CSCW (2019): 1-21.

[3] Anselin, Luc. "Interactive techniques and exploratory spatial data analysis." (1996).

[4] Arraiz, Irani, David M. Drukker, Harry H. Kelejian, and Ingmar R. Prucha. "A spatial Cliff‐Ord‐type model with heteroskedastic innovations: Small and large sample results." *Journal of Regional Science* 50, no. 2 (2010): 592-614.

[5] Belk, Russell. "You are what you can access: Sharing and collaborative consumption online." *Journal of business research* 67, no. 8 (2014): 1595-1600.

[6] Berdalet, Elisa, Lora E. Fleming, Richard Gowen, Keith Davidson, Philipp Hess, Lorraine C. Backer, Stephanie K. Moore, Porter Hoagland, and Henrik Enevoldsen. "Marine harmful algal blooms, human health and wellbeing: challenges and opportunities in the 21st century." *Journal of the Marine Biological Association of the United Kingdom* 96, no. 1 (2016): 61-91.

[7] Billingsley, Sunnee, Maria Brandén, Siddartha Aradhya, Sven Drefahl, Gunnar Andersson, and Eleonora Mussino. "COVID-19 mortality across occupations and secondary risks for elderly individuals in the household: A population register-based study." *Scandinavian journal of work, environment & health* 48, no. 1 (2022): 52.

[8] Brown, Charles. "Equalizing differences in the labor market." *The Quarterly Journal of Economics* 94, no. 1 (1980): 113-134.

[9] Burtch, Gordon, Seth Carnahan, and Brad N. Greenwood. "Can you gig it? An empirical examination of the gig economy and entrepreneurial activity." *Management Science* 64, no. 12 (2018): 5497-5520.

[10] Byrne, Stephen, Dermot Coates, Enda Keenan, and Tara McIndoe-Calder. "The initial labour market impact of COVID-19." *Central Bank Economic Letter* 4 (2020): 25.

[11] Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. "Labor supply of New York City cabdrivers: One day at a time." *The Quarterly Journal of Economics* 112, no. 2 (1997): 407-441.

[12] Chen, M. Keith, and Michael Sheldon. "Dynamic pricing in a labor market: Surge pricing and flexible work on the Uber platform." *Ec* 16 (2016): 455.

[13] Chou, Yuan K. "Testing alternative models of labour supply: Evidence from taxi drivers in Singapore." *The Singapore Economic Review* 47, no. 01 (2002): 17-47.

[14] Cramer, Judd, and Alan B. Krueger. "Disruptive change in the taxi business: The case of Uber." *American Economic Review* 106, no. 5 (2016): 177-182.

[15] Dahlgran, Roger A., and Dean G. Fairchild. "The demand impacts of chicken contamination publicity—a case study." *Agribusiness: An International Journal* 18, no. 4 (2002): 459-474.

[16] Didier, Aloysius J. The Pacific coast Dungeness crab fishery. *Pacific States Marine Fisheries Commission*, 2002.

[17] Dills, Angela K., and Sean E. Mulholland. "Ride-sharing, fatal crashes, and crime." *Southern Economic Journal* 84, no. 4 (2018): 965-991.

[18] Dryhurst, Sarah, Claudia R. Schneider, John Kerr, Alexandra LJ Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, David Spiegelhalter, and Sander Van Der Linden. "Risk perceptions of COVID-19 around the world." *Journal of Risk Research* 23, no. 7-8 (2020): 994-1006.

[19] Einav, Liran, Chiara Farronato, and Jonathan Levin. "Peer-to-peer markets." *Annual Review of Economics* 8 (2016): 615-635.

[20] Elinson, Zusha. “Toxic Algae Pinch California’s Crab Season.” *Wall Street Journal*, February 21, (2016).

[21] Farber, Henry S. "Is tomorrow another day? The labor supply of New York City cabdrivers." *Journal of political Economy* 113, no. 1 (2005): 46-82.

[22] Foster, William, and Richard E. Just. "Measuring welfare effects of product contamination with consumer uncertainty." *Journal of Environmental Economics and Management* 17, no. 3 (1989): 266-283.

[23] Greenwood, Brad N., and Sunil Wattal. "Show me the way to go home." *MIS quarterly* 41, no. 1 (2017): 163-188.

[24] Groshen, Erica L. "COVID-19's impact on the US labor market as of September 2020." *Business economics* 55, no. 4 (2020): 213-228.

[25] Hackett, STEVEN C., MATTHEW J. Krachey, CHRISTOPHER M. Dewees, DAVID G. Hankin, and K. R. I. S. T. E. N. Sortais. "An economic overview of Dungeness crab (Cancer magister) processing in California." *California Cooperative Oceanic Fisheries Investigations Report* (2003): 86-93.

[26] Hagenbuch, Brian. "Dungeness Crab Fishermen's Strike Continues on US West Coast." *SeafoodSource*, January 6, (2017).

[27] Hamermesh, Daniel. "Economic aspects of job satisfaction." *Essays in labour market analysis,* Toronto, John Wiley et Sons (1977).

[28] Hausman, Jerry A. "Valuation of new goods under perfect and imperfect competition." *In The economics of new goods*, pp. 207-248. University of Chicago Press, (1996).

[29] Holland, Paul W. "Statistics and causal inference." *Journal of the American statistical Association* 81, no. 396 (1986): 945-960.

[30] Hong Kong Special Administrative Region. "Import and Sale of Dungeness Crabs and Rock Crabs Harvested from Part of California Coast Prohibited Immediately." (2015).

[31] Hughes, Ryan, and Don MacKenzie. "Transportation network company wait times in Greater Seattle, and relationship to socioeconomic indicators." *Journal of Transport Geography* 56 (2016): 36-44.

[32] James, Kenneth, Camelia Thompson, Cameal Chin-Bailey, Kayon Donaldson Davis, Desmalee Holder Nevins, and Dawn Walters. "COVID-19 related risk perception among

taxi operators in Kingston and St. Andrew, Jamaica." *Journal of transport & health* 22 (2021): 101229.

[33] Jensen, Robert. "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector." *The quarterly journal of economics* 122, no. 3 (2007): 879-924.

[34] Jin, Di, Eric Thunberg, and Porter Hoagland. "Economic impact of the 2005 red tide event on commercial shellfish fisheries in New England." *Ocean & Coastal Management 51*, no. 5 (2008): 420-429.

[35] Just, Richard E., Darrell L. Hueth, and Andrew Schmitz. The welfare economics of public policy: a practical approach to project and policy evaluation. *Edward Elgar Publishing*, (2005).

[36] Kalaitzandonakes, Nicholas, Leonie A. Marks, and Steven S. Vickner. "Media coverage of biotech foods and influence on consumer choice." *American Journal of Agricultural Economics* 86, no. 5 (2004): 1238-1246.

[37] Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian. "Vertical integration, appropriable rents, and the competitive contracting process." *The journal of Law and Economics* 21, no. 2 (1978): 297-326.

[38] Lee, Min Kyung, Daniel Kusbit, Evan Metsky, and Laura Dabbish. "Working with machines: The impact of algorithmic and data-driven management on human workers." *In Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pp. 1603-1612. 2015.

[39] Leigh, J. Paul. "Compensating wages for occupational injuries and diseases." *Social Science Quarterly* 62, no. 4 (1981): 773.

[40] Lloyd, Tim, Steve McCorriston, C. W. Morgan, and Anthony J. Rayner. "The impact of food scares on price adjustment in the UK beef market." *Agricultural Economics* 25, no. 2-3 (2001): 347-357.

[41] Mao, Junwei, and Sunny L. Jardine. "Market impacts of a toxic algae event: The case of california dungeness crab." *Marine Resource Economics* 35, no. 1 (2020): 1-20.

[42] Marques, António, Maria Leonor Nunes, Stephanie K. Moore, and Mark S. Strom. "Climate change and seafood safety: Human health implications." *Food Research International* 43, no. 7 (2010): 1766-1779.

[43] Martin-Buck, Frank. "Driving safety: An empirical analysis of ridesharing's impact on drunk driving and alcohol-related crime." *University of Texas at Austin Working Paper* (2017).

[44] Mazzocchi, Mario. "No news is good news: Stochastic parameters versus media coverage indices in demand models after food scares." *American Journal of Agricultural Economics* 88, no. 3 (2006): 727-741.

[45] McCabe, Ryan M., Barbara M. Hickey, Raphael M. Kudela, Kathi A. Lefebvre, Nicolaus G. Adams, Brian D. Bill, Frances MD Gulland, Richard E. Thomson, William P. Cochlan, and Vera L. Trainer. "An unprecedented coastwide toxic algal bloom linked to anomalous ocean conditions." *Geophysical Research Letters* 43, no. 19 (2016): 10-366.

[46] Mehranbod, Christina A., Ariana N. Gobaud, and Christopher N. Morrison. "Ridesharing and alcohol-related assaults in NYC: A spatial ecological case-crossover study." *Drug and alcohol dependence* 232 (2022): 109321.

[47] Menéndez, Cammie KC, Harlan E. Amandus, Parisa Damadi, Nan Wu, Srinivas Konda, and Scott A. Hendricks. "Effectiveness of taxicab security equipment in reducing driver homicide rates." *American journal of preventive medicine* 45, no. 1 (2013): 1-8.

[48] Nafilyan, Vahe, Piotr Pawelek, Daniel Ayoubkhani, Sarah Rhodes, Lucy Pembrey, Melissa Matz, Michel Coleman et al. "Occupation and COVID-19 mortality in England: a national linked data study of 14.3 million adults." *Occupational and Environmental Medicine* 79, no. 7 (2022): 433-441.

[49] Nielsen, Max, Jos Smit, and Jordi Guillen. "Price effects of changing quantities supplied at the integrated European fish market." *Marine Resource Economics* 27, no. 2 (2012): 165-180.

[50] Park, Jiyong, Junetae Kim, and Byungtae Lee. "Are Uber really to blame for sexual assault? Evidence from New York city." *In Proceedings of the 18th annual international conference on electronic commerce: e-Commerce in smart connected world*, pp. 1-7. 2016.

[51] Park, Jiyong, Min-Seok Pang, Junetae Kim, and Byungtae Lee. "The deterrent effect of ride-sharing on sexual assault and investigation of situational contingencies." *Information Systems Research* 32, no. 2 (2021): 497-516.

[52] Rayle, Lisa, Danielle Dai, Nelson Chan, Robert Cervero, and Susan Shaheen. "Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco." *Transport Policy* 45 (2016): 168-178.

[53] Region, Marine. "Review of selected California fisheries for 2014: Coastal pelagic finfish, market squid, groundfish, pacific herring, dungeness crab, ocean salmon, true smelts, hagfish, and deep water ROV surveys of MPAs and surrounding nearshore habitat." (2015).

[54] Shokoohyar, Sina. "Ride-sharing platforms from drivers' perspective: Evidence from Uber and Lyft drivers." International Journal of Data and Network Science 2, no. 4 (2018): 89-98.

[55] Shulstad, Robert N., and Herbert H. Stoevener. "The effects of mercury contamination in pheasants on the value of pheasant hunting in Oregon." *Land Economics* 54, no. 1 (1978): 39-49.

[56] Smith, Mark E., Eileen O. Van Ravenswaay, and Stanley R. Thompson. "Sales loss determination in food contamination incidents: an application to milk bans in Hawaii." *American Journal of agricultural economics* 70, no. 3 (1988): 513-520.

[57] Sommer, Lauren. "Why A Neurotoxin Is Closing Crab Season In California." *National Public Radio*, November 7, (2015). https://www.npr.org/sections/thesalt/2015/11/07/455049713/why-a-neurotoxin-is-closing-crab-season-in-california.

[58] Stafford, Tess M. "What do fishermen tell us that taxi drivers do not? An empirical investigation of labor supply." *Journal of Labor Economics* 33, no. 3 (2015): 683-710.

[59] Stephanie, Moore, Trainer Vera, Mantua Nathan, Parker Micaela, Laws Edward, Backer Lorraine, and Fleming Lora. "Impacts of climate variability and future climate change on harmful algal blooms and human health." (2008).

[60] Sundararajan, Arun. The sharing economy: The end of employment and the rise of crowd-based capitalism. MIT press, 2017.

[61] Swartz, David G., and Ivar E. Strand. "Avoidance costs associated with imperfect information: The case of kepone." *Land Economics* 57, no. 2 (1981): 139-150.

[62] Talley, Ken. "No Title." *Seafood Trend Newsletter*, November (2011).

[63] Thaler, Richard, and Sherwin Rosen. "The value of saving a life: evidence from the labor market." *In Household production and consumption*, pp. 265-302. NBER, 1976.

[64] Thebault-Spieker, Jacob, Loren Terveen, and Brent Hecht. "Toward a geographic understanding of the sharing economy: Systemic biases in UberX and TaskRabbit." *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, no. 3 (2017): 1-40.

[65] Tobias, Lori. "Oregon's Dungeness Crab Learn to Fly to China for a Whole New Market." *The Oregonian*, March (2011).

[66] Viscusi, W. Kip. "Union, labor market structure, and the welfare implications of the quality of work." *Journal of Labor Research* 1, no. 1 (1980): 175-192.

[67] Viscusi, W. Kip. Employment hazards: An investigation of market performance. No. 148. Harvard University Press, 1979.

[68] Wallsten, Scott. "The competitive effects of the sharing economy: how is Uber changing taxis." *Technology Policy Institute* 22, no. 3 (2015).

[69] Weber, Bryan S. "Uber and urban crime." *Transportation research part A: policy and practice* 130 (2019): 496-506.

[70] Webster, Allan, Sangeeta Khorana, and Francesco Pastore. "The labour market impact of COVID-19: early evidence for a sample of enterprises from Southern Europe*." International Journal of Manpower* (2021).

[71] Wessells, Cathy Roheim, Christopher J. Miller, and Priscilla M. Brooks. "Toxic algae contamination and demand for shellfish: a case study of demand for mussels in Montreal." *Marine Resource Economics* 10, no. 2 (1995): 143-159.

[72] Wessells, Cathy Rohein, and James E. Wilen. "Seasonal patterns and regional preferences in Japanese household demand for seafood." *Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie* 42, no. 1 (1994): 87-103.