Can Monitoring and Citations Make Food Safer?

Jason Huang
Department of Economics
Stanford University
Stanford, CA 94305
jhuang99@stanford.edu

July 16, 2017

Abstract

Many local and state governments conduct regular health inspections on retail food establishments to ensure proper hygiene standards, and they hand out citations when the standards are not upheld. But how effective are these citations in incentivizing the restaurants to improve their sanitation practices? By analyzing food inspection data from New York City's Department of Health and Mental Hygiene, along with 311 complaint calls, I examine how restaurants respond to citations given out during regular food inspections. I find that more citations leads to better subsequent inspections. I also find that the establishments not only address the areas for which they received citations but also improve in other areas, suggesting complementarity across multiple dimensions of cleanliness. Furthermore, I find that consumers also respond to these improved sanitation conditions by making 311 complaint calls less frequently. To deal with the endogeneity of the inspection results, I exploit the random assignments of inspectors and construct inspector-specific measures of stringency as an instrumental variable. I find that this variable is highly predictive of not only the total scores but also the specific violations that are cited, despite the random assignment process that results in similar restaurant characteristics across inspectors.

1 Introduction

Food hygiene at restaurants is a great public health concern. The CDC estimates that over 3 million incidents of food-borne illness occur in the US annually, with 60% of the cases resulting from dining in restaurants.[1] Unlike food quality or service, which consumers can easily observe, food sanitation remains opaque to consumers since most diners cannot venture back to the kitchens and monitor the food preparation process. Local health departments have long conducted regular food inspections to ensure that restaurants meet certain hygiene standards.

This paper estimates the causal impact of NYC's food inspection results on restaurant compliances and customer perceptions, measured by frequencies of 311 complaints calls. Inspection results can potentially influence restaurant behaviors through for two reasons. First one is monetary. The scores determine the letter grades restaurants post on their entrance windows. These grades can influence consumer demand and affect revenues. The second one is behavior. restaurants may improve by learning from previous mistakes through the citations. Restaurants may commit violations not because they are profit maximizing by cutting enough corners to equate the marginal savings with the marginal costs from penalties. Instead, restaurants may be simply making mistakes because of lack of knowledge or experience.[2] Given that cleanliness is a multi-dimensional task, I examine how the distribution of citations impact the distributions of efforts across these dimensions. In addition, I test whether consumers behaviors respond to changes in restaurant hygiene efforts by measuring the changes in the frequencies of 311 complaint calls to DOHMH.

---

[1] Surveillance for Food borne Disease Outbreaks United States: 2013 Annual Report. Available: https://www.cdc.gov/foodsafety/pdfs/foodborne-disease-outbreaks-annual-report-2013-508c.pdf

[2] A potential reason is that scores are directly associated with monetary penalties. Each violation is associated with a fine, and the amount of fine depends on the severity cited by the inspector. However, the amounts of these fines are very small, ranging from 200 to 1000 per violation.

I find that more citations results in improved outcomes during subsequent inspection results. Receiving an one standard deviation worse overall sanitation score leads to an 0.23 standard deviation improvement in the overall score, a 9 percent increase in probability of earning an A, and 62 percent decrease in probability of being temporarily shutdown for unresolved critical violations during the subsequent inspection. While most of the improvement that a citation caused the restaurant to make is in the same area, citations in some areas led to improvements in other areas. For example, I find that an additional citation in Facility Design led to 0.195 fewer citation in Food Protection, in addition to 0.2 fewer citation in Facility Design. After merging the inspection data with the 311 complaint call data, I find that getting a worse inspection score also leads to fewer complain calls, suggesting that customers also notice the improved sanitation qualities.

Even though understanding how small businesses respond to citations is of interest to academics and policy makers, there is a dearth of studies on this topic. A major empirical challenge is that the variable of interest is endogenous. The inspection results depend on both the effects of monitoring and compliance, and we need to disentangle the two to estimate the causal effect of citations. For example, suppose we find that restaurants that had high numbers of citations tend to have high number of citations in the future. Do more citations lead to worse compliance or do restaurants that did poorly in the past continue to do poorly in the future? This paper uses the random assignment of food inspectors to inspections for restaurants in New York City as an instrumental variable to overcome the endogeneity of inspection results. As different inspectors have different levels of stringencies, the results that restaurants receive are partly determined by chance.[3]

A growing literature that complements this paper has focused on the role of transparency, rather than citation and citation, to improve food sanitation and reduce food-borne illness. The first type is certification, or what Ho (2012) calls targeted transparency. Letter grade system, first introduced in Los Angeles and later adopted by New York City is one example of certification. Jin and Leslie (2003) and Simon et al. (2005) study the introduction of the grading system in Los Angeles in 1998 and find strong evidence for a decrease in hospitalization for foodborne illnesses in jurisdictions after the enactments of the grading systems. Jin and Leslie (2003) argue that the reduction is driven by people cooking more at home and eating out less. Wong et al. (2015) and Meltzer et al. (2015) study New York City's food inspection program. They use inspection scores to measure restaurant cleanliness. After finding that improved sanitation scores and lower violation fines followed the reform, they conclude that the reform improved sanitation practices. These papers looks at impact of the overall inspection program while I study the effect of individual inspections. Filion and Powell (2009) document the growing popularities of these grading programs across the globe. The second type of transparency arise from customer reviews. Kang et al. (2013) use Yelp reviews to predict health code violations, and they argue that inspection selection can be optimized based on information from social media. [TODO: Say more about indirect]]However, sending food inspectors to restaurants for inspections remains as the most direct tool of enforcement, and understanding the impact of these inspections on compliance remains important. Hence, I see this study about monitoring and detection as a complement to instead of a substitute of studies about the role of transparency.

This paper also relates to a broad literature that studies the connection between government monitoring and compliance. Researchers have explored this issue in other areas such as taxation (Feinstein (1991), Kleven et al. (2010), Gemmell and Ratto (2012)), workplace safety (Levine and Johnson, 2012), nuclear power plants (Feinstein, 1989), and environmental protection (Duflo et al., 2016). While Jin and Lee (2014a) and Jin and Lee (2014b) also examines the food inspection programs, the two papers focus on the inspectors' behaviors and detection technology. Jin and Lee (2014b) provides suggestive reduced form evidence that improved detection technology lowers the probability of food borne disease outbreaks, it does not provide evidence that links inspector behaviors with restaurant compliance, which this paper does.

2   New York City Food Inspection System

New York City's DOHMH oversees the food safety program. It sends out food inspectors to restaurants to conduct inspections. Each restaurant receives at least one inspection a year. Starting in 2005, the city

---

[3]Other areas of research that have used this random assignments of decision makers to recover causal impacts of include the impact of incarceration on recidivism (Bhuller et al., 2016), the effect of disability insurance on labor supply (Maestas et al., 2013), and the impact of foster care system on child outcomes (Doyle, 2007).

implemented a numerical scoring system. It outlined ninety-eight violations, broken into eight categories. Each violation is associated with a range of scores, and the inspector assigns the numerical value that reflects the severity of the violation. For example, a violation 10F for not non-food contact surface improperly constructed is worth from 2 to 5 points, depending on how many of these surfaces the inspector finds.[4]

Starting in 2010, the sum of the scores were converted to letter grades, A, B, C. A score between 0 to 13 earns an A, between 14 to 27 earns a B, and 28 or above earns a C. In addition, the city required the restaurants to post these letter grades on the entrance so that they are clearly visible to the customers. The grading policy also changed the inspection cycles. It introduced dual inspection cycles, which involves the a random initial inspection and a possible re-inspection within a month if the initial inspection results in a letter grade lower than an A. During the time in between the initial inspection and re-inspection, restaurants can either post their assigned letter grade or a sign saying "Grade Pending." The score from re-inspection is then used to determine the letter grade the restaurant posts until the next initial inspection, and the previous initial inspection result has no effect.
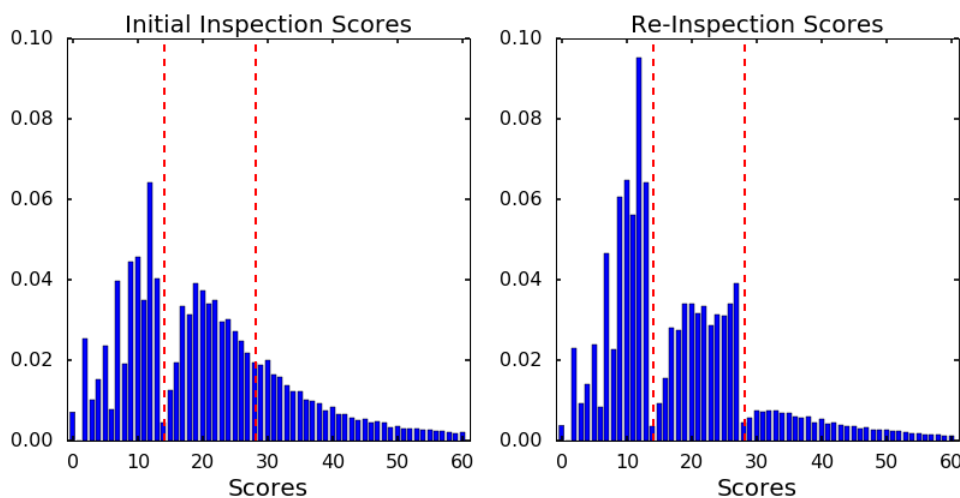
As a result of these letters, ... bunching

Figure 1: Score Distribution of Initial Inspections and Re-inspections.

DOHMH may temporarily close a restaurant for two reasons. The first reason is that the establishment poses a public health hazard that cannot be addressed by the end of the inspections. Such hazard includes issues such hot food not cooked to high enough temperature or cold food not stored at low enough temperature, presence of unpasteurized milk, or harmful noxious gas detected. The second reason is that restaurants receive scores above 28 for three consecutive inspections. If closed, DOHMH immediately posts a closure sign on the window. To reopen, the establishment must submit a written statement to DOHMH that describes the steps it has taken to correct the violations. Then re-opening inspection occurs. The median time between closure and re-opening inspection is 4 days, with 95% of the cases occurring within 24 days. The establishment can re-open if it passes the re-opening inspection, but it receives its next initial inspection within three months. Repeat violations may prompt DOHMH to shutdown an establishment for longer or permanently.

3 Empirical Analysis

3.1 Stringencies of Inspectors as Instrumental Variable

An ideal experimental design to estimate the causal effect of citations on restaurant compliance behavior is to randomly vary the number of citations each restaurant receives. However, inspection results are not

---

[4]See Food Service Establishment Inspection Procedures Handbook:
http://www1.nyc.gov/assets/doh/downloads/pdf/about/healthcode/health-code-chapter23.pdf.

random and are potentially correlated with variables, such as the restaurant's management quality or financial health, that are also correlated with restaurant's sanitation quality, regression estimates suffer from omitted variable bias. To address this issue, I exploit the two institutional features of NYC's food inspection program to create a natural experiment. First, according to the staff of DOHMH, inspectors are randomly assigned to restaurant for each inspection to ensure equity. They do not specialize in any particular geography or type of restaurant. Second, the levels of stringencies among the inspectors vary widely, with some inspectors being very strict while other being more lenient.

To empirically implement this natural experiment, I create an instrumental variable that reflect the level of stringency of each inspector that a restaurant faces during an inspection. Specifically, I define the instrument between restaurant i and inspector j as

$$Z_{ij} = \frac{1}{n_j - I_{ij}} \left(\sum_{kjtp \neq ijt} SCORE_{kjt} \right) \tag{1}$$

where n_j is the total number of initial inspections conducted by inspector j, I_{ij} is the total number of initial inspections inspector j has conducted at restaurant i, and $SCORE_{kjt}$ is the score inspector j gave to restaurant k on date t.

Figure 2 plots the distribution of the inspector average scores. The histogram shows a considerable amount of heterogeneity in terms of inspectors' levels of stringencies.[5] Figure 2 also shows the fitted value of locally weighted regression of the inspection scores on the inspectors' leave-out average scores.[6] The inspection scores and the inspector average scores follow a monotonic relationship, closely tracing along the 45° line. The plot suggests that being inspected by an inspector with a 1 point higher average score increases the expected inspection score by 1 point.
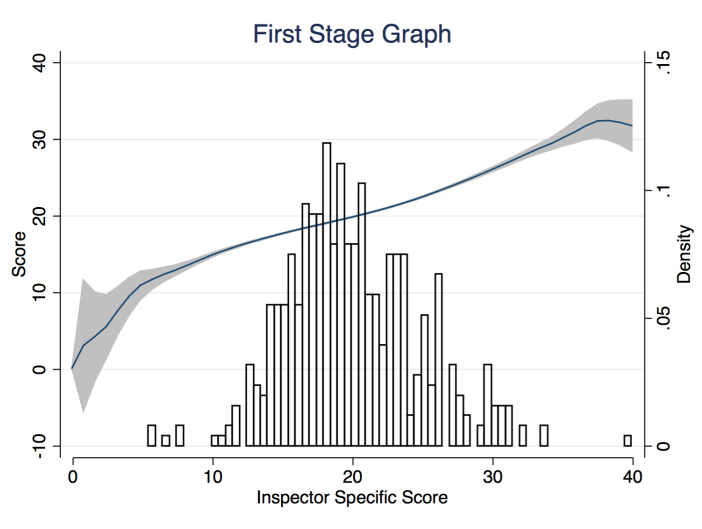

Figure 2

3.2 Impact of Inspection Scores on Subsequent Inspection Results

In this section, I examine the effects of the inspection results on subsequent initial inspection results. I consider three outcomes of health inspections as the dependent variables: the overall score, whether the restaurant receives an A, and whether that restaurant gets temporally shutdown for a critical violation that it cannot correct by the end of the inspection. While the subsequent initial inspection's overall score offers

---

[5]See Feinstein (1989), Feinstein (1991), and Macher et al. (2011) for additional documentation of inspector heterogeneity pertaining to IRS examiners, nuclear power plant inspectors, FDA regulators, respectively.

[6]I use a bandwidth of 0.5.

VARIABLES	(1) Score (OLS)	(2) Score (IV)	(3) Closure (OLS)	(4) Closure (IV)	(5) Grade A (OLS)	(6) Grade A (IV)
Score	-0.139***	-0.242***	-0.000763***	-0.00101***	-8.49e-05	0.00222***
	(0.00492)	(0.0162)	(8.26e-05)	(0.000185)	(0.000127)	(0.000556)
Observations	149,831	149,831	138,674	138,674	149,831	149,831
Inspection Date FE	YES	YES	YES	YES	YES	YES
Restaurant FE	YES	YES	YES	YES	YES	YES
dependent mean	21.08	21.08	0.0162	0.0162	0.372	0.372

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

Table 1: Impact of Inspection Scores on Temporary Closure and Obtaining Grade A

a summary of the number of violations cited and the severities of those violations, the two other variables should also be of interest. Given the bunching of scores around the 13-14 threshold, we do not expect the treatment effect to be constant across different score ranges. In addition, increasing the proportion of initial inspections that result in As can reduce administrative costs for DOHMH by reducing the resources devoted to re-inspections. We should also examine temporary shutdowns since these are associated with critical violations that may pose the biggest threat to public health.

To measure the effect of scores from initial inspections on subsequent initial inspection outcomes, I estimate the following equation:

$$Y_{i,t^{next}} = \beta Score_{it} + \delta_i + \tau_t + \tau_{t^{next}} + \varepsilon_{it}, \tag{2}$$

where $Score_{i,t}$ is the score earned by restaurant i on date t, δ_i and τ_t are the restaurant and date fixed effects, respectively, and t^{next} is the date of the subsequent initial inspection. The error term ε_{it} is two-way clustered at the inspector and zipcode levels. The term $Score_{it}$ is endogenous, so I instrument for it using the following first stage equation:

$$score_{it} = \gamma Z_{it} + \delta_i + \tau_t + \epsilon_{it} \tag{3}$$

where Z_{it} is defined in (1).

Table 1 reports the results. Given that the standard deviation of scores are around 14, the result suggests that one standard deviation increase in inspection scores increases the chance of getting an A during next initial inspection by about 3%, which is almost a 10% increase relative to the average 36%. And one standard deviation increase in score decreases the chance of getting closed [TODO: More discussion here.]
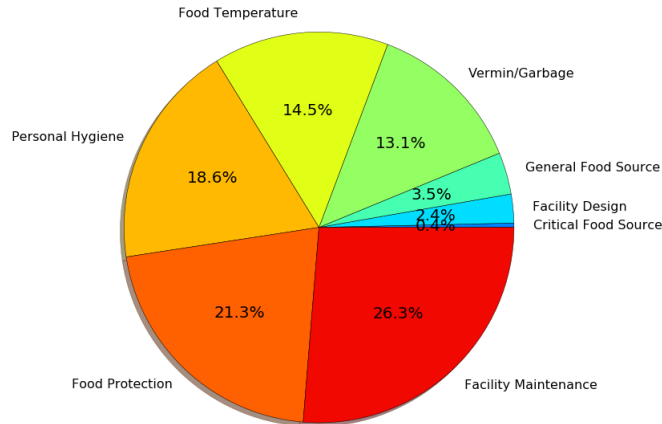
3.3   Multi-task

Figure 3: Breakdown of Frequencies of Violation Groups

So far, I have used the total score as an one dimensional summary of the inspection results. However, cleanliness is a multi-dimensional task. There are 82 violation codes, grouped into eight categories: food temperature, food source, food protection, facility design, personal hygiene, vermin, and facility maintenance. Figure 3 provides a visual breakdown of the frequencies by which these groups are cited. I measure the result of an inspection using an eight-dimensional binary vector. Each element of the vector represents one of the eight violation groups and encodes whether that inspection had at least one citation within that group.

I test how restaurants' effort allocation responds to the citation profile using the following regression.

$$Pr(Cited_{igt^{next}}) = \sum_{g' \in \mathcal{G}} \beta_{gg'} Cited_{igt} + \delta X_i + \tau_t + \varepsilon_{igt}, \tag{4}$$

where $Cited_{igt}$ is an indicator for whether restaurant i receives a citation of violation in group g on date t, τ_t is the date fixed effect and X_i is the restaurant controls. As before, $Cited$ is an endogenous, so need to instrument it using inspector propensity. The coefficient $\beta_{gg'}$ measures how a restaurant's effort in group g responds to a citation in group g'. To instrument eight endogenous variables, I construct each inspector's tendency to cite at least one violation in group g by using a leave-out strategy similar to the one done in section 3.1. More specifically, I construct the instrument of inspector j's tendency to cite group g at restaurant i as

$$Z_{ijg} = \frac{1}{n_j - n_{ij}} \left(\sum_{kjgt \neq ijgt} Cited_{ijgt} \right),$$

where n_j is the number of inspections made by inspector j and n_{ij} is the number of inspections inspector i has made at restaurant j. The first stage is a system of equations defined as

$$Pr(Cited_{ijgt}) = \sum_{g' \in \mathcal{G}} \theta_{gg'} Z_{ijg'} + \varepsilon_{ijgt}, \tag{5}$$

where $\mathcal{G}$ is the set of violation groups.

Table 2 presents the results of (4) instrumenting with (5). We see that the diagonal coefficients are all negative, and six of the eight groups are statistically significant at 90% level. I find that restaurants improve the most in the areas of Facility Maintenance in response to citations in the same area, with one more citation leading to a 0.236 fewer expected citations in the subsequent inspection.

VARIABLES	(1) Facility Maintenance	(2) Food Protection	(3) Personal Hygiene	(4) Food Temperature	(5) Vermin/Garbage	(6) Gen. Food Source	(7) Facility Design	(8) Crit. Food Source
Facility Maintenance	-0.236***	0.00733	0.000184	-0.00506	-0.000922	-0.00159	0.000981	-0.000509
	(0.00961)	(0.0119)	(0.0107)	(0.00944)	(0.00766)	(0.00380)	(0.00295)	(0.00140)
Food Protection	-0.0776**	-0.177***	-0.00731	-0.0142	-0.0397*	-0.0100	0.00372	0.00780*
	(0.0378)	(0.0306)	(0.0410)	(0.0253)	(0.0227)	(0.0113)	(0.0108)	(0.00400)
Personal Hygiene	0.0222	-0.00484	-0.189***	-0.00771	-0.000551	0.00458	0.00129	0.00217
	(0.0156)	(0.0103)	(0.0148)	(0.00957)	(0.00892)	(0.00489)	(0.00320)	(0.00173)
Food Temperature	0.0148	-0.0261	0.0133	-0.214***	-0.00923	0.00613	-0.00622	-0.00153
	(0.0216)	(0.0163)	(0.0162)	(0.0134)	(0.0128)	(0.00722)	(0.00573)	(0.00243)
Vermin/Garbage	0.132***	-0.0784*	0.00134	0.0446	-0.174***	0.0128	-0.0105	-0.00126
	(0.0439)	(0.0404)	(0.0544)	(0.0318)	(0.0283)	(0.0159)	(0.0138)	(0.00404)
Gen. Food Source	-0.0114	-0.00139	-0.0456	-0.0103	-0.0395*	-0.191***	-0.0147	-0.000895
	(0.0291)	(0.0348)	(0.0374)	(0.0334)	(0.0202)	(0.0199)	(0.00933)	(0.00395)
Facility Design	0.0265	-0.195**	-0.00742	0.0168	-0.0235	0.00421	-0.199***	0.0169*
	(0.0787)	(0.0925)	(0.0776)	(0.0551)	(0.0612)	(0.0296)	(0.0239)	(0.00977)
Crit. Food Source	0.274	-0.284	0.279	0.125	-0.0346	-0.0233	-0.0643	-0.185***
	(0.248)	(0.207)	(0.244)	(0.195)	(0.142)	(0.0848)	(0.0682)	(0.0298)
Observations	149,831	149,829	149,831	149,829	149,829	149,831	149,829	149,831
Dependent mean	0.989	0.802	0.842	0.645	0.503	0.125	0.0694	0.0126

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

Table 2

The signs of the diagonals shed insight on the production function of restaurant cleanliness. We see that many off-diagonal elements are negative and statistically significant, which suggests that a citation in one dimension leads to an improvement in another dimension. This finding makes sense in light of the fact that various dimensions may be complementary. One example of such complementarity is that a citation in Food Protection also reduces the expected number of citations from subsequent inspections in Facility Maintenance and Vermin/Garbage by -0.0776 and 0.0397, respectively. A citation in Food Protection may

lead the restaurant to improve along this dimension, which might include fixing a broken thermometer (violation 10E) that improves the Facility Maintenance dimension or better maintaining contact surface (violation 10F) that makes the kitchen more vermin proof.

3.4 Impact of Inspection Results on 311 Complaint Calls

The ultimate goal of any food inspection programs is to reduce food-borne illness, but directly studying food poisoning is difficult. This outcome variable suffers from under-reporting, and more severe cases that result in hospitalization, the origin of the illness cannot be traced back to a specific establishment. Food inspection results serve as a convenient proxy for restaurant cleanliness. A couple studies have found evidence of association between inspection results and the likelihood of food-borne illness outbreaks.[7] In this section, I use 311 complaint calls of restaurants to supplement the findings in my previous sections.

I transform the inspection level data to a restaurant-month level data where each observation represents a month for each restaurant between the first and last recorded inspection. For each month, I calculate the inspection score as the score from the most recent inspection. I estimate the following equation.

$$Pr(called_{it}) = \delta_i + \tau_t + \beta_0 Score_{it} + \beta_1 Month\_Since\_Inspection_{it}$$
$$+\beta_2 Month\_Since\_Inspection_{it} \times Score_{it} + \varepsilon_{it} \tag{6}$$

where $called_{it}$ a binary variable that equals one if DOHMH received at least one complaint call for restaurant i in month t. For the IV regression, I again instrument the score using equation (3).

VARIABLES	(1) Prob Call (OLS)	(2) Prob Call (OLS)	(3) Prob Call (OLS)	(4) Prob Call (IV)	(5) Prob Call (IV)	(6) Prob Call (IV)
SCORE	-4.67e-05***	-6.95e-06	-9.01e-05***	-8.98e-05*	-9.18e-05*	-6.21e-05
	(1.53e-05)	(1.57e-05)	(1.78e-05)	(4.77e-05)	(4.99e-05)	(6.17e-05)
Months Since Inspection		0.000540***	-1.81e-05		0.000498***	0.000735*
		(3.72e-05)	(6.61e-05)		(4.33e-05)	(0.000435)
c.mon_from_inspect#c.SCORE			5.02e-05***			-2.16e-05
			(5.80e-06)			(3.98e-05)
Observations	1,223,207	1,223,207	1,223,207	1,223,207	1,223,207	1,223,207
Year-Month FE	YES	YES	YES	YES	YES	YES
Restaurant FE	YES	YES	YES	YES	YES	YES
Dependent Mean	0.0145	0.0145	0.0145	0.0145	0.0145	0.0145

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

Table 3: Impact of Inspection Scores on the Probability of Receiving 311 Complaint Call

Table 3 presents the result of (6). We see that getting a one standard deviation worse inspection score lowers the probability of a 311 call in a month by around 0.01%, which is almost a 10% reduction relative to the sample average of 1.45%.

4 Conclusion and Discussion

This paper combines multiple administrative level data from New York City to study how restaurants hygiene practices respond to citations given out during health inspections. While studying the causal link between citations and later behaviors is difficult when the citations are not randomly assigned, this paper overcomes this issue by creating a quasi experiment that uses the random assignment of inspectors with differing levels of stringency as an instrumental variable. I find evidence that restaurants respond to citations of health code violations by improving their sanitation qualities, measured by improved results during subsequent inspections and reduced frequencies of customer complaint calls.

---

[7]See Irwin et al. (1989) and Petran et al. (2012a) and Petran et al. (2012b).

References

Bhuller, M., Dahl, G. B., Lken, K. V., and Mogstad, M. (2016). Incarceration, Recidivism and Employment . *Mimeo*.

Doyle, J. (2007). Child Protection and Child Outcomes: Measuring the Effects of Foster Care. *American Economic Review*, 97(5):1583–1610.

Duflo, E., , Greenstone, M., Pande, R., and Ryan, N. (2016). The Value of Regulatory Discretion: Estimates from Environmental Inspections in India. *NBER Working Paper Series*.

Feinstein, J. (1989). The Safety Regulation of U.S. Nuclear Power Plants: Violations, Inspections, and Abnormal Occurrences. *Journal of Political Economy*, 97(1):115 – 154.

Feinstein, J. (1991). An Econometric Analysis of Income Tax Evasion and Its Detection. *The RAND Journal of Economics*, 22(1).

Filion, K. and Powell, D. A. (2009). The Use of Restaurant Inspection Disclosure Systems as a Means of Communicating Food Safety Information. *Journal of Food Service*, 20(6):287–297.

Gemmell, N. and Ratto, M. (2012). Behavioral Responses to Taxpayer Audits: Evidence from Random Taxpayer Inquiries. *National Tax Journal*, 65(1):33–58.

Ho, D. H. (2012). Fudging the Nudge: Information Disclosure and Restaurant Grading. *The Yale Law Journal*, 112(3):522 – 851.

Irwin, K., Ballard, J., Grendon, J., and Kobayashi, J. (1989). Results of Routine Restaurant Inspections Can Predict Outbreaks of Foodborne Illness: The Seattle-King County Experience. *American Journal of Public Health*, 79(5):586–590.

Jin, G. Z. and Lee, J. (2014a). A Tale of Repetition: Lessons from Florida Restaurant Inspections. *Mimeo*.

Jin, G. Z. and Lee, J. (2014b). Inspection Technology, Detection, and Compliance: Evidence from Florida Restaurant Inspections. *The RAND Journal of Economics*, 45(4):885917.

Jin, G. Z. and 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.

Kang, J. S., Kuznetsova, P., Choi, Y., and Luca, M. (2013). Where Not to Eat? Improving Public Policy by Predicting Hygiene Inspections Using Online Reviews. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2010). Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark. *Econometrica*, 79(2):651–692.

Levine, D. I. . M. W. T. and Johnson, M. S. (2012). Randomized Government Safety Inspections Reduce Worker Injuries with No Detectable Job Loss. *Science*, 336(6038):907–911.

Macher, J. T., Mayo, J. W., and Nickerson, J. A. (2011). Regulator Heterogeneity and Information Asymmetry. *The Journal of Law and Economics*, 54:25 – 54.

Maestas, N., Mullen, K. J., and Strand, A. (2013). Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt. *American Economics Review*, 103(5):1797–1829.

Meltzer, R., Calabrese, T., Silver, D., and Mijanovich, T. (2015). Is Public Grading Worth the Costs? An Evaluation of New York Citys Restaurant Grades Policy. *Mimeo*.

Petran, R. L., White, B. W., and Hedberg, C. W. (2012a). Health department inspection criteria more likely to be associated with outbreak restaurants in minnesota. *Journal of Food Protection*, 75(11):2007–2015.

Petran, R. L., White, B. W., and Hedberg, C. W. (2012b). Using a Theoretical Predictive Tool for the Analysis of Recent Health Department Inspections at Outbreak Restaurants and Relation of this Information to Foodborne Illness Likelihood. *Journal of Food Protection*, 75(11):2016–2027.

Simon, P. A., Leslie, P., and Run, G. (2005). Impact of Restaurant Hygiene Grade Cards on Foodborn-Disease Hospitalization. *Journal of Environmental Health*, 67(7):32– 36.

Wong, M. R., McKelvey, W., Ito, K., and Jacobson, J. B. (2015). Impact of a letter-grade program on restaurant sanitary conditions and diner behavior in New York City. *American Journal of Public Health*, 105(3):81–87.