FOODSERVICE BUSINESS CLOSURES IN ATHENS, GEORGIA: A SURVIVAL ANALYSIS

by

Bishal Neupane

(Under The Direction of Daniel Remar)

ABSTRACT

Foodservice businesses- restaurants, caterers, bars, recreational centers, bakeries and similar other establishments suffer closures due to internal and external factors. This study compares survival rates and evaluates the impact of chain affiliation, and health inspection scores on the closures of four different segments of foodservice businesses in Athens, Georgia, from the year 2000 to 2020. It utilizes non-parametric Kaplan-Meier Survival analysis to determine the survival rates of total 841 businesses born, died and censored within the study period. It was observed that the restaurants having chain affiliation and higher inspection grades positively affected the probability of survival and, limited-service restaurants were found to have greater longevity than full-service restaurants, bars, pubs, taverns and the other segments of foodservice sector. A Cox's proportional- hazard regression was also utilized which identified significant impact of chain affiliation and grades on the closure of foodservice businesses in Athens.

Keywords: survival analysis, foodservice business closure, restaurants, bars, Athens

FOODSERVICE BUSINESSES CLOSURES IN ATHENS, GEORGIA: A SURVIVAL ANALYSIS

by

BISHAL NEUPANE

B.S., Agricultural and Forestry University, Nepal, 2017

A Thesis Submitted to the Graduate Faculty of the University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2021

FOODSERVICE BUSINESSES CLOSURES IN ATHENS, GEORGIA: A SURVIVAL ANALYSIS

by

BISHAL NEUPANE

Major Professor: Daniel Remar

Committee: John Salazar

Susana Ferreira

Electronic Version Approved by:

Ron Walcott, Vice Provost for Graduate Education and Dean of the Graduate School

The University of Georgia

August 2021

This thesis is dedicated to Mukti Pandit, Sharmila Neupane and Sarbagya Dhimal

ACKNOWLEDGEMENT

I would like to express my sincere gratitude towards my committee chair Dr. Daniel Remar for his support and guidance throughout this research. His inputs and suggestions were extremely vital for the development and completion of this study. I am also grateful towards my committee members Dr. John Salazar and Dr. Susana Ferreira for their constructive criticisms. I am thankful to Andrea Kerr, Environmental Health County Manager, Clarke County, for providing me with Health inspection data and for patiently explaining all of my queries.

Thank you my parents, and Amul for every little thing, and Mr. Chandra Bhakta and Mrs. Tara Bhakta for my guardians here in Athens.

Finally, special appreciation goes to my favorite Bulldawgs and friends- Shweta, Dhiraj, Bimal, Dikshit, Tulsi, Shreena, Swikriti, Sujan, Shiwani, Subin dai, Rosan, Sabin dai, Manoj dai and Ritika.

TABLE OF CONTENT

ACKNOWLEDGEMENT .... **V**
LIST OF FIGURES .... **VIII**
LIST OF TABLES .... **IX**
CHAPTER:1 INTRODUCTION **1**
1.1 Foodservice industry in the USA 1
1.2 Categories of foodservice businesses in the USA 3
1.3 Foodservice business closures 5
1.4 Athens-Clarke County (demographic information) 6
1.5 Foodservice businesses in Athens and their closures 7
1.6 Research Objectives 10
CHAPTER:2 LITERATURE REVIEW **11**
2.1 Survival Analysis 11
2.2 Chain and independent foodservice businesses 13
2.3 Health inspection scores 14
2.4 Foodservice segments 15
CHAPTER:3 METHODOLOGY **17**
3.1 Data retrieval 17
3.2 Data Cleaning 17
3.3 Grade variable 18
3.4 Chain variable 19
3.5 Segment variable 19
3.6 Data Analysis 21
CHAPTER:4 FINDINGS AND RESULTS **25**
4.1 Kaplan-Meier Survival Analysis 25
4.2 Summary Statistics 26
4.3 Interpretation of Survival function 29

4.4 Testing of equality of survivor function for chain variable 32
4.5 K-M survival analysis based on Chain 32
4.6 Testing of equality of survivor function for Grades variable 35
4.7 Kaplan-Meier Survival Analysis based on Grades 35
4.8 Testing of equality of survivor function for segment 1 38
4.9 Kaplan-Meier Survival analysis based on segment 1 38
4.10 Testing of equality of survivor function for segment 2 40
4.11 Testing of equality of survivor function for segment 3 41
4.12 Cox's proportional hazard Regression 42
4.13 Test of proportional-hazard assumption 43
CHAPTER:5 DISCUSSIONS 44
5.1 Survival function for Athens foodservice establishments 44
5.2 Survival function and chains 45
5.3 Survival function and health inspection grades 45
5.4 Survival functions and segments 46
5.5 Implications 48
5.6 Future Research 48
References 49
Appendix A 53

LIST OF FIGURES

Figure 1 Value added to GDP by agriculture and related industries, 2009-19..............................1

Figure 2 Employment in agriculture, food and related industries, 2019.......................................2

Figure 3 Restaurant industry sales, 1970-2020; source: NRA 3

Figure 4 Businesses in Athens based on their creation date, 2019; source: ACC finance dept. 7

Figure 5 proportion of restaurant businesses in Athens, 2019; source: ACC finance dept.............8

Figure 6 K-M survival curve for foodservice businesses in Athens, 2000-2020 30

Figure 7 K-M survival curve for chain variable 34

Figure 8 K-M survival curve for Grades variable...36

Figure 9 K-M survival curve for Grades variable without confidence interval 37

Figure 10 K-M survival curve for Segment 1...39

LIST OF TABLES

Table 1 Sales from different segments of foodservice industry 2019-20; source: NRA 4

Table 2 Examples of variables created by stset 26

Table 3 Description of survival dataset 27

Table 4 Summary of survival dataset 28

Table 5 Description of survival variables 28

Table 6 K-M survivor function for foodservice businesses in Athens 31

Table 7 Log-rank test, chain variable 32

Table 8 K-M survivor function for chain variables 33

Table 9 Log-rank test for Grades variable 35

Table 10 Log-rank test for Segment 1 38

Table 11 Log-rank test for Segment 2 40

Table 12 Log-rank test for Segment 3 41

Table 13 Cox-regression table with coefficients 42

Table 14 Cox regression table with hazard ratio 43

Table 15 Schoenfeld test for proportional-hazard assumption 43

CHAPTER: 1

INTRODUCTION

1.1 Foodservice industry in the USA

According to (USDA, 2020), Agriculture, food and related industries had a contribution of $1.109 trillion, 5.2 percent of the total share, to the U.S. Gross Domestic Product (GDP) in 2019. These sectors include food and beverage manufacturing, food and beverage stores, food services and eating and drinking places, textiles, apparel, and leather products and forestry and fishing. The

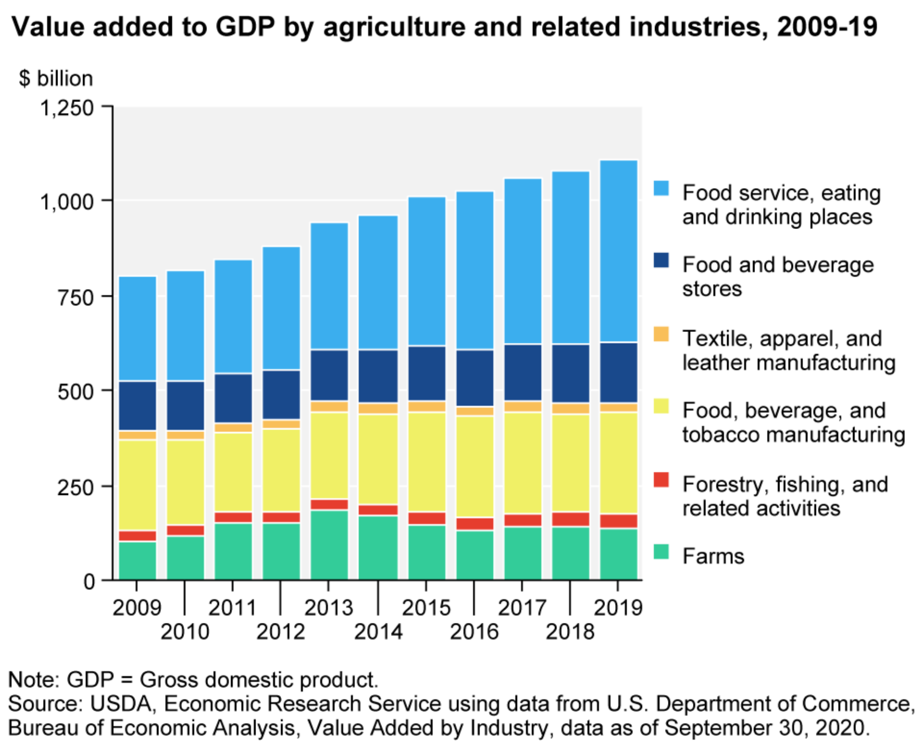

Figure 1 Value added to GDP by agriculture and related industries, 2009-19

annual sales of food service, eating and drinking places alone amounted to $864.3 billion in 2019 (NRA, 2021).

Agriculture, food and related industry also accounted for 22.2 million jobs, approximately 10.9 percent, of the total employment in the US. The largest employment share was from foodservices and eating places – 13 million jobs. The sector's employment is larger than oil, mining, construction, and transportation industries (Bureau of Labor Statistics, 2020).

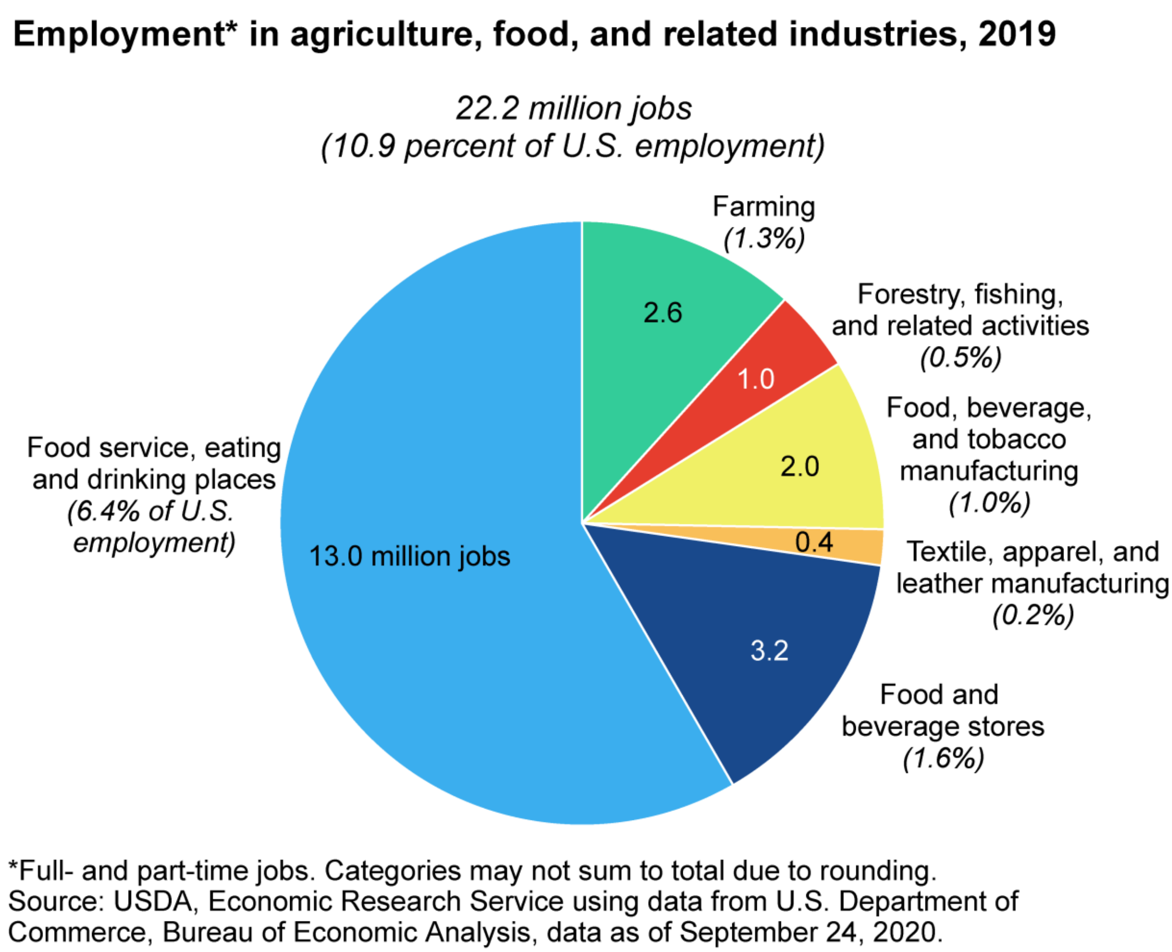

Figure 2 Employment in agriculture, food and related industries, 2019

With average American spending about 13 percent of their total income on food, it is ranked third among the American Household expenditures, behind housing and transportation. With its substantial contribution to the economy, employment, and per capita expenses, it is sufficient to say that foodservice industry is vital for the American economy. So, any form of inherent factors or economic crisis that leads to the closure of foodservice industry can have substantial impact on the economy, be it small or large. Even a small college-town like Athens, Georgia can have

1.2 Categories of foodservice businesses in the USA

The National Restaurant Association lumps full-service restaurants (NAICS code 722511), limited-service restaurants (NAICS code 722513), bars and taverns (NAICS code 722410) and other establishments into one category of 'restaurant and food service industry'. Restaurant sector is further divided into Full-service and Limited-service segment. There were over 1 million restaurants in the United Sates as of 2020. The restaurant and foodservice industry employs about 13 million people (NRA, 2021). The (NRA, 2019) report shows that restaurant industry sales

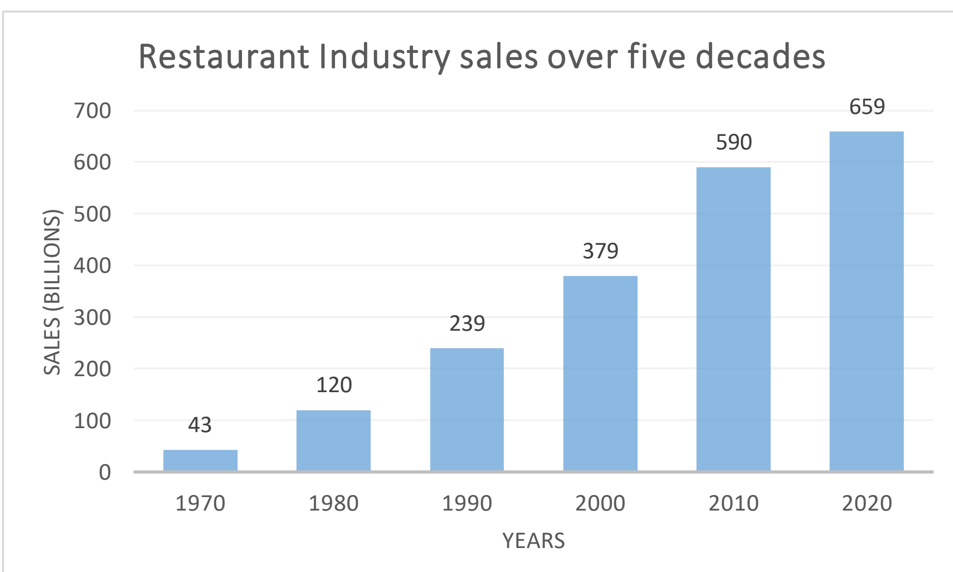

Figure 3 Restaurant industry sales, 1970-2020, inflation adjusted; source: NRA

during 1970 was only $43 billion. The sales figure has increased by almost 18 folds from 1970 to 2020 as shown in figure: 3.

Of the total restaurant and foodservice sales in the United States in 2019, the limited-service segment contributed the highest, followed by full-service segment, bars and taverns, and all other foodservice establishment.

Table 1 Sales from different segments of foodservice industry 2019-20; source: NRA

	2019 Sales (Billions)	**2020 Sales (Billions)**	**2021 Sales (Billions) Projection**
Eating and Drinking Places	**$615.9**	**$497.6**	**$548.3**
Full-service segment	$285.0	$199.5	$220.8
Limited-service segment	$308.9	$290.4	$313.6
Bars and Taverns	$22.0	$7.7	$13.9
Other establishments	$248.4	$161.5	$183.2
Total	**$864.3**	**$689.0**	**$731.5**

There are also, single/independent and chain foodservice establishments. In 2018 alone, there were 352,815 independent restaurants and 307,914 chain restaurants (Statista, 2019). In 2019, the chain restaurants in the US generated about $159 billion and single-location full-service establishments generated a revenue of more than $184 billion. With the onset of pandemic however, these figures dwindled and for the first time since 2008/09, there was a negative annual change of (34 percent) net revenue in 2020 from 2019 in the chain restaurant sector. The independent sector also suffered a loss of 14 percent which was also the first negative annual net revenue growth figure since 2009 (IBISWorld, 2021).

1.3 Foodservice business closures

Restaurants and hospitality businesses are capital intensive, labor intensive and highly leveraged which makes them susceptible to effects of the vicissitudes of economy (Parsa.H.G., Kreeger.J.C., Rest.J.P, Xie.L, & Lamb.J., 2019). It was previously believed that 90 percent of the restaurants failed in the first year of their establishment. This myth plagued the industry for several years until it was found that only 17 percent of independently owned full-service restaurant startups failed in the first year, compared to 19 percent for all other service providing startups. The median lifespan of the restaurants was also found to be 4.5 years and the median lifespan of restaurant which had 5 or less than 5 employees was 3.75 (Luo & Stark, 2014).

The closure of these businesses could result from several different aspects. It could have managerial, ecological, psychological and many other marketing reasons and implications. One of the reasons for fatality of restaurant businesses is due to mistakes in marketing mix strategies such as expenditures, products, and services (Bertagnoli, 2005). Other scholars suggest that restaurants get weeded out due to natural selection (Hannan B. &., 1989).

Nevertheless, restaurant failure and closure are an inevitable part of the foodservice industry. Failure of businesses refers to bankruptcy or going out of business, it occurs when the business is suffering and not doing well. On the other hand, closure can occur even if the businesses are successful. In fact, 29 percent of the restaurants who closed their operation were successful at the time of closure (Luo & Stark, 2014). Closure can occur because of health-related issues or logistic reasons or when the owners venture out to new businesses or maybe because of an economic recession. This study specifically identifies the effect of health inspection scores, and chain affiliation on the survival rate of four different segments of foodservice industry in Athens, Georgia from 2000 to 2020.

1.4 Athens-Clarke County (demographic information)

Athens is the smallest of the 159 counties in Georgia, located 70 miles northeast from the major city of Atlanta. Clarke was created in the year 1801 as the twenty-fifth county of Georgia while the town of Athens was chartered in 1806 A.D. In the year 1991, the Clarke County and the city of Athens governments unified and created Athens-Clarke unified government.

As per the 2014-2019 census data, there are 126,913 residents in Athens-Clarke including the students in UGA. Of the total population, 48 percent are male, and the rest are female. As for the racial composition, 66 percent of the residents in Athens-Clarke are white, a little more than 29 percent are black, almost 5 percent are Asian and the remaining are Hispanic, Latinos and other races. The median age of the population is 28 years, and the median household income is $38,311. Almost, 30 percent of the residents in this unified government live below poverty level and of the total population, about 5 percent are unemployed. There are 15 pre-kindergartens, 14 elementary, 4 middle and 3 high-schools in the county along with the biggest employment provider- The University of Georgia (Athens-Clarke County Unified Government, 2021).

There are 19 zip-codes associated with Athens-Clarke County Metro Area Georgia, however, this study is limited to the town of Athens with zip codes, 30601, 30602, 30605, 30606, 30607 and 30609. According to Discoverer, 2020, it is the coolest small town in America and as per Southern Living, 2020, it is the South’s best college town. With 37,283 students enrolled in the University of Georgia, Athens has number of bars, pubs, and other restaurant establishments located in the downtown and its outskirts that serve the students and the residents.

1.5 Foodservice businesses in Athens and their closures

According to Athens-Clarke County Finance Department Report on business licensing and tax listings, there were 5393 businesses in operation in the year 2019. The annual report also has information on the creation dates of the businesses along with the name of the proprietors or owners and the sector of the industry they are in.

Fig: 4 displays the number of total businesses in the town of Athens that were in operation in 2019 and their dates of creation as listed in the Athens Business Occupation Tax listings. The lowest number of businesses in 2019 were from the year 2001 and the highest were from 2019 itself.

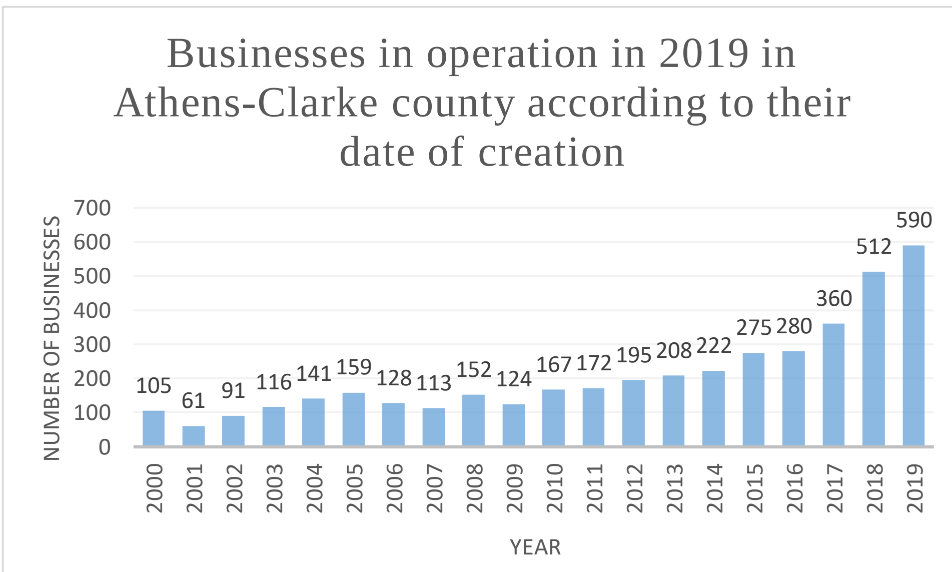

Figure 4 Businesses in Athens based on their creation date, 2019; source: ACC finance dept.

According to Fig 5, the highest percentage of restaurant with respect to the businesses belonged to the year 2014 and 2015, and the lowest figure is from the year 2010. On average, for every hundred businesses in operation in Athens in the year 2019, about 6 were restaurant establishments.

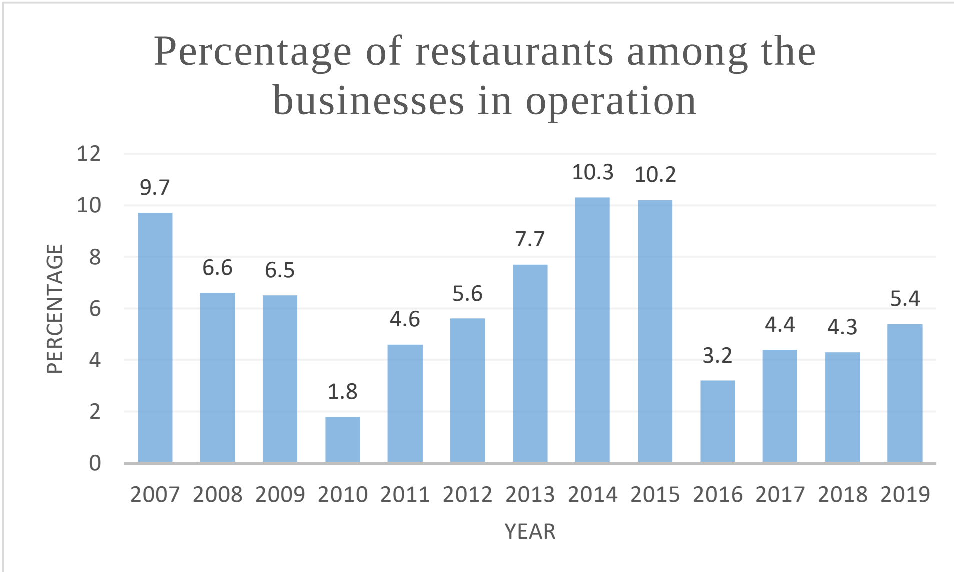

Figure 5 proportion of restaurant businesses in Athens, 2019; source: ACC finance dept.

The restaurants and bars have been one of the major contributors to the Athens' economy. From January through October in 2019, they earned on an average $19.5 million through liquor sales only. In 2016, the bars and restaurant sector had seen a rise of $300,000 on liquor sales from July to August in 2016. The sales soared to $500,000 in 2018 and 2019 during the same time period (The Red & Black, 2020).

However, with the onset of novel corona virus in late 2019, Governor Brian Kemp announced Statewide 'Shelter in place' Executive order (georgia.gov, 2021). Shelter in place, April 2,2020 encouraged citizens to remain in their houses except when they needed to carry out essential activities like going to the grocery store or doctors or carrying out essential business travels.

For businesses, they were categorized into essential (grocery stores, health providers, hardware stores, banks, etc.) and non-essential (includes bars, nightclubs, amusement parks, and live performance venues). The essential businesses could operate while the non-essential were temporarily closed. According to the August reports of Atlanta Hospitality Alliance, the restaurant business in Georgia was down by ninety-five percent; the situation was even worse for pubs and bars who faced stricter restrictions. In Athens, after the hospitality industry lost almost forty percent of yearly revenue due to cancellation of graduation ceremony and spring closures (Grady Newsource, 2020). Most of the bars, pubs and hotels in Athens were closed after March, and a few even went out of business. With looser social distancing regulation at the beginning of the fall session, and limited capacity football games in effect, many businesses started to reopen but a few of them shut down permanently (Athens Banner Herald, 2020). So, naturally few questions were raised, are these reports of restaurant closures as serious as it was mentioned in the media? Is this situation worse in Athens? How does it fare against the national average?

Since, restaurant and food service businesses are so crucial for Athens, it is essential to study the survival and risk of closures of such businesses. And as these closures occurred due to a health crisis, this study observed the relationship between health inspection scores and closures. Likewise, multiple reports of independent foodservice closures (National Restaurant Association, 2020) presented an avenue to compare the survival/closure rates among chains and independent sectors and further within four different sectors of foodservice businesses- full-service restaurants, limited-service restaurants, bars/pubs/taverns and 'other'. Which is why this study explored the impact of major explanatory variables like health inspection scores, chain affiliation status and type of restaurant segment on the survival rate of foodservice establishments. To determine this the research utilized non-parametric Kaplan-Meier model along with Cox's proportional hazard regression. This will be its new contribution to the literature.

1.6 Research Objectives

- to compare survival rates of foodservice businesses based on their health inspection grades
- to compare the survival rates of foodservice businesses based on their affiliation with a chain/franchise
- to compare the survival rates of four different segments of foodservice businesses
- to compare the survival rates before and during the COVID-19 pandemic

CHAPTER:2

LITERATURE REVIEW

2.1 Survival Analysis

Survival analysis comprises several techniques that are used to measure a positive value random variable. The random variable generally refers to the time to the failure of a physical component. Here failure might include death, shutdown, and closure and the physical component might include patient, animals, cells, businesses etc. This method was developed centuries ago to work on mortality tables but in a modern world it was used during World War II to simulate the interest in the reliability of military equipment and machineries. However, since the sixties it has been extensively used in the field of medical research (Miller, Gong, & Munoz, 1981).

In the field of medical research, survival analysis has been used to determine the rate of survival of patients who had been given a specific treatment. One such research studied the survival rates of 100,000 Japanese cancer patients who had received gastric surgical treatment, this study was conducted between 2001and 2007. It was found that the death rate of patients within 30 days of surgery was 0.5 percent whereas, the 5-year survival rates of patients who received the operation was 71.1 percent (Katai, et al., 2018). Another study conducted by Taiwanese researchers identified the survival rates of pediatric dialyses patients in Taiwan. It was found that the 1-,5-, and 10-years rates of survival for peritoneal dialysis (PD) patients were 98.1, 88.0 and 68.4 percent respectively (Lin, et al., 2012).

Similarly, in the field of Hospitality and Tourism, this technique has been utilized to determine the length of stay of tourists. This was first conducted by Gokovali, Bahar, & Kozak (2007) to measure the length of stay of Turkish tourists who were departing from the region of Bodrum. This was followed by other studies like the one conducted in Spain regarding the length of stay of tourists, destination choice and the effect of High-speed rail using continuous survival model. It determined that HSR played small role on tourist's decision to visit Costa-Daurada in Spain. And their length of stay was influenced by HSR only when they stayed overnight in second residences (Gutie´rrez, Miravet, Saladie, & Clave, 2020). Likewise, another study by Aguilar, & Diaz (2019) identified the length of stay of tourists and the effect of socio-demographic variables on the dependent variable.

Survival analysis technique has also been used to predict the closure and bankruptcy of hospitality businesses. A study conducted by Gemar, Soler, & Guzman-Parra (2019) assessed the influence of explanatory variables on the resort hotel's closure and the degree of risk posed for the bankruptcy of the establishment. They found that the business closure was depended on size, location, executive management and business cycle. However, the survival was neither affected by the type of hotels nor their financial structure. Likewise, previous research by Lado-Sestayo, Vivel-Bua, & Otero-Gonzalez (2016) on the survival of 6494 lodging hotels in Spain between 2005-2011 applied and determined that location had a significant hand on the probability of survival of such businesses. Moreover, they also found that level of occupancy had a positive effect on the dependent variable whereas, low level of competition actually decreased the average expected survival. The research also identified that the total effect related to destinations reduced expected survival by about 90 percent.

When examining restaurant business closures, Parsa, Self, Sydnor-Busso, & Yoon (2011) used ecological perspective to determine the failure rates of restaurant establishments in Cobb County, Georgia. They studied the impact of affiliation of restaurants with a franchise, location (based on zip-codes) and size (employee number) on restaurant failures. The researchers concluded that all the variables significantly contributed to the mortality of restaurant business. In terms of affiliation, a chain restaurant had a significantly lower rate of failure when compared with an independently owned restaurant. Additionally, restaurants with larger employment and enhanced services, had lesser failure rates. It was also established that restaurants belonging to certain zip codes had higher failure rates than others.

Similar study conducted by Luo, & Stark (2004) on mortality of startup restaurants utilized nonparametric maximum likelihood estimator of survival function i.e., Kaplan-Meier Product Limit estimator. The research was done using the US. Bureau of Labor Statistics data of 81,000 full-service restaurants over the period of 20 years. It rejected the myth that 90 percent of restaurant businesses failed in their first year of establishment and concluded that only 17 percent of the restaurant startups failed during their first year of inception which was 2 percent lower than other service-providing startups. They also calculated the median lifespan of the restaurants as 4.5 years which was higher than the median lifespan of other service providers. They opined that unlike the commonly held belief restaurant businesses were actually safer to start which was complimented by the fact that they had very nominal barriers of entry.

2.2 Chain and independent foodservice businesses

Studies on chain and independent foodservice businesses range from dining experiences to food safety and even social media. For instance, a research on comparison of casual dining experience in chain and independent restaurants by, Young, Clarke, & McIntyre (2008) found that consumers considered the dining experience in chain establishments to be more favorable to independent ones. Another study by Harris, DiPietro, Murphy, & Rivera (2014) delved into health inspection scores and chain as well as non-chain statuses of restaurants- both our research interests. It indicated that the number of health safety violations by the restaurants are affected by their locations as well as their chain and independent status.

English (1996) concluded that the chain and franchise restaurants have better investments and greater success rate. Similarly, study on failure rates and chain/independent nature of businesses by Parsa, Self, Njite, & King (2005) found that the failure rates for independent restaurants were higher than that of chain-restaurants. This could be because chains and franchises obtain congenital information, knowledge, skills and brand recognition/trademarks from their owners which the independent businesses do not. Local chains and franchises have knowledge regarding suitable business location and there is tacit information exchange between managers at close proximity which might give them an upper hand over independent foodservice businesses (Klanins, & Mayer, 2004).

H1: Chain foodservice businesses have higher survival rates than independent ones

2.3 Health inspection scores

Researches on health inspection scores are mostly related to food borne diseases, consumer perception and training of the workers. Some of the studies have been conducted by Jones, Pavlin, LaFleur, & Schaffner, 2004, Lee, Almanza, Nelson, & Ghiselli, 2009, and Mathias, Sizto, Hazlewood, & Cockedge, 1995.

A paper by Behler, Norton, & Sen, (2008) found that franchised stores receive significantly better scores than company owned outlets. There were also few chains where company owned stores received better scores than franchised counterparts.

In terms of health inspection scores and failures, a study by Parsa, Kreeger, Rest, Xie, & Lamb (2019) found that the rise in number of health code violations by restaurants increases their probability of failures. It suggests that greater the health inspection grades, better will be the chances of survival of foodservice businesses.

One of the reasons for this could be because consumer prefer foodservice businesses that have greater inspection scores. They also opine that those businesses that receive poor scores should also be shut down until they meet with all the health requirements (Jones, & Grimm, 2008). Also consumers have access to reviews online, a study by Kong, 2020 also found a significant relationship between better restaurant reviews and lower critical violations which is why customers do not prefer visiting restaurants with lower review scores.

H2: Foodservice businesses which receive greater inspection grades have higher rates of survival

2.4 Foodservice segments

An article published in Tourism Economics by Kim, Zhong, Chen, & Karadag (2009) on three restaurant segments- buffet, quick-service and full-service concluded that quick-service restaurants outperform the other two segments in terms of risk-adjusted performance. Another relevant article by Zheng, Farrish, & Wang (2012) also illustrated that limited-service restaurants outperform full-service restaurants during economic crisis and that the former is actually recession proof.

Lee, & Ha (2004) found a correlation between customer's disposable income and sales of full-service restaurants. Since, the major customers of Athens' foodservice businesses are students, this might affect the sales of full-service restaurants.

There are no empirical studies that examine the survival rates of the foodservice businesses and their segments however, based on the above studies it can be hypothesized that the survival rates of limited-service restaurants would be better than other segments.

H3: The survival rates of limited-service restaurant are higher than other segments

CHAPTER:3

METHODOLOGY

3.1 Data retrieval

Data on Health inspection scores for this study was provided by Andrea Kerr, Environmental Health County Manager, Northeast Health District- Environmental Health Office, Clarke County. Similarly, data on business listings was obtained from Athens-Clarke County Unified Government's Business Occupation Tax listings (https://www.accgov.com/138/Business-Occupation-Tax). And database on small businesses in Athens was retrieved from Reference-USA Business Database also referred as Data axle (http://www.referenceusa.com/UsBusiness/Result/d64a1a6028a345b4969f061d74b8bb0f). The latter two were used as reference for the first dataset.

3.2 Data Cleaning

Health Inspection scores from the year 2000 to 2020 was obtained from the Northeast Health District, Clarke County. Obtained raw data had information on foodservice establishment's ID (given by the DPH office), name of the business, inspection ID, date of inspection, type of inspection (initial, follow-up and routine) and grade. There were altogether 17,833 observations from the establishments within a period of two decades.

However, the dining halls from different institutions like schools, correctional facilities, Athens-Clarke County offices, UGA, Aramark and Sanford were eliminated as they have a different business model and policy.

Likewise, old businesses whose location or segment could not be identified using the reference databases, YELP (https://www.yelp.com/collection/xRzTqrKHWJ1xz9RC90QoQg) , Trip advisor (https://www.tripadvisor.com/Restaurants-g29209-Athens_Georgia.html) , Athens dining (http://www.athensdining.com/) and Athens Area Chamber of Commerce (https://business.athensga.com/list) were also eliminated. This brought down the number of observations to 14,015 over the period of twenty years.

3.3 Grade variable

As per the policy of Environmental Health Office, health inspection of foodservice businesses is conducted at least once a year, they are termed as routine inspection. Initial inspection is organized right after the establishment of businesses. And if a foodservice business scores low, then a follow-up inspection is also done. Thus, a single establishment can undergo several inspections within a year, sometimes even routine inspections are performed more than once for a single business. Previously the scores would be numbers ranging from 1 to 100 but since 2020 the Clarke County health inspection office has started grading those inspections where scores from 90-100 are graded as A, 80-90 as B and then 70-80 as C. Each health inspection score for every institution is tabulated in its original value and then categorized into three grades under Grade variable during the survival analysis.

3.4 Chain variable

The Chain dummy is concerned whether a business is associated with a chain or not. Chains should not be confused with franchises; to be a franchise a business has to obtain a license which gives them an access to their franchisor's proprietary information, technologies and trademarks and such businesses operate under the franchisor's name, and in return the franchisee has to pay a start-up and annual fees (Hayes, 2021). And as for chains they operate under a parent company which does not sell its proprietary rights, knowledge or skills but expands businesses in more than one location (CHRON, 2020). Since chains and franchises have an advantage over an independent business in terms of skills, knowledge and trademarks both of them are considered a single unit and the independent ones are considered another unit.

In this study in order to be considered a chain, a business should have obtained more than one business occupation license under the same company name or should be a part of a franchise. Which is why a chain dummy ($\delta_0$chain) is created for this variable. If a business has establishments in more than one location or if the establishment is part of a franchise, 'Chain' is assigned 1, 0 otherwise.

3.5 Segment variable

In this study, the foodservice sector in Athens has been divided into four segments- Full-service restaurants, Limited-service restaurants, Bars/Pubs/Taverns/Nigh clubs, and others. The NAICS (North American Industry Classification System) categorizes restaurants into Full-service restaurants (NAICS code 722511) and Limited-service restaurants (NAICS code 722513). According to NAICS Association, Full-service restaurant "comprises establishments primarily engaged in providing food services to patrons who order and are served while seated (i.e., waiter/waitress service) and pay after eating. These establishments may provide this type of food service to patrons in combination with selling alcoholic beverages, providing carryout services, or presenting live nontheatrical entertainment". Similarly, Limited-service restaurant "comprises establishments primarily engaged in providing food services (except snack and nonalcoholic beverage bars) where patrons generally order or select items and pay before eating. Food and drink may be consumed on premises, taken out, or delivered to the customer's location. Some establishments in this industry may provide these food services in combination with selling alcoholic beverages." Likewise, for bars/taverns/pubs and nightclubs are categorized under NAICS code 722410 and they are defined as "establishments known as bars, taverns, nightclubs, or drinking places primarily engaged in preparing and serving alcoholic beverages for immediate consumption. These establishments may also provide limited food services." These are the three main categories of foodservice businesses in Athens. The rest of the establishments like ice-cream place, yogurt shops, theatres, bowling alleys, candy stores, and baking goods retailers (cookies, pretzels, cookies and bagels) are classified as 'others.' It has been done because these businesses do not have enough observations to be categorized as a single segment as their number will be smaller than the number of variables in the model.

Since, there are four categories in Segment variable, they are assigned with three (k-1) dummies, i.e., δ_1Segment1, δ_1Segment2 and δ_1Segment3. If a business is listed as Limited-service restaurant, then it is assigned under Segment1. Likewise, if a business is listed as Bars/Pubs/Taverns/Nightclubs, then it is categorized as Segment 2. Finally, if a business is listed as ice-cream place, yogurt shop, theatre, bowling alley, candy store and baking goods retailer then they are referred as Segment3.

3.6 Data Analysis

Closures of restaurants occur continuously over time, which means that the dependent variable is time-variant and also there is presence of right censoring which causes OLS to produce biased results hence it cannot be used for determining the closures of foodservice establishments. The best method to analyze foodservice business closure is survival analysis. It accounts for time-variations, right-censoring and cross-sectional data (Cameron & Trivedi, 2005).

Survival Analysis

The dependent variable (Y) indicates the time till the occurrence of an event, for example- death of patients, failure of equipment or closure of businesses. The dependent variable 'Y' or the outcome is comprised of two parts,

1) Time = t

   The time from the beginning of the study till the occurrence of the desired outcome is called survival time, in this case it is the time till the closure of restaurants.

2) Event = e, (e=1 refers to closure and 0 otherwise)

   Here it refers to the closure of restaurants. But sometimes the closure may not occur during the study period and sometimes there is a lack of follow-up, and the occurrence of event remains unknown.

Thus, Y= Surv (t,e), survival outcome is a function of time that the subjects are followed for and the occurrence of an event.

Fundamentals of survival analysis

Censoring - For some observations 'e=1' within time 't', however, for some 'e=0' within the study period, which means those subjects or foodservice establishments that do not get closed during 2000 to 2020 are termed as 'censored' observations. It is necessary to include the businesses that have not failed in the study because they do have a huge impact on the survival estimates, even though we will not know their exact time of failure in the future.

Right censoring - If the failure occurs after the study period but it is not known exactly when, it is called right-censoring. It considers the fact that some businesses will still be in operation even after the study period and with some businesses the health inspection department will be unable to follow-up because of an unprecedented event like a pandemic. It is the most common form of censoring.

Left-censoring - If the failure had occurred before the observation period but it is not known when, then such observation is called left-censored. Left censoring has not been included in this study as already closed businesses will not have records during the study period in the health department data.

Interval censoring - If the closure of businesses occurs during two follow-ups within the study period, we know the occurrence of the event between two intervals but not exactly the date then this is called interval-censoring.

Survival function $S(t)= P(T>t)$ - It measures the probability of survival beyond the time t. Here, it measures the survival of foodservice restaurants beyond 2020.

Survival analysis using different models,

1) Kaplan-Meier estimator

This estimator was developed and published by Edward Kaplan and Paul Meier in June 1958 in the Journal of American Statistical Association. It is a univariate model and can only estimate survival but not the hazard ratio. It is also useful when the predictor variable is categorical. The K-M estimator was designed to measure the survival rates in patients which takes into account the ones that died and those that survived the study period. The resulting proportion of survival of patients at a given time is then plotted over a given time which finally results in a K-M survival curve. It is a stepwise declining curve.

In case of business closure, the K-M Analysis or Product-limit method can be used as,

Survival probability at time t_i,

$S(t_i) = S(t_{i-1})\ (1- d_i/n_i)$

$S(t_{i-1})$ – the probability of being in operation at t_{i-1}

n_I – number of businesses in operation just before t_i

d_i – the number of closures at t_i

Assumptions of K-M analysis

- The probability of survival is same for censored and uncensored subjects
- The likelihood of occurrence is independent to the time of enrollment in the study
- Same probability of censoring for different groups
- The event is assumed to occur at a defined time (Etikan, Bukirova, & Yuval, 2018)

2) Cox-proportional Hazard Method

It was developed and published by D.R. Cox as 'Regression models and life-tables' in the Journal of the Royal Statistical Society in 1972. It is used in medical research to determine the effect of multiple variables on the survival of patients. In general, it relates many risk factors or exposures to survival time simultaneously. It is useful for both quantitative and categorical variables. This model measures the hazard rate; in this this case it is the risk of closure of foodservice business.

If we consider a hazard function h(t), it can be estimated as,

$h(t) = h_0(t) \exp (b_1x_1 + b_2x_2 + \ldots + b_kx_k)$

Where,

h(t) – hazard function estimated by sets of k covariates,

$h_0(t)$ – baseline hazard, represents the hazard when all of the predictors are zero,

exp – exponential function

$b_1, b_2, \ldots, b_k$ – coefficients

$x_1, x_2, \ldots, x_k$ – independent variables

Assumptions

- Ratio of the hazards for any two groups is constant over time
- There is a non-linear covariate relationship
- Survival time among the individual subject is independent from one another (Simon, 2019)

CHAPTER: 4

FINDINGS AND RESULTS

4.1 Kaplan-Meier Survival Analysis

In order to overcome biases in K-M survival estimate, all the foodservice businesses that were established before the year 2000 were removed from the dataset. Only those businesses that were born and died during the study period were considered along with those that were also born within 2000-2020 but were right censored. Thus, only 8,700 observations were left for K-M analysis.

STATA 14 was then used to set up the dataset as survival data using the following code:

"stset Date, id(ClientID) origin(time Birth) fail(Event==1) scale(365.25)

Here, inspection dates were assigned as the time variable in the survival data (represented by Date). The 'id' tracked the Client ID of each business entity during the analysis period. Similarly, date of establishment was assigned to origin variable, finally, business failure was considered when the Event variable equaled to 1.

Assigning the excel dataset as survival data, STATA created four different variables _st, _d, _t, and _t0. In this case, _st was 1 if the record was going to be used and it was zero otherwise. Another variable _d was equal to 1 for failure event and was 0 if censored. Similarly, _t was analysis time when the record ended whereas, _t0 showed the analysis time when the record began.

Table 2 Examples of variables created by stset

Event	_st	_d	_origin	_t	_t0
0	1	0	15310	2614	2490
0	1	0	15310	1560	1397
1	1	1	15310	4084	3943
0	1	0	15310	1733	1560
0	1	0	15310	1397	1145

4.2 Summary Statistics

After, editing out duplicate data and probable errors, 841 different subjects had 8466 remaining observations in the final survival dataset. There were 483 failure events during the study period. The total analysis units of time that the subjects remained under risk was 1489847.

8700 total observations
157 multiple records at same instant
22 ignored because never entered
49 observations end on or before enter()
6 observations begin on or after (first) failure

8466 observations remaining, representing 841 subjects

483 failures in single-failure-per-subject data

1489847 total analysis time at risk and under observation

at risk from t = 0

earliest observed entry t = 0

last observed exit t = 6843

However, when the survival dataset was changed to annual time rather than days by using the scale of 365.25 (to convert it from days to year), the cumulative analysis time lowered to 4079 years. This was done so that the results would be cleaner, more concise and interpretable.

After scaling the dataset into years and describing it (with stdescribe command), we found that there were 841 total subjects with an average of more than 10 records per business with the total of 8466 records. The mean entry time for a business was 0 while the mean exit time was 4.9 years. The maximum exit time was in the nineteenth year. Overall, there were 483 failure events within a period of 20 years.

Table 3 Description of survival dataset

Category	Total	Mean	Min	Median	Max
No. of subjects	841				
No. of records	8466	10.06659	1	7	42
(first) Entry time		0.000469	0	0	0.073922
(final) Exit time		4.885309	0.002738	3.351129	18.73511
Subjects with gap	40				
Time on gap if gap	29.1718	0.648262	0.271047	0.553046	1.396304
Time at risk	4078.979	4.850153	0.002738	3.326489	18.72964
Failures	483	0.574316	0	1	1

The ‘stsum’ command in STATA summarized the survival data which indicated all 841 subjects came under risk during the study period which was pretty apparent as they were considered to have been at risk once they entered the inspection records. Likewise, their cumulative time of risk was 4079 years. The survival time for the twenty fifth percentile was 1.6 years, for fiftieth percentile was more than 5 years and for the seventy fifth percentile was almost 15 years.

Table 4 Summary of survival dataset

				------ Survival time -----		
	Time at risk	Incidence rate	No. of subjects	25%	50%	75%
Total	4078.97878	0.118412	841	1.623546	5.204654	14.92676

Finally, with ‘stvary’ command we identified those variables that varied over the period of 20 years. In our survival dataset, none of the variables remained constant over time. While, only one record in chain variable changed, multiple records for name, birth date, grades and segments variables changed over time. Moreover, none of the records were missing for all the variables except for the origin date which was represented by Birth.

Table 5 Description of survival variables

Variable	constant	varying	Never missing	Always missing	Sometimes missing
Name	824	17	841	0	0
Birth	803	2	80	36	725
Year	134	707	841	0	0
Pandemic	535	306	841	0	0
Grade	310	531	841	0	0
Grade2	310	531	841	0	0
Chain	840	1	841	0	0
Segment1	836	5	841	0	0
Segment2	836	5	841	0	0
Segment3	838	3	841	0	0
Segment4	833	8	841	0	0

4.3 Interpretation of Survival function

With command sts list, the survivor function for the full data was calculated. Below is a portion of the table for K-M survival analysis. It consists of eight columns. The first column represents the specific point in time (age of the establishments) and the second consists of number of businesses in operation at that particular time, for example, 833 businesses of age 0.0027 years. The third column represents the number of failures at that specific time, in this case 3 failures at age 0.0027. Likewise, net lost indicates number of censored data at a particular age- 3 at 0.0027 or 0 at 0.0055 years. The survivor function indicates the probability of survival beyond a point in time, for example the probability of survival beyond age 0.0027 is 99.6 percent. The subsequent columns represent the standard errors and confidence interval of the survival time. The table ends when the last surviving business either closes down or is censored.

The Kaplan-Meier survival function of Athens foodservice businesses from 2001 to 2020 was also plotted in a graph, given below. On the X-axis is the analysis time from year 2001 to 2020, where the age of the establishments has been represented in an increasing order from 1 to 20. And the proportion of survivor businesses is on the Y-axis. The survivor estimate curve is stepwise as the number of surviving businesses go on decreasing at different points in time and it is not gradual. At age less than 1, all of the businesses are in operation, but as they grow older their survival drops in a stepwise fashion and less than 25 percent of them remain until the age of 20.

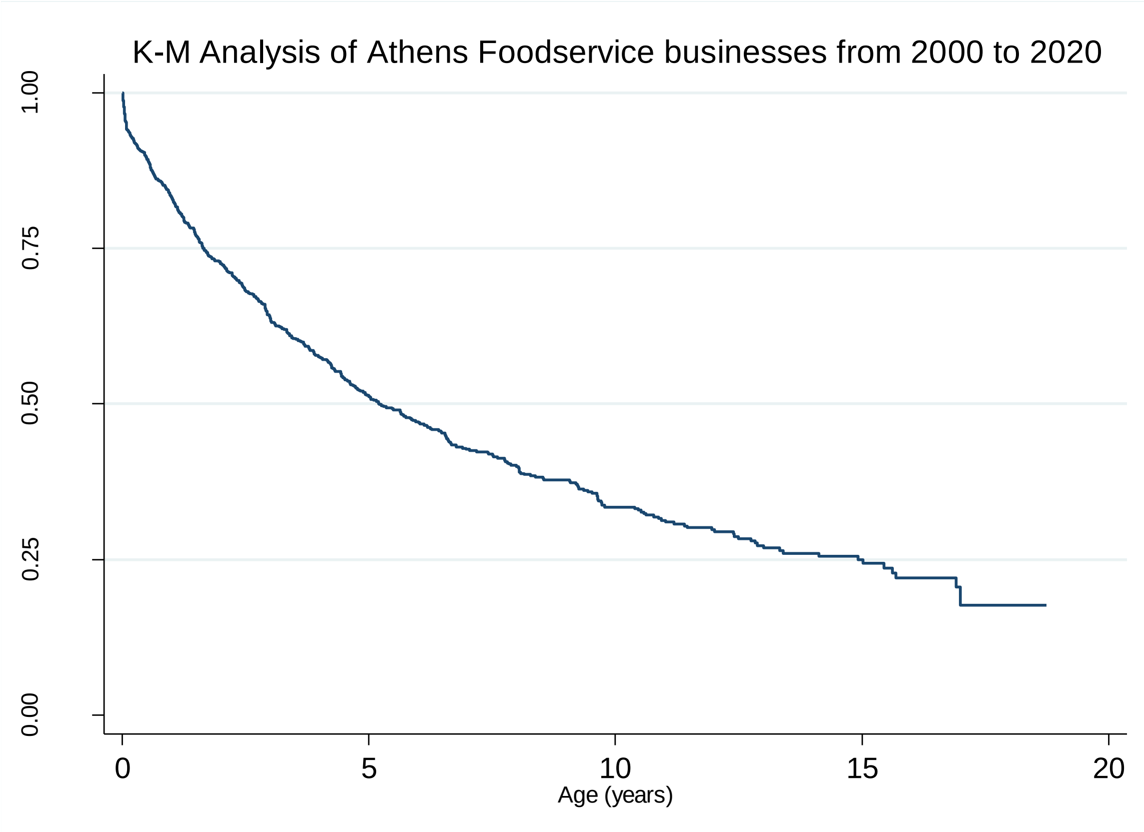

Figure 6 K-M survival curve for foodservice businesses in Athens, 2000-2020

Since, many subjects with multiple observations were studied for two decades, the survival function table became rather long and complicated. Which was why sts list, at (1(1)20) was used, it displayed the results from age 1 to 20 with an increment of 1, these were also not intervals of time rather specific points. It should be noted that survival function is calculated with full data and evaluated at indicated time, the table below is a representation of aggregates and not a true analysis.

According to this table, the probability of survival of foodservice businesses beyond age 1 year was 83 percent. Similarly, the probability of survival beyond age 10 years was about 34 percent and the probability of survival of foodservice businesses beyond age 18 years was almost 20 percent. Since, the businesses remaining after age 18 was zero, we could not asses the survival function beyond that point as there was no observed closures after that point in records.

Table 6 K-M survivor function for foodservice businesses in Athens

Time	Beg. Total	Fail	Survivor Function	Std. Error	[95% Conf. Int.]	
1	651	136	0.8302	0.0133	0.8023	0.8545
2	536	80	0.7254	0.016	0.6927	0.7553
3	442	63	0.6371	0.0175	0.6017	0.6702
4	378	42	0.5748	0.0182	0.5382	0.6096
5	317	40	0.5122	0.0187	0.4748	0.5482
6	264	25	0.4696	0.019	0.4319	0.5064
7	228	23	0.427	0.0193	0.3890	0.4644
8	192	14	0.3992	0.0194	0.3610	0.4369
9	165	10	0.3777	0.0195	0.3395	0.4158
10	134	18	0.3346	0.0198	0.2962	0.3734
11	112	8	0.3131	0.0199	0.2746	0.3524
12	89	5	0.298	0.0201	0.2592	0.3376
13	70	7	0.2725	0.0205	0.2331	0.3133
14	61	3	0.2605	0.0208	0.2207	0.3019
15	42	2	0.2498	0.0213	0.2092	0.2923
16	23	4	0.2203	0.0234	0.1763	0.2676
17	13	3	0.1763	0.0295	0.1229	0.2376
18	6	0	0.1763	0.0295	0.1229	0.2376
19	1	0	.	.	.	.
20	1	0	.	.	.	.

4.4 Testing of equality of survivor function for chain variable

H0= The survivor function is equal across Chain variable

A default log-rank test was used to test the equality of survivor function across groups. Of the failed 483 businesses, 319 belonged to independent establishments and 164 were chains. The Chi-square at 1 degree of freedom was 53.47 and the p-value was 0.0000. Thus, we reject the null of equality of survivor function across Chain variable. Which means, the survivor function of single/independent business was statistically different from that of the chains.

Table 7 Log-rank test, chain variable

Chain	Events observed	Events expected
0	319	239.01
1	164	243.99
Total	483	483

chi2(1) = 53.47

Pr>chi2 = 0.0000

4.5 K-M survival analysis based on Chain

As for the aggregate survivor function for two different kinds of Chain variable- independent and chain, a K-M survival analysis was conducted. Even though the complete analysis was done with a comprehensive data, this succinct analysis shows points in time from 1 to 20 for two groups of foodservice businesses based on their independence and association with a chain. The first column displays the point in time (age), the second represents the total businesses that are in operation at the given time. Likewise, the third column portrays the number of failure events at the given time whereas, the remaining columns give the survivor functions along with its standard error and confidence interval. Here, Chain=0 refers to independent business and Chain=1 refers to chains.

According to the table, the probability of survival of independent businesses beyond age 1 year is 75 percent whereas, of that of chains is 94 percent. Similarly, the probability of survival of independent foodservice establishments beyond age 9 years in Athens is almost 30 percent and that of chain services is nearly 50 percent. The survival rate of non-chain businesses beyond age 17 yers is 13 percent, and the survival rate of chain businesses beyond the similar point of time is nearly 25 percent (almost double). Hence, this confirms that the survival rates of chain foodservice businesses in Athens is higher than the independent businesses of the same sector.

Table 8 K-M survivor function for chain variables

Time	Total	Fail	SF	Error	(95%	Conf. Int.)
Chain=0						
1	339	115	0.7485	0.0203	0.706	0.7858
3	215	99	0.5173	0.0239	0.4694	0.563
5	151	45	0.4025	0.024	0.3554	0.4492
7	105	27	0.3246	0.0236	0.2789	0.3711
9	72	9	0.293	0.0236	0.2477	0.3397
11	51	12	0.2397	0.0238	0.1946	0.2876
13	28	5	0.209	0.0245	0.1631	0.2588
15	18	4	0.1742	0.026	0.1267	0.228
17	4	3	0.1315	0.0298	0.08	0.1959
19	1	0	.	.	.	.
Chain=1						
1	313	21	0.9386	0.013	0.9074	0.9596
3	228	44	0.7983	0.0225	0.7499	0.8384
5	167	37	0.6612	0.0277	0.6037	0.7124
7	124	21	0.568	0.0304	0.5061	0.6251
9	94	15	0.4936	0.0319	0.4296	0.5544
11	62	14	0.4136	0.0332	0.3481	0.4777
13	43	7	0.3595	0.0346	0.2922	0.4272
15	25	1	0.3508	0.0349	0.2831	0.4191
17	10	4	0.2505	0.0501	0.1592	0.3526
19	1	0	.	.	.	.

The survival analysis graph based on the Chain is given below, the X-axis represents age of the foodservices and the Y-axis displays the survivors. The red curve which belongs to chain foodservice businesses is always higher than the blue curve belonging to non-chain establishments.

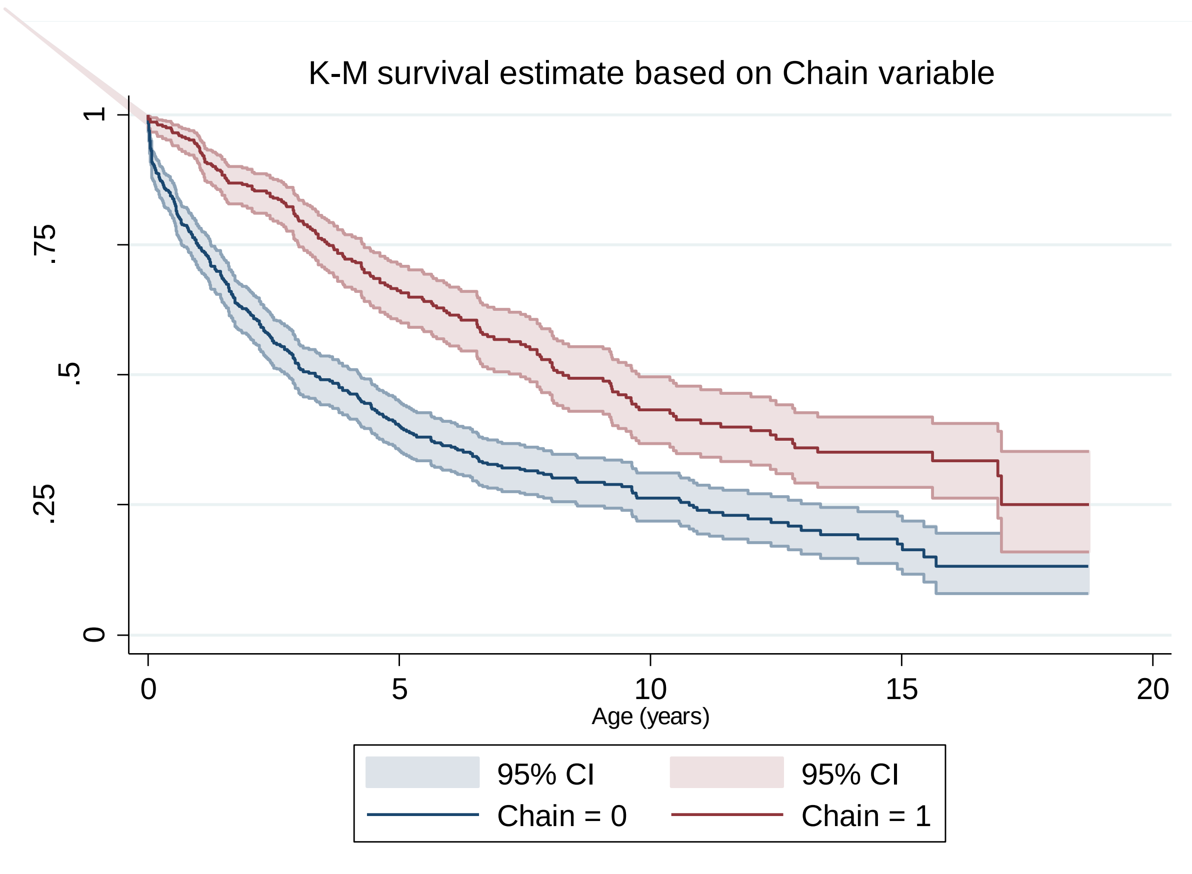

Figure 7 K-M survival curve for chain variable

4.6 Testing of equality of survivor function for Grades variable

H0= The survivor function is equal across all Grades

A default log-rank test was again used to test the equality of survivor function across groups. Of these 483 failure events, 287 were in Grade A group, 166 were in B and 30 in Grade C group. The Chi-square at 1 degree of freedom was 28.92 and the p-value was 0.0000. Thus, we reject the null of equality of survivor function across Grades variable. Which means, the survivor function of businesses who received Grade A is different from the survivor function of establishments who received Grade B or Grade C.

Table 9 Log-rank test for Grades variable

Grades	Events observed	Events expected
A	287	338.66
B	166	126.62
C	30	17.72
Total	483	483

chi2(2) = 28.92

Pr>chi2 = 0.0000

4.7 Kaplan-Meier Survival Analysis based on Grades

According to the aggregate survival function table below, the first column represents the point of time (age), the second displays number of businesses in operation at that time, the third contains the number of failed businesses and from fourth onwards it contains the survivor function, its standard error and confidence interval. Here, “A” refers to business who fall under grade A category at that point in time; the same applies to B and C.

The K-M survivor table for grades indicates that the business which receives higher grade during inspection has a higher rates of survival. For instance, the probability of survival of businesses with A grades beyond age 1 is 86 percent while the probability for establishments who received Grade B and C beyond the same point in time is 73 and 79 percent (marginally higher) respectively. As for beyond age 9, the survival function for businesses with Grades A is 43 percent and for B and C it is 28 and 19 percent respectively. Beyond age 17 the probability of survival for business with the highest grade slump to 25 percent, business with grades B have only 9 percent chances of survival; as for restaurant businesses with grades C, none of them survive beyond time 15.

The K-M survival graph for foodservice businesses in Athens based on their inspection grades is given below. The X-axis represents the time of analysis whereas the Y-axis represents the proportion of survivors. The blue curve which is always higher is the survival curve for businesses with grade A followed by Grade B and C that are represented by red and green curves respectively.

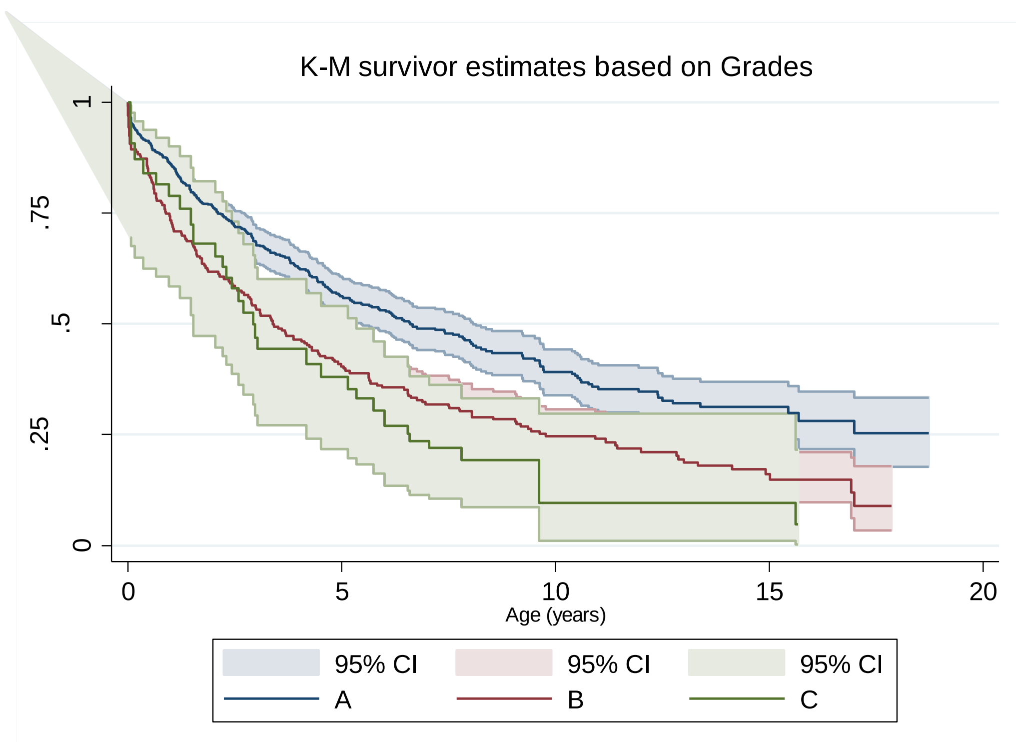

Figure 8 K-M survival curve for Grades variable

The flat line of businesses with Grade C from analysis time 10 to 15 indicates no closure between those two specific points in time. This graph is much clearer for Grade B and C recipients without the confidence interval, as it is given below. It indicates that the businesses with Grade B had a higher probability of survival than those with Grade C when they are younger than 5 years.

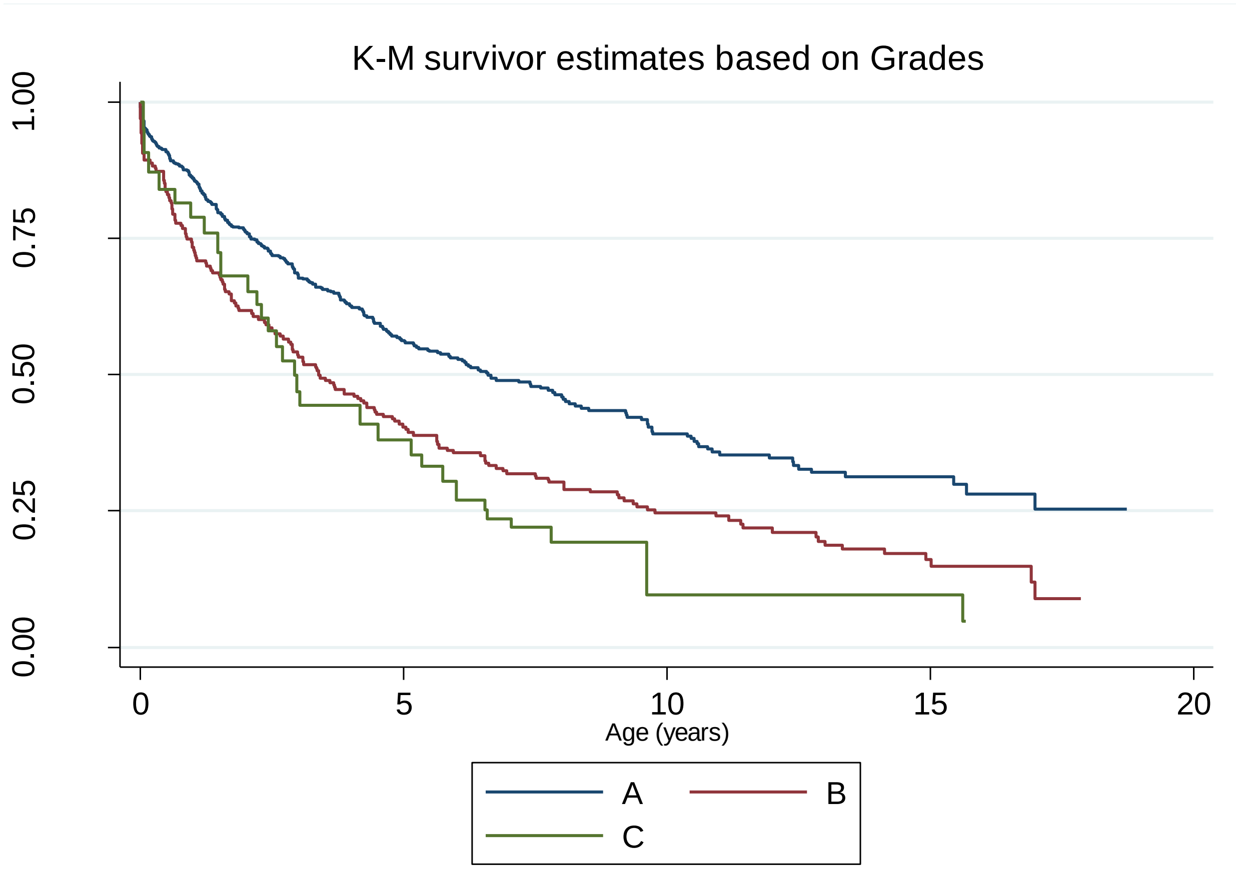

Figure 9 K-M survival curve for Grades variable without confidence interval

4.8 Testing of equality of survivor function for Limited-service restaurants

H0= The survivor function is equal across limited-service restaurants

Another log-rank test was used to test the equality of survivor function across groups. Of the total 483 failure events, 166 belonged to segment 1 and 317 belonged to others. The Chi-square at 1 degree of freedom was 9.47 and the p-value was 0.0021. Thus, we reject the null of equality of survivor function across limited-service restaurants at 5 percent level of significance. Which means, the survivor function of Limited-service restaurants is different from the survivor function of establishments which belonged to other segments.

Table 10 Log-rank test for Segment 1

Segment1	Events observed	Events expected
0	317	283.72
1	166	199.28
Total	483	483

chi2(1) = 9.47

Pr>chi2 = 0.0021

4.9 Kaplan-Meier Survival analysis on Limited-service restaurants

The following aggregate survival analysis table consists of several columns- for age of businesses, the total proportion of survivors at that age, failure events, and survival function with standard error and confidence interval. Here, Segment 1=1 refers to limited-service restaurants and 0 otherwise.

It can be inferred from the survivor table that the probability of survival of limited service restaurants is higher than other segments. After 1 year of age, the probability of survival for limited-service restaurant is almost 89 percent while, it is about 80 percent for other segments. Likewise, after 9 years of age, the probability of survival of limited-service restaurant is 42 percent and that of other segments is 35 percent. Finally, the survival function for limited-service restaurants beyond age 17 is 23 percent while the survival function for other segments of foodservice businesses beyond the same point in time is almost 15 percent.

This is well represented in the graph given below. Here, the limited-service restaurant segment is indicated by the red curve which is always higher than the blue survival curve that represents other business segments.

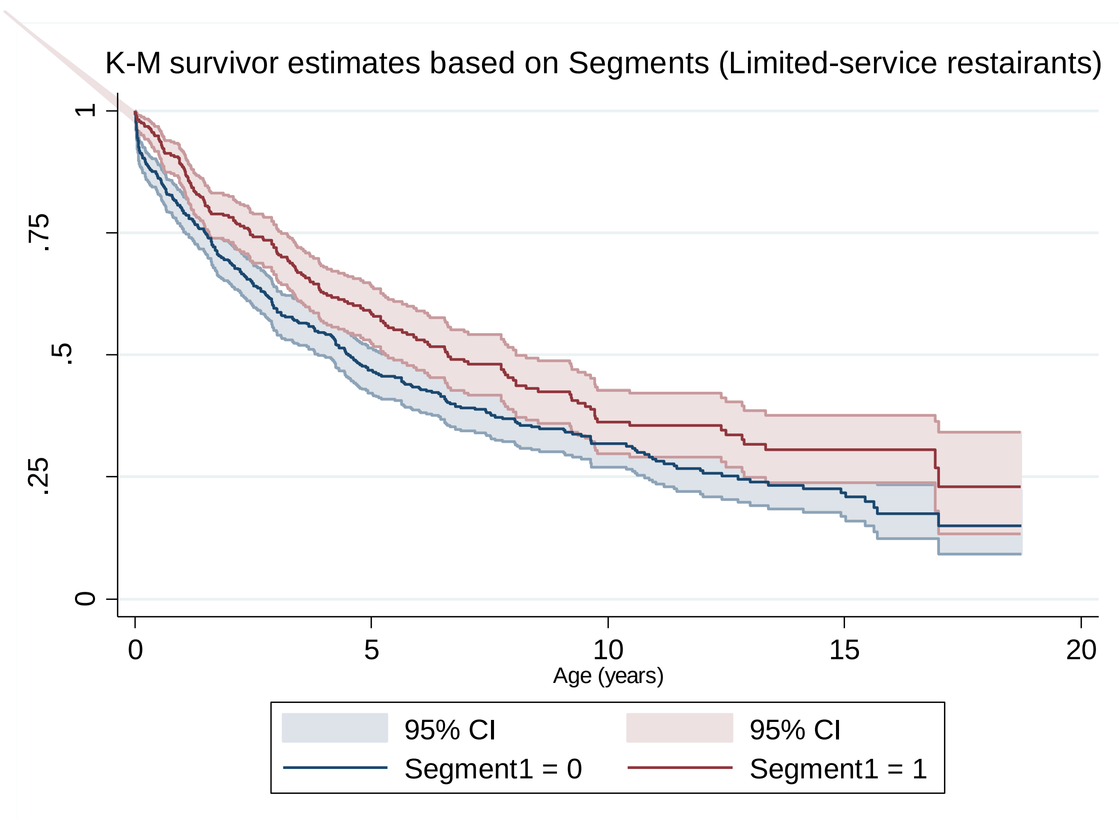

Figure 10 K-M survival curve for Segment 1

4.10 Testing of equality of survivor function for bars/pubs/taverns

H0= The survivor function is equal across bars/pubs/taverns

Again log-rank test was used to test the equality of survivor function across groups. For the concerned group, i.e. Segment 2 there were a total of 483 failure events. Of these failed establishments, 79 were bars, pubs, and taverns and 404 were other segments. The Chi-square at 1 degree of freedom was 1.82 and the p-value was 0.1778. Thus, we fail to reject the null of equality of survivor function across bars, pubs and taverns. Which means, the survivor function of bars, pubs and taverns is statistically same with other segments of foodservice sectors.

Thus, there is no interpretation of survivor function for bars, pubs and taverns.

Table 11 Log-rank test for Segment 2

Segment2	Events observed	Events expected
0	404	414.33
1	79	68.67
Total	483	483

chi2(1) = 1.82

Pr>chi2 = 0.1778

4.11 Testing of equality of survivor function for 'other' segment

H0= The survivor function is equal across 'other' segment

A default log-rank test was used to test the equality of survivor function across groups. For the concerned group, i.e. Segment 3 there were a total of 483 failure events. Of these failed establishments, 66 were hotels, bakeries, ice-cream parlors, recreational centers, theatres and movie halls and 417 were other segments. The Chi-square at 1 degree of freedom was 2.24 and the p-value was 0.1346. Thus, we fail to reject the null of equality of survivor function across Segment 3 variable. Which means, the survivor function of "other" forms of foodservice establishments are statistically similar to other segments.

Thus, there is no interpretation of survivor function for other forms of foodservices.

Table 12 Log-rank test for Segment 3

Segment3	Events observed	Events expected
0	417	427.48
1	66	55.52
Total	483	483

chi2(1) = 2.24

Pr>chi2 = 0.1346

4.12 Cox's proportional hazard Regression

Table 13 illustrates a simple Cox's proportional regression result with three covariates- chain, grades and segments. The Chi-square value of the regression function is 81.94 and the p-value is 0.000 which means the model is significant. For this multivariate Cox model, chain and grades variables are significant with p-value<0.05, however, the segments variable fails to be significant with p-value=0.727.

Table 13 Cox-regression table with coefficients

_t	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
Chain	-0.7117	0.10171	-7	0	-0.911	-0.5123
Grades	0.40278	0.0735	5.48	0	0.25871	0.54684
Segments	-0.0129	0.03703	-0.35	0.727	-0.0855	0.05967

LR chi2(3) = 81.94

Log likelihood = -2862.9805, Prob > chi2 = 0.0000

Hazard-ratio

With, p=0, and HR= exp(β) = 0.71, chain variable indicates a strong relationship with decreased risk of closure. It means holding, other covariates constant, having an affiliation with chain (chain=1) decreases the likelihood of closure by a factor of 0.49 or 51 percent.

Also, with p-value= 0 and HR= exp(β)= 1.5, Grades variable displays a strong relationship with increased likelihood of closure. Here, greater value of grade increases the likelihood of closure of foodservice business by a factor of 1.5. This finding seems counter-intuitive at first, however, the grades here have been coded in such a way that (A=1, B=2 and C=3) which means having higher numerical value (2>1) refers to lower Grade (B<A) which increases the likelihood of closure by 50 percent.

As for the Segments variable, it has a p-value of 0.727 and HR= exp(β)= 0.98 with confidence interval between 0.92 to 1.06. Since, the confidence interval of Segments variable includes 1, it does not have a significant contribution towards the closure of foodservice businesses in Athens.

Table 14 Cox regression table with hazard ratio

_t	Haz. Ratio	Std. Err.	z	P>z	[95% Conf.	Interval]
Chain	0.490829	0.04992	-7	0	0.4021228	0.599103
Grades	1.495971	0.10996	5.48	0	1.295258	1.727786
Segments	0.9871762	0.036555	-0.35	0.727	0.9180686	1.061486

4.13 Test of proportional-hazard assumption

H0= The hazard ratio between any two groups is not proportional

Schoenfeld residual test is carried out to examine the proportionality assumption of hazard-ratio. Here, the Chi-square value is 13 and the degree of freedom is 3, however, the p-value is 0.0046, which means we fail to reject the null-hypothesis. So, our survival dataset could not fulfill the proportional hazard assumption, this implies that our inference holds low power (Simon, 2019).

Table 15 Schoenfeld test for proportional-hazard assumption

	chi2	df	Prob>chi2
global test	13	3	0.0046

CHAPTER:5

DISCUSSIONS

A K-M analysis on 8700 observations of Athens foodservice establishments from year 2000 to 2020 was conducted. Of those total observations belonging to businesses born and died and censored between 2000 and 2020, 8466 remained after removing duplicates and errors. Those observations related to 841 total subjects of which 483 experienced a failure event, in this case-closure. The cumulative time under risk was approximately 4079 years as establishments were considered to have been at risk once they entered the inspection records.

5.1 Survival function for Athens foodservice establishments

For food service businesses in Athens, the probability of survival beyond age 1 was 83 percent, this is in line with the findings of (Luo & Stark, Only the bad die young: Restaurant Mortality in the Western US, 2014) that in the first year 17 percent of the independently owned full-service restaurants failed in the US (which means 83 percent survived). Similarly, the survival rate beyond the age 10 was 34 percent and beyond age 18 was nearly 20 percent. This high failure rate during the initial years of establishments may be because restaurants are highly leveraged and are vulnerable to economic turmoil. Still, they have a higher rate of survival than other forms businesses (Luo & Stark, Only the bad die young: Restaurant Mortality in the Western US, 2014) which makes them a better investment and livelihood option than other businesses.

5.2 Survival function and chains

According to the log-rank tests and K-M analysis, the survival of chain foodservice was higher than that of independently owned businesses. Similar results were reported by (Parsa, Self, Njite, & King, 2005) after studying the restaurant ownership turnover data in Columbus, Ohio from 1996 through 1999 found that the rate of failure of independent restaurants was higher than that of franchise chains. They also reported a 26.16 percent failure of independently owned restaurant which is similar to that of the Athens foodservice businesses where in their first year only 75 percent of the independently owned businesses survived which translates to 25 percent failure rate. This finding is also supported by (English, 1996) who conducted a five-year longitudinal study on restaurant attrition in El-Paso Texas and concluded that franchise and corporate chain businesses had higher investments and higher success rate as compared to independently owned restaurants. This may be due to the transfer of knowledge, skills and trademarked information to the owners from the franchises and chains which makes them less susceptible to failures. Similarly, independent businesses have smaller number of staffs as well as food choices which also lowers their chances of success (Parsa, Self, Njite, & King, 2005).

5.3 Survival function and health inspection grades

Businesses who receive higher grades in Athens have better rates of survival than those who receive lower grades. When we compare the log-rank test, foodservice businesses have greater chances of survival if they receive A grade rather than B or C. Surprisingly, before third year, the survival rates for businesses with B grade is similar to those who receive C grade, later on there is a marginal improvement in the survival rate of the former as compared to the latter. Even though there has been no empirical study that has measured the impact of restaurant inspection scores om their survival rate. But a 2019 study conducted by (Parsa, Kreeger, Rest, Xie, & Lamb, 2019) confirms that in Denver, Colorado, the increase in number of health code violations by a restaurant increases its likelihood of failure irrespective of its density locations.

There has been a research that assesses the relationship between QSC (Quality, Service and Cleanliness Inspections) and financial performance those institutions. A study by (DiPietro, Parsa, & Gregory, 2011) discovered that the relationship between QSC variable and restaurant performance is weak. This is not enough to determine the closure of restaurants. However, the closure might be related to public perception of inspection scores and their reluctance to visit restaurants with low health inspection scores. A survey conducted by (Jones & Grimm, 2008) on 2000 respondents concluded that 77 percent of participants were unwilling to visit restaurants with score less than 80 and 45 percent said they prefer restaurants with scores above 90. Of the total respondents, 37 percent also opined that the low scoring establishments should be closed immediately after health violations.

5.4 Survival functions and segments

Another interesting finding from this study is that only the survival rate of limited-service restaurant is statistically different and higher to other all segments like full-service restaurants, bars, taverns and pubs and theatres, recreational centers, etc. According to the K-M survival analysis and log-rank test, beyond age 1 the probability of survival of limited-service restaurants is 89 percent.

This could be because most of the limited-service restaurants are chains and since chains have greater longevity at a given point in time, naturally this segment of foodservice institutions will have greater probability of survival. However, further interaction of chains and segments of restaurants must be carried out to study their impact on the survival rates. As, K-M analysis is non-parametric, it is limited to single variable at a time and cannot provide the necessary information.

5.5 Cox's P-H ratio and closure of foodservice businesses in Athens

A multivariate Cox-regression indicates a strong relationship between Chain and Grades variable with the closure of foodservice businesses in Athens. As per the analysis, affiliation with a chain decreases the likelihood of closure by 51 percent and receiving a lower grade (from A to B) increases the closure rate by 50 percent. These findings are in conjunction with findings of (Parsa H. G., Self, Sydnor-Busso, & Yoon, 2011) and (Parsa, Kreeger, Rest, Xie, & Lamb, 2019). As for Segments, it is not a significant contributor to the closures of foodservice businesses in Athens, which means being a limited-service restaurant has an equal probability of closure than being a full-service restaurant.

5.5 Implications

This study recommends investors especially those who are willing to invest on a sustainable restaurant business, they should focus on gaining a franchise license or opening branches of foodservices in multiple locations than running an independently owned business. Also, for restaurateurs, they should target for greater inspection scores as higher inspection grades have been directly linked with longevity of foodservice businesses. Although, multiple regression analysis proposed no significant relationship between any segments of restaurant and their closures, other literatures have found that limited-service restaurants have greater probability of survival than other segments during recessions and recoveries. So, establishing and maintaining a quick-service and fast casual foodservice business is beneficial in the long run rather than full-service restaurants, bars, taverns and other forms of businesses especially during economic crisis.

5.6 Future Research

Due to time-constraints and limitation of data, this study could not incorporate several research possibilities such as, Athens being a college town and comparing it with another town without a university as a control for the effect of UGA and its proximity on business closures. Likewise, the survival rates of foodservice businesses with other forms of businesses like retail could also be compared using similar methods. There is also an interesting prospect of including the street addresses of foodservice or any other businesses and finding their closure rates in specific locations of county. Improvement upon the methodology like satiation of proportional-hazard assumption or utilizing Restricted Mean Square Test (RMST) could also strengthen the power of inference. These are some of the possible avenues in survival analysis and foodservice business closures that can be explored in the future.

References

A.W.Bartik., M.Bertrand., Z.B.Cullen., E.L.Glaeser., M.Luca., & C.T.Stanton. (2020). HOW ARE SMALL BUSINESSES ADJUSTING TO COVID-19? *EARLY EVIDENCE FROM A SURVEY*.

Aguilar, M. I., & Diaz, B. (2019). Length of stay of international tourists in Spain: A parametric survival analysis. *Annals of Tourism Research*.

Athens Banner Herald. (2020, 09 13). A look at downtown Athens businesses: Who reopened, who remains closed during COVID-19. *Athens Banner Herald*.

Athens Banner Herald. (2020, 05 18). Georgia hardest hit state with COVID-19 job losses. *Athens Banner Herald*.

Athens-Clarke County Unified Government. (2021). *Athens-Clarke county by numbers*. Retrieved from Athens-Clarke County Unified Government: https://www.accgov.com/DocumentCenter/View/535/Athens-Clarke-County-By-the-Numbers

Bartik, A., Bertrand, M., Cullen, Z., Glaeser, E., Luca, M., & Stanton, C. (2020). How Are Small Businesses Adjusting to COVID-19? Early Evidence from a Survey. *National Bureau of Economic Research*.

Beheler, R., Norton, S., & Sen, K. (2008). A Comparison of Company Owned and Franchised Fast Food Outlet Performance: Insights from Health Inspection Scores. *Strategy and Governance of Networks*, 113-125.

Bertagnoli, L. (2005). How long can they go? *Restaurants and Institutions,* 115(3):57-59.

Bureau of Labor Statistics. (2020). *Labor Force Statistics from the Current Population Survey.* Washington, DC: Bureau of Labor Statistics.

Cameron, A., & Trivedi, P. (2005). *Microeconometrics Method and Application.* New York: Cambridge University Press.

CDC. (2020, January 21). *First Travel-related Case of 2019 Novel Coronavirus Detected in United States*. Retrieved from Centers for Disease Control and Prevention: https://www.cdc.gov/media/releases/2020/p0121-novel-coronavirus-travel-case.html

CDC. (2020, 03 15). *Guidance for COVID-19*. Retrieved from Centers for Disease Control and Preventions: https://www.cdc.gov/coronavirus/2019-ncov/communication/guidance.html

CDC. (2021, 05 24). *COVID Data Tracker*. Retrieved from Centers for Disease Control and Prevention: https://covid.cdc.gov/covid-data-tracker/#datatracker-home

CHRON. (2020, 11 20). *Difference Between a Chain and Franchised Restaurant*. Retrieved from CHRON: https://smallbusiness.chron.com/difference-between-chain-franchised-restaurant-17288.html

CNN. (2020, 03 18). What is and isn't allowed during a 'shelter-in-place' order. *CNN*.

CNN. (2020, 07 25). Yikes! Yelp says 60% of restaurant Covid-19 closures are permanent. *CNN Business*.

D.Gursory., & C.G.Chi. (2020). Effects of COVID-19 pandemic on hospitality industry: review of the current situations and a research agenda. *Journal of Hospitality Marketing & Management*.

DiPietro, R., Parsa, H., & Gregory, A. (2011). Restaurant QSC inspections and financial performance: an emperical investigation. *International Journal of Contemporary Hospitality Management*.

English, W. (1996). Restaurant attrition: a longitudinal analysis of restaurant failures. *International Journal of Contemporary Hospitality Management*, 17-20.

Etikan, I., Bukirova, K., & Yuval, M. (2018). Choosing statistical tests for survival analysis. *Biometrics & Biostatistics International Journal*, 477-481.

FOX Business. (2020, 05 08). What is a non-essential business? *FOX Business*.

Gemar, G., Soler, I. P., & Guzman-Parra, V. F. (2019). Predicting bankruptcy in resort hotels: a survival analysis. *International Journal of Contemporary Hospitality Management*, 1546-1566.

georgia.gov. (2021, 05 25). *2020 Executive Orders*. Retrieved from Governer Brian P. Kemp, Office of the Governer: https://gov.georgia.gov/executive-action/executive-orders/2020-executive-orders

Gokovali, U., Bahar, O., & Kozak, M. (2007). determinants of length of stay: A practical use of survival analysis. *Tourism Management*, 736-746.

Grady Newsource. (2020, 03 16). Some Local Athens Restaurants Face Extinction During Pandemic. *Grady Newsource*.

Gursoy, D., & Chi, C. G. (2020). Effects of COVID-19 pandemic on hospitality industry: review of the current situations and a research agenda. *Journal of Hospitality Marketing and Management*, 527-529.

Gutie´rrez, A., Miravet, D., Saladie, O., & Clave, S. A. (2020). High-speed rail, tourists' destination choice and length of stay: A survival model analysis. *Tourism Economics*, 578-597.

Hannan, B. &. (1989). *Organizational Ecology.* Cambridge: Harvard University Press.

Hannan, M. &. (1977). The population ecology of organizations. *American Journal of Sociology,* 82(5), 929-964.

Harris, K., DiPietro, R., Murphy, K., & Rivera, G. (2014). Critical food safety violations in Florida: Relationship to location and chain vs. non-chain restaurants. *International Journal of Hospitality Management*, 57-64.

Hayes, A. (2021, 02 28). *Business Essentials*. Retrieved from Investopedia: https://www.investopedia.com/terms/f/franchise.asp

IBISWorld. (2021, 04). *US Industry (NAICS)*. Retrieved from IBISWorld: https://my.ibisworld.com/search/?q=report%20on%20single%20location%20full-service%20restaurants%20in%20the%20U.S

J.Salazar., J.C.Bergstrom., & D.Poudel. (2020). Estimating the Economic Impacts of the COVID19 Pandemic on the State of Georgia Economy_A Quick-Response Method. *Research in AGricultural and Applied Economics*.

Jones, T., & Grimm, K. (2008). Public knowledge and attitudes regarding Public Health Inspections of Restaurants. *American Journal of Preventive Medicine*, 510-513.

Jones, T., Pavlin, B., LaFleur, B., Ingram, L., & Schaffner, W. (2004). Restaurant Inspection Scores and Foodborne Disease. *Emerging infectious diseases*, 688-692.

Katai, H., Ishikawa, T., Akazawa, K., Isobe, Y., Miyashiro, I., Oda, I., . . . al., e. (2018). Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001–2007). *Gastric Cancer*, 144-154.

Kim, W., Zhong, J., Chen, M., & Karadag, E. (2009). Risk-Adjusted Performance of Three Restaurant Segments in the USA. *Torusim Economics*, 139-152.

Lado-Sestayo, R., Vivel-Bua, M., & Otero-Gonzalez, L. (2016). Survival in the lodging sector: An analysis at the firm and location levels. *International Journal of Hospitality Management*, 19-30.

lee, J., Almanza, B., Nelson, D., & Ghiselli, R. F. (2009). Using Health Inspection Scores to Assess Risk in Food Services. *Journal of Environmental Health*, 29-35.

Lin, H. H., Tsai, C. W., Lin, P. H., Cheng, K. F., Wu, D. W., Wang, I. K., & al., e. (2012). Survival analysis of pediatric dialysis patients in Taiwan. *Nephrology*, 621-627.

Luo, T., & Stark, B. (2014). Only the bad die young: Restaurant Mortality in the Western US. *Cornell University* .

Luo, T., & Stark, P. (2014). *Only the bad die young: restaurant mortality rate in the US.*

Mathias, R., Sizto, R., Hazlewood, A., & Cockedge, W. (1995). The effects of inspection frequency and food handler education on restaurant inspection violations. *Canadian Journal of Public Health*, 46-50.

Miller, R. J., Gong, G., & Munoz, A. (1981). Survival Analysis. In R. J. Miller, *Survival Analysis* (pp. 1-10). New York: John Wiley & Sons.

NRA. (2019). *2019 National Restaurant Industry Factbook.* Washington, DC: National Restaurant Association.

NRA. (2021). *State of the Restaurant Industry, 2021.* Washington, DC: National Restaurant Association.

Parsa, H. G., Self, J., Sydnor-Busso, S., & Yoon, H. J. (2011). Why Restaurants Fail? Part II - The Impact of Affiliation, Location, and Size on Restaurant Failures: Results from a Survival Analysis. *Journal of Foodservice Business Research*, 360-379.

Parsa, H., Kreeger, J., Rest, J.-P., Xie, L., & Lamb, J. (2019). Why restaurants Fail? Part V: Role of Economic Factors, Risk, Density, Location, Cuisine, Health Code Violations and GIS Factors. *International Journal of Hospitality and Tourism Administration*, 142-167.

Parsa, H., Self, J., Njite, D., & King, T. (2005). Why restaurants Fail. *Cornell Hotel and Restaurant Asministration Quarterly*.

Song, H., Yeon, J., & Lee, S. (2021). Impact of the COVID-19 pandemic: Evidence from the U.S. restaurant industry. *International Journal of Hospitality Management*.

Statista. (2019). *Travel, Tourism and Hospitality*. Retrieved from Statista.com: https://www.statista.com/statistics/374866/number-of-restaurants-by-type-us/

Statista. (2021, May 23). *Monthly sales of foodservice and drinking places in the United States from January 2018 to January 2021*. Retrieved from Statista: https://www.statista.com/statistics/894051/monthly-sales-of-us-eating-and-drinking-places/

The Red & Black. (2020, 12 29). *Economic swings continue for bars and restaurants in Athens as UGA's fall football season ends*. Retrieved from The Red & Black: https://www.redandblack.com/athensnews/economic-swings-continue-for-bars-and-restaurants-in-athens-as-uga-s-fall-football-season/article_89f629be-49a7-11eb-ba9f-fb27f83bddfc.html

US Census Bureau. (2021, 5 14). *Estimates of Monthly Retail and Food Services Sales by Kind of Businesses:2021.* Retrieved from United States Census Bureau: https://www.census.gov/retail/mrts/www/mrtssales92-present.xls

USDA. (2020). *Ag and Food Sectors and the Economy.* U.S. Department of Agriculture.

Young, J., Clark, P., & McIntyre, F. (2008). An Exploratory Comparison of the Casual Dining Experience. *Journal of Foodservice Business Research*, 87-105.

Zheng, T., Farrish, J., & Wang, X. (2012). How Did Different Restaurant Segments Perform Differently through the Recession? An ARIMA with Intervention Analysis on US Restaurant Stock Indices. *Journal of Hospitality Financial Management*.

Appendix A

Survivor Function by Grades

	Time	Beg. Total	Fail	Survivor Function	Std. Error	[95% Conf. Int.]	
A							
	1	474	83	0.8606	0.0143	0.8300	0.8861
	3	306	89	0.6838	0.0202	0.6422	0.7216
	5	221	49	0.5627	0.0229	0.5165	0.6063
	7	148	25	0.4893	0.0243	0.4410	0.5358
	9	110	14	0.4342	0.0256	0.3836	0.4838
	11	71	17	0.358	0.027	0.3053	0.4109
	13	41	6	0.3201	0.0283	0.2655	0.3760
	15	26	1	0.3123	0.0287	0.2571	0.3690
	17	10	3	0.2525	0.0402	0.1779	0.3339
	19	1	0	.	.	.	.
B							
	1	146	47	0.7331	0.0333	0.6613	0.7920
	3	120	43	0.5322	0.0357	0.4599	0.5992
	5	79	30	0.4039	0.034	0.3371	0.4696
	7	64	17	0.3181	0.0325	0.2556	0.3823
	9	50	8	0.2845	0.0312	0.2251	0.3467
	11	38	8	0.2402	0.03	0.1839	0.3009
	13	27	6	0.1944	0.0296	0.1401	0.2554
	15	14	4	0.1606	0.0292	0.1084	0.2219
	17	4	3	0.0895	0.0371	0.0341	0.1783
	19	1	0	.	.	.	.
C							
	1	31	6	0.7878	0.078	0.5838	0.8997
	3	16	11	0.469	0.0878	0.2926	0.6272
	5	16	3	0.3796	0.085	0.2178	0.5404
	7	16	6	0.2355	0.0708	0.1138	0.3821
	9	3	2	0.1932	0.0648	0.0861	0.3320
	11	1	1	0.0966	0.0756	0.0111	0.2974
	13	0	0	0.0966	0.0756	0.0111	0.2974
	15	0	0	0.0966	0.0756	0.0111	0.2974
	17	1	1	.	.	.	.
	19	1	0	.	.	.	.

Survivor function by Segments (Limited-Service restaurants)

Time	Beg. Total	Fail	Survivor Function	Std. Error	[95% Conf. Int.]	
Segment1=0						
1	385	101	0.7958	0.0181	0.7575	0.8288
3	257	93	0.5913	0.0228	0.5453	0.6343
5	184	50	0.4688	0.0238	0.4216	0.5146
7	132	28	0.3909	0.024	0.3438	0.4376
9	95	13	0.3487	0.0241	0.3018	0.3959
11	64	15	0.2865	0.0246	0.2392	0.3354
13	40	8	0.2452	0.0251	0.1977	0.2956
15	27	4	0.2173	0.0259	0.1688	0.2699
17	7	5	0.1502	0.0336	0.0916	0.2222
19	1	0	.	.	.	.
Segment1=1						
1	267	35	0.8855	0.0182	0.8441	0.9164
3	186	50	0.7113	0.0266	0.6555	0.7598
5	134	32	0.5828	0.03	0.5217	0.6390
7	97	20	0.4857	0.0319	0.4218	0.5466
9	71	11	0.4246	0.0328	0.3597	0.4878
11	49	11	0.3556	0.0335	0.2906	0.4211
13	31	4	0.3169	0.035	0.2497	0.3860
15	16	1	0.3063	0.0354	0.2387	0.3764
17	7	2	0.2297	0.0539	0.1339	0.3409
19	1	0	.	.	.	.

Survivor function before and during the pandemic

Time	Beg. Total	Fail	Survivor Function	Std. Error	[95% Conf. Int.]	
Pandemic=0						
1	596	130	0.8249	0.014	0.7956	0.8505
3	408	136	0.6267	0.0183	0.5898	0.6613
5	273	79	0.4932	0.0196	0.4541	0.5310
7	190	46	0.4036	0.02	0.3642	0.4427
9	142	23	0.351	0.0202	0.3115	0.3906
11	84	23	0.2862	0.0206	0.2466	0.3270
13	46	9	0.2468	0.0216	0.2056	0.2901
15	22	4	0.2183	0.0236	0.1739	0.2661
17	1	5	.	.	.	.
19	1	0	.	.	.	.
Pandemic=1						
1	56	6	0.9008	0.0386	0.7919	0.9543
3	34	7	0.7681	0.0573	0.6319	0.8593
5	43	3	0.7139	0.0612	0.5741	0.8150
7	37	2	0.67	0.0649	0.5255	0.7793
9	22	1	0.649	0.0662	0.5035	0.7616
11	27	3	0.5763	0.0709	0.4259	0.7006
13	25	3	0.5093	0.0725	0.3604	0.6401
15	19	1	0.4881	0.0726	0.3407	0.6201
17	13	2	0.4268	0.0757	0.2779	0.5677
19	1	0	.	.	.	.