The Effect of Public Inspection Information Disclosure for Consumers on Restaurants' Hygiene Quality

Shuiqin Yu
Colorado State University
shuiqin.yu@colostate.edu
Marco Costanigro
Colorado State University
marco.costanigro@colostate.edu

Selected Paper prepared for presentation at the 2019 Agricultural & Applied Economics Association Annual Meeting, Atlanta, GA, July 21 – July 23

Introduction

As technology develops, so does the way consumers search for food away from home. More and more consumers are using online search website and reading other consumers' reviews to determine the final choice of restaurant. Yelp.com is one of the most popular website consumers relying on restaurants search and choices. Despite the recent stagnant growth of desktop users, Yelp captures a rapid growth of mobile users, making it the top one website for diners.

Yelp began to create the Local Inspector Value-Entry Specifications (LIVES) program with San Francisco and New York City in 2012 (Yelp, 2018). This LIVES program enables municipalities to disclose restaurant inspection information on Yelp. As of September of 2018, 30 districts were partnering with Yelp to publish restaurant inspection information on the Yelp platform. This paper addresses the economic question of whether restaurants respond to this form of online disclosure regulation and how economically significant these changes are in restaurant hygiene.

A growing body of food safety policies emerged in the wake of increasing public concerns over the foodborne illness outbreaks due to compromised ingredients, unsanitary environments, and improper food handling in the United States in recent years. One such policy is the legislated Food Safety Modernization Act (FSMA) in 2011, which provides a systematic approach to food safety regulations (Hoffman, 2011). Policies aimed at reducing these foodborne illness outbreaks through increasing product quality information to consumers have been gradually implemented across the nation at different municipality levels. For example, Los Angeles County has mandated restaurants to display hygiene quality grade cards on windows since 1998 (Jin and Leslie, 2003). This study examines the impact of public disclosure of restaurant inspection information or LIVES program on restaurant hygiene quality. Using a difference-in-difference approach and a regression discontinuity design, this study finds evidence of positive and significant restaurant hygiene quality improvements as a result of increased access to information for consumers.

I will present the development of the LIVES program in the background section. Previous theoretical and empirical studies on the impact of information on product quality are discussed in the literature review section. In the conceptual model section, I adopt a theoretical framework to model the dynamic process of restaurant quality choice when consumers have increased access to hygiene information. I describe the data in use in the data section and review the empirical specification and results in the empirical specification section. The last section concludes and summarizes the main findings.

Background

The regulatory framework for the oversight of restaurants includes the Food and Drug Administration (FDA) and state/territorial regulatory agencies. The FDA developed the Food Code, a model that aims at protecting public health and guaranteeing safe and uncontaminated food for consumers (FDA, 2017). While the adoption of the Food Code is encouraged, but not mandatory, most states adopted it except California, New York, and Vermont as of 12/31/2017 (FDA, 2017). Despite the disparities in state or territorial regulations, restaurants in the U.S. are generally subject to inspections at least once a year. Typically, inspection of a restaurant or retail food service could be a routine inspection, a follow-up inspection, or a complaint-initiated inspection. Other types of inspections vary within each state/territory. The regulations or codes that are applied to inspect restaurants vary from state to state, and in some cases, from city to city. For example, cities in North Carolina share the same inspection criteria for food and facility inspections, while counties in Colorado use different criteria to inspect restaurants.

There are three main restaurant grading systems: the points-deduction system, the letter grade system, and the violation accumulation system. The points-deduction system scores start with 100 points. Points are deducted based on the critical levels and categories of the violations. The letter grade system uses letter grades A, B, or C to signal the sanitary conditions and food handling practices from Excellent (A), Good (B) or Fail (C).

In some cities, the inspector will issue a grade card or a score card to the restaurant manager at the end of the inspection. These cards are required to be displayed so that consumers know the result of the most recent inspection. The specific requirements on the placement of the card varies from city to city. Some cities require the card to be displayed on a window while other cities require the card to be posted near the public entrance. The purpose of these displaying requirements is to increase consumers' awareness about the hygiene of the restaurants. In addition, hygiene reports for restaurants in some cities can often be found in local newspapers or broadcasted by local radio stations. In many cities, restaurant inspection reports are posted on the state or local municipality websites.

Starting from 2012, Yelp began to work with the City of San Francisco and City of New York to launch the LIVES program (Yelp, 2018). Restaurant inspection results can be retrieved on the LIVES page on Yelp. This program allows government agencies that carry out restaurant inspections to share their inspection data with Yelp's mobile and desktop users. By clicking on the Yelp website or application, consumers find the inspection data at a restaurant's yelp page along with other operating information. They learn whether a restaurant passed its previous inspection, what the violations were, and whether they were critical or non-critical (in other cases, high-risk or low-risk). Disclosing the restaurant inspection information on Yelp reduces the information asymmetry about restaurant hygiene quality between consumers and restaurants. Therefore, it increases the saliency of hygiene quality in the market.

By September of 2018, 30 municipalities in the U.S. have partnered with Yelp to bring inspection information to consumers. These municipalities include cities and counties from Alaska, California, Colorado, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, and Texas. Although this program is a success in many municipalities, technological issues are potential barriers for some cities to post their inspection results on Yelp. According to research from the open data company Socrata (Shueh, 2015), a large number of U.S. cities neither publish restaurant inspection information, nor do they use a digital-friendly format to present it. Based on my communication to a local county officer, satisfaction of the status quo could be another reason for municipalities to stay away from this program.

Literature review

As Stigler (1961) points out, a search for quality information about a product could be more difficult than a search for price information. Currently, the price information for meals in a restaurant can be easily found by clicking on the restaurant's website or walking in to the restaurant and reading through the menu. However, details about the quality of the food and service in a restaurant are more challenging for consumers to discover. The fact that sellers are more knowledgeable about the quality of their products than buyers results in information asymmetry (Akerlof 1970). The asymmetric information between restaurants and consumers puts those consumers at a disadvantage in the market (Akerlof 1970).

There are numerous ways to mitigate asymmetric information between sellers and buyers, such as guarantees, brand-name goods, chains, and licensing practices (Akerlof 1970). These methods provide a standard for the quality of products and services. Product/service quality reports function similarly by increasing product/service quality information to consumers. This could also diminish asymmetric information and impact consumer and producer behaviors.

A growing body of literature examines how consumers respond to information entailing product/service quality, most of which explores the health economics field (Beaulieu 2002; Idig and Tai-Seale 2002; Jin and Sorensen 2006; Dafny and Dranove 2008; Bundorf et al. 2009). Consumers' responsiveness to service quality information in terms of health plan choices was confirmed by Beaulieu (2002) and Dafny and Dranove (2008). Wedig and Tai-Seale (2002) and Jin and Sorensen (2006) verified their openness to service quality information concerning health insurance choices, while Bundorf et al. (2009) further confirmed this responsiveness regarding clinic selections. These studies investigate the introduction of report cards, published ratings on consumers choices, and find evidence of behavioral change of consumers in response to healthcare service quality information. In addition, Pope (2006) finds evidence of patients' and students' responses to hospital and college rankings. Similarly, Hastings and Weinstein (2008) confirm the positive effect of school test score information on public school choices of low-income families.

Another array of research examines consumers' perception and behavior on restaurant hygiene and Yelp. Recent empirical works conducted by Henson et al. (2006) and Aksoydan (2007) suggest that consumers determine their restaurant choices based on their observed judgement of restaurant hygiene. Utilizing feedback from focus groups and a postal survey in Ontario, Canada, Henson et al. (2006) conclude that in addition to inspection results, the number of customers in a restaurant contributes to consumers' dining choices. Similarly, Aksoydan (2007) summarizes that restaurant hygiene influences consumers' decisions on restaurants the most after studying 243 academic staff in Ankara, Turkey. Consumers' responses to letter grade hygiene information and numerical hygiene scores are studied by Kang (2015). He finds that A grades and larger inspection scores are associated with higher perceived quality. Luca (2016) shows a Yelp rating improvement of one star leads to a 5-9% increase in the restaurant revenue.

Producers value product/service quality information because consumers' learning of product/service quality may result in demand change. Due to this, producers may adjust their practice accordingly. A group of empirical research shows how producers react to information disclosure with regards to their product/service quality. The major conclusion is that information disclosure reduces violations of drinking water standards from community drinking water suppliers in Massachusetts (Bennear and Olmstead, 2008) and pollution from pulp and paper plants in India (Powers et al. 2008). Information disclosure also improves restaurant hygiene quality in Los Angeles County (Jin and Leslie 2003,2009), and service quality of nursing homes (Lu, 2012).

Contrary to the major conclusions of the positive effects of information disclosure, Dranove et al. (2013) find that the introduction of report cards on health care providers entice doctors to reject treating grievously ill patients in order to achieve better outcomes on report cards. This kind of selection behavior resulted in worse health outcomes for sicker patients.

Despite of the research findings on how producers react to health service quality information in the health care industry, there is limited research on how quality information affects producers in the restaurant industry. Unlike the health care industry where the health care providers have market power to select patients that are more likely to produce positive outcomes (Dranove et al. 2013), the regulations on restaurants are enforced by local agencies. This leaves restaurants no room to select inspectors. Therefore, if there are observed improvements in restaurant hygiene, they are not due to selection behaviors.

Although consumers' response to report cards/grades on restaurant hygiene have been studied by economists, there is quite a gap in understanding the disclosure of restaurant hygiene information on restaurants' hygiene quality. Jin and Leslie (2003) study the effect of restaurant hygiene grade cards in Los Angeles from 1996 to 1998 on restaurant hygiene quality and find a positive effect caused by the introduction of grade cards. Their study provides insights on how mandatory and voluntary disclosure affect restaurant hygiene quality. However, their study does not capture new ways of disclosing hygiene quality information online. Makofske (2017) estimates that disclosing restaurant hygiene information on Yelp leads to a restaurant inspection violation score reduction by 12-14% in Louisville, KY. Makofske (2017) 's paper is very similar to this study since both studies examine posting health inspection information on restaurant hygiene quality. While Makofske (2017) only analyzes inspection violation data from one city that has the information disclosure program, this study applies treated and control cities to resemble a natural experiment to estimate the average treatment effect of the LIVES program on restaurant hygiene quality. This could result in the omitted variable bias. This study provides new evidence in understanding the role that information plays in a monopolistic competition market like the restaurant industry.

Empirical framework

The goal of this study is to examine how public information disclosure affects producers' quality choice. In particular, I study the causal effect of the LIVES program on restaurant hygiene quality.

The LIVES program began in 2012 with the purpose of increasing consumers' access to restaurants' hygiene information and helping consumers make their dining choice (Yelp, 2018). Its main function is to let municipalities post restaurant inspection information on Yelp. This practice provides additional and easier access to consumers to navigate restaurant inspection information, as the information on Yelp is timely and easy to find. Prior to the LIVES program, consumers had to search restaurant inspection reports on municipality websites, local newspapers, and other media outlets, which sometimes could be tricky and outdated.

North Carolina is a good choice to study the effect of the LIVES program for many reasons. First, all counties in North Carolina have used the same inspection form since 2013 so that inspection score in one county is comparable to another; whereas this situation does not apply to other states participating in the LIVES program. Second, there are 98 counties in North Carolina, five of which participated in the LIVES program to bring restaurant inspection data to consumers on Yelp. This allows us to pick treated and control groups to mimic a natural experiment. By comparing a treated group that participated in the LIVES program with a similar control group that did not participate, we can conclude that the LIVES program caused changes in the outcome of the treated group.

Figure 1 is the North Carolina State map. Five counties that participated in the LIVES program are highlighted. Orange, Wake, Harnett and Cumberland Counties are located in the middle of

North Carolina, while Mecklenburg County sits on the west side. A list of counties entailing county demographics in North Carolina can be found in Table 1 in the appendix.

The LIVES program has slowly and gradually been implemented across North Carolina. According to Yelp (2018), the steps for a municipality to participate in the LIVES program include: 1) create a feed under Yelp's data requirements, 2) host the feed on either a HTTP or HTTPS, 3) send an email to Yelp attaching a link to the feed. Upon receipt of the email, Yelp will validate, test, and launch the data. Wake County was the first county in North Carolina to participate in the LIVES program in October of 2013, followed by Orange County joined in September 2014. More than two years later, Cumberland County brought inspection data on Yelp in April of 2017 and Mecklenburg did the same in May 2017. Harnett county was the newest member in this group, starting from the first quarter of 2018. According to Susan Cole from the health department in Mecklenburg County, their partnership with Yelp is in line with their objective and emphasis to safeguard the public's health and safety (Douglas, 2017). Perhaps this kind of agreement on helping consumers ensure food safety between local government and Yelp helps explain the expansion of the LIVES program.

In this study, we seek to use two empirical approaches to study the causal effect of the LIVES program on restaurant hygiene quality: the Difference in Difference (DID) approach and the Geographic Regression Discontinuity (GRD) approach. Both approaches require a treated group and a control group. Comparing the difference between a treated group and a control group, with their difference being the absence of treatment, helps establish the causal inference of the treatment (Dunning, 2012). Applying two empirical methods helps us critically evaluate the estimated effect of the LIVES program on restaurant hygiene.

Orange County is selected as the treated group while Durham County is chosen as the control group to study the effect of the LIVES program in North Carolina on restaurant hygiene quality for several reasons. First, Mecklenburg and Wake Counties are the most and second most populous counties. This causes difficulty in finding them a control group that has about the same population so that the treated and control groups are similar. Second, Harnett County has only participated in the LIVES program since the first quarter of 2018, which resulted in a shortage of inspection data after the policy is implemented. A similar situation applies to Cumberland County. Third, the GRD approach requires that the treated and control counties are adjacent to each other. Orange County and Durham County share a county border.

Data for Difference-in-Difference Approach

The dataset for Difference-in-Difference study includes restaurant inspection data and restaurant characteristic data. The restaurant inspection data consists of inspection reports from the treated group (Orange County) and the control group (Durham County). The main sources for restaurant inspection data for Orange and Durham Counties are the LIVES program website and the Durham County website. Each inspection report contains a unique business identification number, inspection date, inspection score, inspection type, and a description of violations. The inspection dates range from 2013 to 2019. Inspection scores range from 0 to 100, with 0 indicating the restaurant completely violates the regulations and 100 indicating the restaurant fully meets all inspection criteria. Inspection types fall into three categories: routine, follow up, and complaint. In this study, only routine inspections are included because they describe the general hygiene quality in restaurants. Violations are divided into critical and non-critical violations. Critical violations indicate incidents related to foodborne disease outbreaks, while non-critical violations concern minor food handling misconduct.

Restaurant characteristics data for Orange and Durham Counties come from Yelp.com. On Yelp, restaurants are divided into four classifications on the basis of price: inexpensive, moderate, pricey, and ultra high-end. These classifications are represented by "$", "$$", "$$$", and "$$$$", respectively. However, the dataset in this study only includes restaurants in the first three categories. A "$" restaurant denotes that the average price for a meal per person in this restaurant is below $10, "$$" indicates that the price is between $11 and $30, and "$$$" means the price ranges from $30 to $60. Moreover, another restaurant characteristic is the restaurant category. A restaurant category can be defined by the food ethnicity, the speed of food served, the food cooking method, the table service, and so on (e.g. "Chinese, fast food").

Using observational data for causal inference often involves dealing with confounding variables that influence pretreatment control variables (Iacus et al. 2011). Matching is used to ensure that the distributions of the covariates in the treated and control groups are more similar, or more balanced (Iacus et al. 2011). To reduce the imbalance of restaurant characteristics between the treated and control groups, I apply Coarsened Exact Matching (Iacus et al. 2011) on the aggregated dataset from inspection data and restaurant characteristic data. First, I choose the covariate restaurant category and coarsen it according to the restaurant classification. For example, Sushi Bar and Japanese are coarsened into the same category. Then, each restaurant inspection in the treated group within a certain restaurant category falls into a stratum, and it is matched to a corresponding restaurant inspection from the control group. Afterwards, I exclude the unmatched observations so that the number of inspections in the treated and control group are the same. The final matching result can be found in Table 2 (Observations in the treated group is 850+2770=3620. This is equal to the observations in the control group, 671+2949=3620).

Table 2 reports the summary statistics of the restaurant inspection data from Orange County and Durham County. After the CEM method, the observations of restaurant inspection for Orange and Durham Counties are the same. Orange County has a larger mean inspection score and a less divergent distribution of scores.

Sample means for the treated and control groups before and after the policy provide a preliminary measure of the effect of the program. Table 3 reports the mean inspection scores for the treated and control groups before and after the LIVES program is implemented. The mean inspection scores for the treated group is 0.22 points less than that of the control group before the policy is in effect. However, after the program is in effect, the mean inspection score for the treated group increases to 97.94 points, which is about 1.32 points more than that of control group. Thus, this change indicates that implementing the LIVES program improves inspection scores in the treated group.

Empirical Specification -Difference in difference approach

To estimate the effect of joining Yelp's LIVES program on restaurants' hygiene, I apply a difference in difference (DID) approach. The DID approach allows researchers to use observational data to simulate the natural experiment to establish causal relationship by comparing the difference between the treated and control groups before and after the treatment (Angrist and Pischke, 2009)

One core assumption for utilizing the DID approach is the parallel trend assumption (Angrist and Pischke, 2009). This assumption requires that the macroeconomics trends, seasonality, and unobserved heterogeneity remains consistent after the treatment. The application of CEM method creates a quasi-control group that guarantees heterogeneity between the control and treated groups are significantly reduced. The careful selection of the treated and control groups and the application of CEM ensure that the main difference between Orange County and Durham County is the treatment of the LIVES program.

I estimate the following regression:

$$y_{ict} = \beta_0 + \beta_1 After_t + \beta_2 After_t \times Treat_c + Restaurant + Year + Zipcode + Month + \varepsilon_{ict} \ (1)$$

where subscript c is county, t is time period, and i is an individual restaurant.

In the regression, y_{ict} denotes the restaurant inspection score in county c for restaurant i at time t. Variable $Treat_c$ is a dummy variable and is 1 for the county that participated in the LIVES program, 0 otherwise. $After_t$ is a dummy variable differentiating the time period before and after the treatment of joining the LIVES program. It captures the aggregated change in hygiene scores.

The variable $Restaurant$ controls for the unobserved time-invariant factors that affect hygiene quality in each restaurant. These time-invariant factors include food type, operation hours, seating capacities, and so on. The restaurant fixed effects rule out the effect of these time-invariant restaurant characteristics on restaurant hygiene by creating a dummy variable for n-1 restaurants. There are 784 restaurants after dropping 61 singleton observations. 783 dummy variables for restaurants are created though their coefficients are not reported. The variable $Zipcode$ denotes ZIP Code-specific fixed effects. It controls for the time-invariant heterogeneity between each ZIP Code area. There are 27 ZIP Code areas in this study. The variable $Year$ represents year fixed effect. The variable $Month$ controls for seasonality within a year.

The difference of the expected inspection score for the treated group after and before the treatment is $E(y_{ict}|c = 1, t = 1) - E(y_{ict}|c = 1, t = 0) = \beta_1 + \beta_2$ (2)
The difference of the expected inspection score for the control group after and before the treatment is $E(y_{ict}|c = 0, t = 1) - E(y_{ict}|c = 0, t = 0) = \beta_1$ (3)

Thus, the difference of the above differences is

$$\{E(y_{ict}|c = 1, t = 1) - E(y_{ict}|c = 1, t = 0)\} - \{E(y_{ict}|c = 0, t = 1) - E(y_{ict}|c = 0, t = 0)\}$$
$$= \beta_2 \ (4)$$

β_2, the primary estimate interest, measures how the LIVES program affects restaurants' hygiene scores in Orange County.

Results

Table 3 provides an overview of how sample means change across treated and control groups after the policy. Table 4 reports the main difference-in-difference results with Orange County being the treated group and Durham County being the control group.

The main estimate of interest, β_2, is 1.135. It demonstrates the effect of joining the LIVES program on average improved Orange County restaurant inspection score by 1.135 point. This evidence suggests that restaurants in Orange County respond to the policy change and make hygiene improvements. Economic theory predicts more information to consumers is usually better (Jin and Leslie, 2003). This empirical result supports such policies that aim to provide more information to consumers.

Test for Parallel Trend Assumption

The core assumption for the difference-in-difference approach is the parallel trends assumption. Parallel trends assumption requires that the treated and control groups have similar macroeconomic trends, seasonality, and unobserved heterogeneity. To test whether this assumption holds, I create a placebo dummy for the two months before the program is implemented. Because the LIVES program is implemented in September 2014 for Orange County, the placebo dummy takes the value of zero in July 2014 and one in August 2014. The placebo dummy takes the place of the After variable in the robustness check and interacts with the Treat variable.

Table 5 presents the robustness check results. The estimates for placebo dummy and placebo interaction with treat variable are both insignificant, which suggests that the treated county and the control county do have the same trends before the change in information disclosure policy.

One may argue that improving the inspection score of one point out of 100 is trivial. However, the mean inspection score for Orange County before the policy is 97.16. In other words, the LIVES program reduced hygiene violation scores by 1.135 points out of 2.84, which is about 39.96% of a reduction in violation scores. This effect of the LIVES program in this study is larger than the 12-14% of a reduction in health violations reported by Makofske (2017).

Empirical Specification-Geographic Regression Discontinuity Design

In addition to the difference-in-difference approach, I apply the geographic regression discontinuity (GRD) design to measure the causal effect of policy change by studying restaurants located on two sides of a geographical boundary. The difference-in-difference approach is widely used to examine the effect of a program in economics. However, Geographic Regression Discontinuity is often used in political science and criminology when a geographic border defines the cutoff of the treated and control groups. In this study, I sample restaurants located next to the Orange-Durham County border to estimate the average treatment effect of the LIVES program. On one side of the Orange-Durham border, restaurants are subject to inspection results revealed on Yelp while restaurants on the other side are not. There must be a distinguishable difference in outcomes to prove significant effect evidence (Thistlewaite and Campbell, 1960).

The regression discontinuity analysis serves as an essential addition to the difference-in-difference approach. The validity of regression discontinuity relies on whether the treatment is randomly assigned to agents around the threshold (Lee, 2008). For restaurants located near the Durham-Orange County border, their locations in or outside the treated county determines whether they are treated or not. Moreover, they cannot manipulate their locations to avoid treatment. Therefore, the treatment assignment could be considered as a random assignment.

Restaurant inspection data come from Orange County and Durham County's websites. I select the most recent four inspection scores that ranged from September 2014 to February 2019 for each restaurant in Orange and Durham Counties after the LIVES program was enforced in September 2014. The dataset contains characteristics of restaurants, such as business identification number, name, street, city, and ZIP Code. Inspection information such as inspection date and inspection score are also in the dataset. I use the Awesome Table in Google Sheets to convert restaurant addresses into latitudes and longitudes that represent locations of the restaurants. In Geographic Information System (GIS) software, I plot these geographic coordinates of restaurants on the map and calculate the distance of these restaurants to the Orange-Durham County border.

For restaurants located within Orange County (treated county), their distances to the Orange-Durham County border are written in positive numbers. The distances of Durham County restaurants to the border are recorded in negative numbers. In this way, the geographic border is the cutoff between the treated and control county. Mathematically,

$$Treat_c = \begin{cases} 1 \ if \ Distance_i \geq 0, \\ 0 \ if \ Distance_i \leq 0, \end{cases} \quad (5)$$

where $Treat_c$ is the dummy variable that distinguishes the treated and control counties. Orange County (treated) takes the value of one and Durham County (control) takes the value of zero. $Distance_i$ denotes the distance of a restaurant to the border. Positive distances represent restaurants in Orange County, and vice versa.

Figure 2 provides a visual illustration of how restaurants in Orange and Durham Counties are spatially located. Restaurants in Orange County scatter in the center and around the border while restaurants in Durham County concentrate in the west side of the county. The populations in Orange and Durham Counties are 144,946 and 311,640, respectively. It helps explain why there are more restaurants in Durham County than in Orange County. In addition, there are some restaurants located very close to the border because interstate 40 and 85 cross the border. This helps with identification strategy because the border help randomize the treated and control restaurants.

Then, I estimate the following specification:

$$\tau = \overline{y_t} - \overline{y_c} \ for \ -h_{opt} < d < h_{opt} \ (6)$$

In equation (6), $\overline{y_t}$ is the mean inspection score restaurants located in the treated county with distances less than the optimal distance after the program is in effect. $\overline{y_c}$ is the mean inspection score for restaurants located in control county with distances less than the optimal distance after the program is in effect. The running variable is distance. h_{opt} is the optimal bandwidth for regression discontinuity analysis. The optimal bandwidth minimizes the asymptotic mean squared error of the treatment effect and its estimate (Li, 1987, Imbens and Kaalyanaraman, 2011.). τ measures the causal effect of the LIVES program on restaurant hygiene.

Figure 3 is a density plot with restaurants' distances on the x-axis and their densities on the y-axis. Restaurants in Orange County are represented with miles greater than zero and restaurants in Durham County are represented with miles less than zero. According to the figure, there is a slight difference in the distribution of restaurants in the treated and control counties. In other words, restaurants in the control county are more concentrated than restaurants in the treated county.

Geographic Regression Discontinuity Results

Three methods are applied to calculate the optimal bandwidth: the Cross-Validation (CV) method, the Imbens and Kalyanaraman (IK) method, and the Calonico, Cattaneo, and Titunik (CCT) method. The Cross-Validation method is proposed by Ludwig and Miller (2005). They find the optimal bandwidth by fitting a model with all observations that result in a minimum of mean integrated squared error (MISE) (Imbens and Kalyanaraman, 2012). Imbens and Kalyanaraman (2012) propose an asymptotically optimal bandwidth that minimizes the first order approximation of mean squared error of the treatment parameter and treatment parameter estimate. In addition, Calonico, Cattaneo, and Titiunik (2014) proposed the bias-corrected estimator from estimating the bias in the distributional approximation and then deducting it from the point estimate.

Table 6 reports the conventional coefficient estimates from three different approach: CV, IK and CCT. A conventional estimate is the point estimate that minimizes the mean squared error between the effect parameter and its estimate. The estimates from IK and CCT methods are both statistically significant at 5% significance level. On the contrary, the Cross-Validation method selected the smallest bandwidth out of the three aforementioned methods and the estimate is not statistically significant. The Cross-Validation method uses only a small portion of observations close to the threshold, which may result in an unprecise bandwidth (Imbens and Kalyanaraman, 2012). The Cross-Validation method here may not be as credible as the other two methods.

The estimates from the GRD approach, 0.58716 (IK estimate) and 0.53169 (CCT estimate), are much smaller than the Difference-in-Difference estimate (1.168). Their statistical significances are not as resilient as the DID estimate and their standard errors outweigh the DID estimate. This evidence lends credibility to the DID estimate. First, the DID approach controls for time invariant restaurant fixed effects, Zip Code fixed effects, and removes the trend, which guarantees the accuracy of the estimate. Second, the GRD approach has low external validity. It only estimates the local treatment effect of restaurants located 7.298 and 9.303 miles away from the border. Therefore, the estimate only reflects the improvement in hygiene scores in those areas that are close to the border. Third, in terms of reducing heterogeneity between restaurants in the treated and control groups, the common trends assumption does a better job than GRD's random assignment assumption.

Discussion

Both the difference-in-difference approach and regression discontinuity provide compelling evidence that implementing the LIVES program improves restaurant inspection hygiene in the treatment group, which is in line with previous studies. Directly comparing the magnitude of the results of this study with previous studies is not reasonable since the inspection criteria vary across states, and sometimes across counties. For example, Jin and Leslie (2003) conclude the mandatory restaurant hygiene disclosure brought from onsite grade cards improve inspection scores by 4.4 points in Los Angeles County. Even though both Los Angeles County and Orange County denote 100 points as perfectly clean, the different distributions of inspection items and different weights on inspection items make the comparison of hygiene quality improvements unrealistic.

This study also supports Makofske (2017)'s study on the effect of the Louisville-Yelp Partnership on restaurant hygiene. Makofske (2017) finds posting inspection reports of Louisville restaurants on Yelp induces restaurants to reduce 12-14% health violation scores (a health violation score is 100 minus the inspection score). Although this paper provides an estimate of different magnitude, both papers confirm the positive effect of disclosing inspection on Yelp in regard to improving restaurant hygiene quality. These empirical studies support the theory that reducing information asymmetry on quality between consumers and producers will prompt producers to improve the quality of their products.

What motivates restaurants to improve quality after information is disclosed to the public? Jin and Leslie (2003) suggest that demand at restaurants with good hygiene may increase and demand at restaurants with bad hygiene may decrease. Therefore, prices may hike in restaurants with high inspection scores and prices may plunge in restaurants with low inspection scores. In this study, prices in restaurants are assumed to stay relatively stable because the study period is in the short run. The driving force for restaurants to improve hygiene quality is the changing demand that comes from Yelp-informed consumers after inspection information disclosure.

The Geographic Regression Discontinuity provides us with another approach to analyze the causal effect of the LIVES program, however, there are several limitations for this research design. First, this study cannot separate the effect caused by the LIVES program from the unobserved compound treatments where two or more treatments were imposed on the treated county, if there is any. That is, the estimate of Geographic Regression Discontinuity measures the effect of multiple policies, including the LIVES program on the treated county. This poses a serious threat to the validity of Geographic Regression Discontinuity.

Conclusion

This paper examines the effect of publicly disclosed restaurant inspection information on Yelp on restaurants' hygiene quality. The theoretical model indicates that after the disclosure of restaurant inspection information on Yelp, restaurants may improve their hygiene quality. The empirical findings support this possibility and find that Orange County restaurants improved restaurant inspection scores after Orange County implemented this practice of disclosing inspection information on Yelp.

With adequate data at hand, this study is capable of using two empirical approaches to examine the effect of the LIVES program on restaurant hygiene. Both approaches support the positive effect of the LIVES program. This result provides new evidence of support for the argument that providing more information about product quality to consumers improves product quality. This study utilizes the strength of both approaches and concludes the positive effect of the LIVES program.

This study confirms the positive effects the LIVES program played in increasing easier access to consumers and improving restaurant hygiene. Counties who are hesitant to participate in the LIVES program now have a reason to join. Compared to the enormous public cost (more than $15.6 billion in 2014) and deaths (2,377 Americans annually) due to foodborne pathogens (Flynn, 2014), the cost of paying officers to organize and transform inspection data into the format that the LIVES program requires is trivial.

This study also confirms the benefits of making government data available to the public. The LIVES program increases consumers' access to government data by putting inspection data on Yelp. Rather than review the county government website, consumers who regularly visit Yelp are more likely to use this platform before choosing a restaurant to eat. With the growing population of Yelp users, Yelp reduces the time consumers spend on searching for restaurants. This study provides support for programs that strive to bring government data to the public.

Appendix

Table 1: List of North Carolina Counties and their demographics

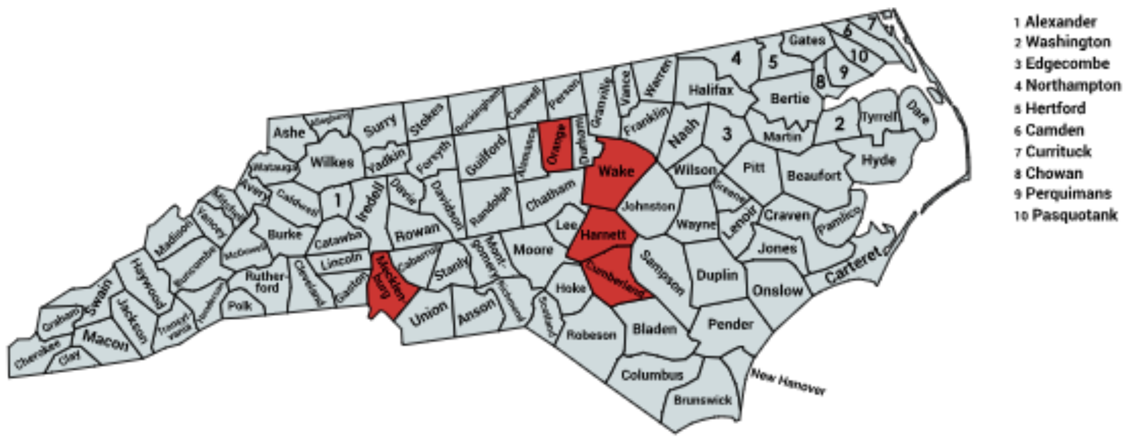
County Name	Population (2017 July 1, estimate)	Households with a computer, percent, 2013-2017	Households with a broadband Internet subscription, percent, 2013-2017	Join LIVES program
Alamance	162,391	0.85	0.74	No
Alexander	37,286	0.76	0.67	No
Alleghany	11,031	0.77	0.69	No
Anson	24,991	0.75	0.62	No
Ashe	26,957	0.76	0.66	No
Avery	17,536	0.76	0.63	No
Beaufort	47,088	0.74	0.61	No
Bertie	19,224	0.67	0.52	No
Bladen	33,478	0.75	0.61	No
Brunswick	130,897	0.89	0.78	No
Buncombe	257,607	0.86	0.79	No
Burke	89,293	0.76	0.66	No
Cabarrus	206,872	0.91	0.85	No
Caldwell	81,981	0.79	0.70	No
Camden	10,581	0.88	0.80	No
Carteret	68,881	0.89	0.79	No
Caswell	22,646	0.74	0.61	No
Catawba	157,974	0.84	0.75	No
Chatham	71,472	0.86	0.76	No
Cherokee	28,087	0.78	0.66	No
Chowan	14,105	0.78	0.64	No
Clay	11,074	0.85	0.74	No
Cleveland	97,334	0.73	0.62	No
Columbus	55,936	0.71	0.58	No
Craven	102,578	0.88	0.78	No
Cumberland	332,546	0.88	0.79	**Yes, April 12, 2017**
Currituck	26,331	0.90	0.80	No
Dare	36,099	0.93	0.85	No

Davidson	165,466	0.82	0.73	No
Davie	42,456	0.85	0.77	No
Duplin	59,039	0.76	0.58	No
Durham	311,640	0.90	0.84	No
Edgecombe	52,747	0.72	0.58	No
Forsyth	376,320	0.87	0.78	No
Franklin	66,168	0.83	0.72	No
Gaston	220,182	0.83	0.75	No
Gates	11,544	0.77	0.64	No
Graham	14,814	0.84	0.72	No
Granville	59,557	0.84	0.73	No
Greene	21,015	0.80	0.69	No
Guilford	526,953	0.85	0.71	No
Halifax	51,310	0.73	0.56	No
Harnett	132,754	0.88	0.75	**Yes, the first quarter of 2018**
Haywood	61,084	0.81	0.68	No
Henderson	115,708	0.85	0.77	No
Hertford	23,906	0.73	0.57	No
Hoke	54,116	0.84	0.76	No
Hyde	5,363	0.72	0.48	No
Iredell	175,711	0.87	0.77	No
Jackson	42,973	0.82	0.65	No
Lee	60,430	0.83	0.74	No
Lenoir	56,883	0.79	0.71	No
Lincoln	82,403	0.84	0.75	No
McDowell	45,159	0.80	0.67	No
Macon	34,732	0.80	0.68	No
Madison	21,746	0.77	0.67	No
Martin	22,789	0.74	0.58	No
Mecklenburg	1,076,837	0.92	0.84	**Yes, May 2017**
Mitchell	15,072	0.75	0.63	No
Montgomery	27,435	0.72	0.61	No
Moore	97,264	0.85	0.79	No
Nash	93,991	0.80	0.69	No
New Hanover	227,198	0.90	0.82	No
Northampton	19,862	0.68	0.53	No
Onslow	193,893	0.92	0.83	No

Orange	144,946	0.93	0.87	**Yes, Sep 16, 2014**
Pamlico	12,689	0.83	0.68	No
Pasquotank	39,743	0.81	0.69	No
Pender	60,958	0.86	0.71	No
Perquimans	13,474	0.80	0.67	No
Person	39,370	0.79	0.70	No
Pitt	179,042	0.86	0.76	No
Polk	20,558	0.84	0.74	No
Randolph	143,282	0.80	0.68	No
Richmond	44,798	0.74	0.65	No
Robeson	132,606	0.68	0.50	No
Rockingham	90,949	0.78	0.67	No
Rowan	140,644	0.84	0.73	No
Rutherford	66,551	0.76	0.63	No
Sampson	63,430	0.76	0.58	No
Scotland	35,093	0.73	0.55	No
Stanly	61,482	0.83	0.73	No
Stokes	45,717	0.79	0.67	No
Surry	72,224	0.75	0.62	No
Swain	14,294	0.61	0.45	No
Transylvania	33,956	0.88	0.75	No
Tyrrell	4,052	0.68	0.55	No
Union	231,366	0.92	0.87	No
Vance	44,211	0.76	0.61	No
Wake	1,072,203	0.95	0.88	**Yes, Oct 2013**
Warren	19,883	0.72	0.54	No
Washington	12,012	0.69	0.56	No
Watauga	55,121	0.88	0.79	No
Wayne	124,172	0.83	0.73	No
Wilkes	68,576	0.78	0.67	No
Wilson	81,671	0.81	0.71	No
Yadkin	37,774	0.77	0.66	No
Yancey	17,744	0.71	0.59	No

Source: United States Census Bureau

Figure 1: Counties Participating in the LIVES Program in North Carolina.



Source: https://mapchart.net/usa-counties.html

Table 2. Summary Statistics of Inspection Data from Cumberland and Guilford Counties

County Name	Restaurants	Inspections	Mean Score	Std. Dev of Score
Orange County	268	3620	97.7565	2.0223
Durham County	577	3620	96.68	2.4052

Table 3. Difference-in-Difference for mean inspection score

Time period	Group (County)	Obs	Mean	Std. Dev
Before	Treated (Orange)	850	97.16	2.2557
	Control (Durham)	671	96.9367	2.3562
	Difference		0.2233	
After	Treated (Orange)	2770	97.9395	1.9085
	Control (Durham)	2949	96.6216	2.4128
	Difference		1.3179	
	Difference-in-Difference		1.0946	

Table 4. Difference-in-Difference Results

VARIABLES	Fixed effects
After	-0.450
	(0.314)
Treat*after	1.135***
	(0.204)
Constant	97.14***
	(0.262)
Observations	7,179
R-squared	0.590
Business_id FE	YES
Zipcode FE	YES
Year FE	YES
Month FE	YES

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

Table 5. Test for Parallel Trends Assumption

VARIABLES	Fixed effects
Placebo	0.279
	(0.169)
Treat*placebo	-0.0551
	(0.160)
Constant	97.21***
	(0.00236)
Observations	7,179
R-squared	0.581
Business_id FE	YES
Zipcode FE	YES
Year FE	YES
Month FE	YES

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

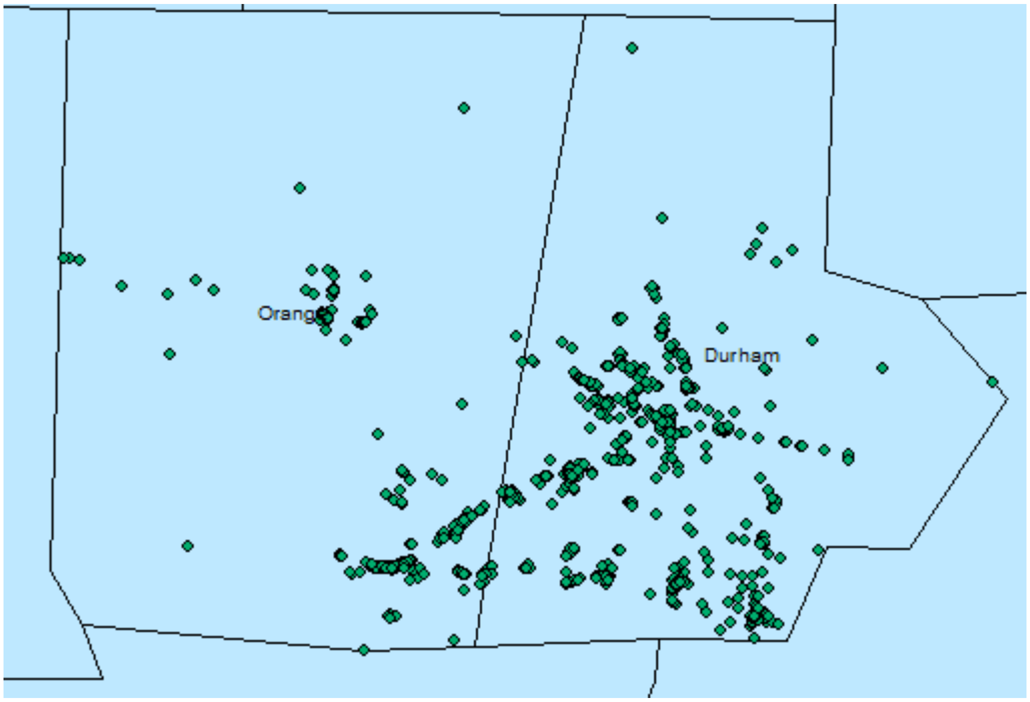


Figure 2: Restaurants in Orange and Durham County

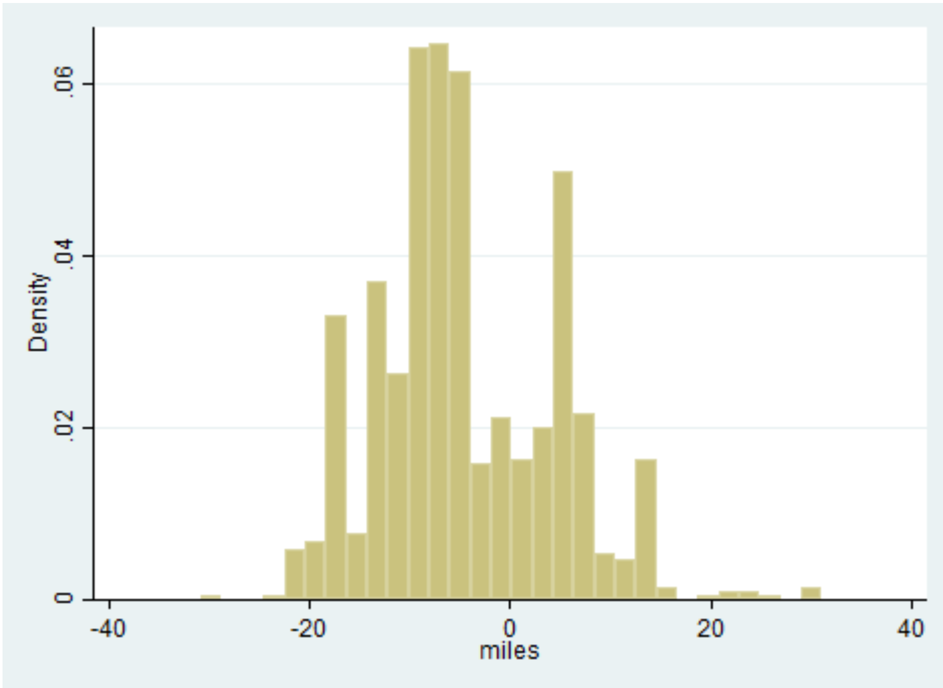


Figure 3: Restaurant density plot

Table 6: Sharp RD Estimates Using Local Polynomial Regression

	CV	IK	CCT
Conventional coefficient	0.5752	0.58716**	0.53169**
Conventional Std. Err.	0.38909	0.26675	0.2407
Number of Observations	2932	3817	3817
Bandwidth (miles)	4.772	7.298	9.303
Kernel Type	Triangular	Triangular	Triangular

*** p<0.01, ** p<0.05, * p<0.1

References

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. The Quarterly Journal of Economics, 84(No.3), 488–500.

Aksoydan, E. (2007). Hygiene Factors Influencing Customers' Choice of Dining-out Units: Finding from a Study of University Academic Stuff. Journal of Food Safety, 27(3), 300–316.

Angrist, J. D.; Pischke, J. S. (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Beaulieu, N. D. (2002). Quality information and consumer health plan choices. Journal of Health Economics, 21(1), 43–63.

Bennear, L. S., & Olmstead, S. M. (2008). The impacts of the "right to know": Information disclosure and the violation of drinking water standards. Journal of Environmental Economics and Management, 56(2), 117–130.

Bundorf, M. K., Chun, N., Goda, G. S., & Kessler, D. P. (2009). Do markets respond to quality information? The case of fertility clinics. Journal of Health Economics, 28(3), 718–727.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs. Econometrica, 82(6), 2295–2326.

Dafny, L., & Dranove, D. (2008). Do report cards tell consumers anything they don't already know? The case of Medicare HMOs. The RAND Journal of Economics, 39(3), 790–821.

Douglas Anna. (2017) What did restaurant inspectors find in Charlotte kitchens? Yelp has the answer. The Charlotte Observer. Retrieved from

https://www.charlotteobserver.com/news/business/article153401854.html

Dranove, D., Kessler, D., McClellan, M., & Satterthwaite, M. (2013). Is More Information Better? The Effects of "Report Cards" on Health Care Providers. Journal of Political Economy, 34.

Dunning, Thad. (2012). Natural Experiments in the Social Sciences: A Design-Based Approach. Cambridge University Press.

FDA. (2017). Adoption of the FDA Food Code by State and Territorial Agencies Responsible for the Oversight of Restaurants and Retail Food Stores 2017. Retrieved from

https://www.fda.gov/media/107543/download

FDA. (2017). Food Code 2017. Retrieved from
https://www.fda.gov/food/fda-food-code/food-code-2017

Lu, Susan Feng. (2012). Multitasking, Information Disclosure, and Product Quality: Evidence from Nursing Homes. Journal of Economics & Management Strategy, 21(3), 673–705.

Flynn Dan. (2014). USDA: U.S. Foodborne Illnesses Cost More Than $15.6 Billion Annually. Food Safety News. Retrived from https://www.foodsafetynews.com/2014/10/foodborne-illnesses-cost-usa-15-6-billion-annually/.

Hastings, J. S., & Weinstein, J. M. (2008). Information, School Choice, and Academic Achievement: Evidence from Two Experiments. The Quarterly Journal of Economics, 1373–1414.

Henson, S., Majowicz, S., Masakure, O., Sockett, P., Jones, A., Hart, Carr, D., Knowles, L. (2006). Consumer assessment of the safety of restaurants: The role of inspection notices and other information cues. Journal of Food Safety, 26(4), 275–301.

Hoffman, Sandra. (2011). U.S. Food Safety Policy Enters a New Era. Retrieved from
https://www.ers.usda.gov/amber-waves/2011/december/us-food-safety-policy/

Iacus, S.M., G. King, & G. Porro. (2011). Casual Inference Without Balance Checking: Coarsened Exact Matching. Political Analysis 20(1): 1-24.

Imbens, G., & Kalyanaraman, K. (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. The Review of Economic Studies, 79(3), 933–959.

Jin, G. Z., & Leslie, P. (2009). Reputational Incentives for Restaurant Hygiene. American Economic Journal: Microeconomics, 1(1), 237–267.

Jin, G. Z., & Leslie, P. (2003). The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards. The Quarterly Journal of Economics, 118(2), 409–451.

Jin, G. Z., & Sorensen, A. T. (2006). Information and consumer choice: The value of publicized health plan ratings. Journal of Health Economics, 25(2), 248–275.

Kang, D. (2015). Effect of Information on Consumer Perception: Evidence from Restaurant Hygiene Grade Cards in New York City. AAWE Working Paper, (No.183).

Lee (2008). Randomized Experiments from Non-random Selection in U.S. House Elections. Journal of Econometrics. 142 (2): 675–697.

Li Ker-Chau. (1987). Asymptotic Optimality for C_p, C_L, Cross-Validation and Generalized Cross-Validation: Discrete Index Set. The Annals of Statistics. 1987, Vol. 15, No. 3, 958-975.

Luca, Michael. (2016). “Reviews, Reputation, and Revenue: The Case of Yelp. Com.” Working Paper 12-016.

Ludwig, J., & Miller, D. L. (2007). Does Head Start Improve Children’s Life Changes? Evidence from a Regression Discontinuity Design. The Quarterly Journal of Economics, 122(1), 159–208.

Makofske, M. P. (2017). The Effect of Information Salience on Product Quality: Louisville Restaurant Hygiene and Yelp.com. MPRA Paper No.79690.

Pope, D. G. (2006). Reacting to Rankings: Evidence from “America’s Best Hospitals and Colleges.” Job Market Paper.

Powers, N., Blackman, A., Lyon, T. P., & Narain, U. (2008). Does Disclosure Reduce Pollution? Evidence from India’s Green Rating Project. Resources for the Future DP 08-38.

Shueh, Jason. (2015). Yelp, Socrata Partner to Make Restaurant Inspections Public. Government Technology. Retrieved from
https://www.govtech.com/data/Yelp-Socrata-Partner-to-Make-Restaurant-Inspections-Public.html

Stigler, G. J. (1961). The Economics of Information. Journal of Political Economy, Vol. 69(No.3), 213–225.

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. Journal of Educational Psychology, 51(6), 309–317.

Wedig, G. J., & Tai-Seale, M. (2002). The effect of report cards on consumer choice in the health insurance market. Journal of Health Economics, 21(6), 1031–1048.

Yelp. (2018). Bringing Restaurant Hygiene Scores to Yelp Pages Everywhere. Retrieved from
https://blog.yelp.com/2018/07/lives-national

Yelp. (2018). Local Inspector Value-Entry Specification (LIVES). Retrieved from
https://www.yelp.com/healthscores