



Munich Personal RePEc Archive

# **Mandatory Disclosure, Letter-Grade Systems, and Corruption: The Case of Los Angeles County Restaurant Inspections**

Makofske, Matthew

21 August 2017

Online at <https://mpra.ub.uni-muenchen.de/80925/>  
MPRA Paper No. 80925, posted 21 Aug 2017 22:14 UTC

# MANDATORY DISCLOSURE, LETTER-GRADE SYSTEMS, AND CORRUPTION: THE CASE OF LOS ANGELES COUNTY RESTAURANT INSPECTIONS

Matthew Philip Makofske\*

August 21, 2017

## Abstract

In 1998, Los Angeles (LA) County adopted a mandatory disclosure policy aimed at inducing restaurant hygiene improvements. LA County restaurants receive numeric scores during unannounced hygiene inspections and then post letter grades in their windows based on broad intervals to which their inspection scores belong. This letter-grade system generates: relatively weak incentives for hygiene improvement at letter-grade thresholds, and relatively strong incentives for score manipulation below those thresholds. Using over 140,000 LA County restaurant inspections spanning October 2014 to September 2016, I test for manipulation by exploiting a feature of the county's scoring criteria. The violation of most health codes carries a prescribed 1, 2, or 4-point deduction. However, there are eleven health code violations where, depending on severity, 2 or 4 points may be deducted. Even when compared with inspections exhibiting better overall hygiene quality, restaurants on the margin of a higher letter grade are 28-40% more likely to receive the lesser point deduction on these violations. Restaurants on the margin are significantly more likely to receive the lesser deduction across all eleven violation types. That, and other characteristics of the data, suggest that these results do not reflect restaurants electively bunching at letter-grade thresholds. I find that scores were manipulated to improve letter grades in as many as 5,921 inspections (4.2% of the sample, and 26.56% of inspections where scoring decisions had letter-grade implications).

**JEL:** D82, L15, I18, K32

**Keywords:** mandatory disclosure, product quality, manipulation, restaurant hygiene

---

\*Department of Economics, Miami University. 2054 Farmer School of Business, 800 E. High St., Oxford, OH 45056. Email: makofsmp@miamioh.edu. Phone: (513) 529-3066. I thank Carl Kitchens for many helpful comments. Any remaining errors are mine.

# 1 Introduction

Despite being largely preventable, the Centers for Disease Control and Prevention (CDC) estimated in 2011 that 48 million Americans (about 1 in 6) contract foodborne illness each year, resulting in 128,000 hospitalizations and 3,000 deaths. The CDC also estimated that, in 2013, restaurants accounted for 60 percent of the foodborne illness outbreaks in the US with a single known food-preparation source.<sup>1</sup> In an effort to prevent foodborne illnesses, many governments have mandated disclosure of hygiene quality information by restaurants.

Mandatory disclosure has become a popular regulatory tool in industries where consumers and producers have asymmetric information over product quality. If consumers prefer higher product quality *cet. par.*, mandatory disclosure of quality information should result in consumers substituting toward higher quality producers. Anticipating this substitution producers should, on average, improve product quality. Lending to its popularity, mandatory disclosure has been shown to cause quality improvements in drinking water, restaurant hygiene, and schools.<sup>2</sup>

However, the recent history of disclosure policies is hardly replete with success stories. In some instances, quality improvements are not apparent; when product quality is multi-dimensional, producers have been found to improve quality along disclosed dimensions and reduce it along unreported dimensions; and in other cases, the heightened stakes brought on by mandatory disclosure have led to manipulative, “gaming”, or even corrupt behavior.<sup>3</sup> Across governments, there can be considerable variety in the design and implementation of mandatory disclosure policies even within the same industry. With the effects of introducing mandatory disclosure having been studied in several settings, attention is turning to how a

---

<sup>1</sup>See <http://www.cdc.gov/foodsafety/foodborne-germs.html> for 2011 estimates. See Centers for Disease Control and Prevention (2013), or <http://www.cdc.gov/features/foodborne-diseases-data/> for 2013 estimates.

<sup>2</sup>See Benneer and Olmstead (2008) on drinking water, Jin and Leslie (2003) on restaurant hygiene, and Hanushek and Raymond (2004) and Jacob (2005) on schools. An extensive review of the quality disclosure literature is found in Dranove and Jin (2010).

<sup>3</sup>See Lu (2012) regarding multidimensional quality. See Dranove et al. (2003), and Figlio and Getzler (2006) regarding disclosure and gaming, and Jacob and Levitt (2003) regarding disclosure and corruption.

disclosure policy’s effectiveness relates to specific details of its design or implementation.<sup>4</sup>

In this paper, I show that one of the success stories of mandatory disclosure, the 1998 requirement that Los Angeles County restaurants post hygiene grade cards in their windows,<sup>5</sup> is still much less effective than it could be due to a design feature: the requirement that restaurants post a letter grade corresponding to their most recent inspection score, rather than the score itself. The letter-grade system discloses less information than is available, and makes the inspection process ripe for corruption because the benefit to restaurants of score manipulation sharply increases below letter-grade score thresholds.

Disclosure policies typically require firms to report some signal of product quality (*e.g.*, water contaminant levels, restaurant health inspection results, or school performance on a standardized test). Producer responses will be limited if the signal has little informational content or is disclosed in a manner such that few consumers will notice.<sup>6</sup> Moreover, if the signal is manipulable, producers may choose to invest in manipulation rather than actual quality improvement. In response to school accountability measures, Figlio and Getzler (2006) find evidence that Florida schools manipulated their test pools by classifying additional low-performing students as disabled, and Jacob and Levitt (2003) find evidence of outright cheating by teachers and administrators in some Chicago schools.

In the cases documented by Jacob and Levitt (2003), Figlio and Getzler (2006), and Forbes et al. (2015), producers were able to practice manipulation independently. When disclosed information is collected by inspection, manipulation likely requires the inspector’s involvement. While this need for complicity creates an initial hurdle to manipulation, corruption could be quite pervasive if that hurdle is cleared. After all, it is the inspectors who

---

<sup>4</sup>Notably, Forbes et al. (2015) assess mandatory disclosure by US airlines of the fraction of their flights which are more than 15 minutes late. They find that some airlines respond strongly, shifting attention toward flights expected to be around 15 minutes late, and some do not respond at all. Among flights expected to be 15-16 minutes late, they find that airlines using manual (as opposed to automatic) arrival-time reporting were 20 to 120 percent more likely to have flights arrive earlier than expected, suggesting misreporting.

<sup>5</sup>Jin and Leslie (2003) show that this resulted in substantial restaurant hygiene improvements.

<sup>6</sup>In Louisville, which adopted grade-card disclosure in 1996, and where restaurant inspection scores have been available on the city’s website since 2009, Makofske (2017) finds that the posting of inspection scores on Yelp.com (a consumer-review forum) in 2013, resulted in significant hygiene improvements among independent restaurants.

collect the quality measures and often see to the policy’s enforcement. Moreover, inspectors who practice manipulation may use institutional knowledge of oversight procedures to avoid detection by supervisors. This makes it especially important to understand if seemingly minor features of disclosure policies may compromise regulatory inspections.

To better understand how disclosure policy designs may unintentionally promote corruption, this paper studies restaurant health inspections from Los Angeles (LA) County, California. During health inspections, LA County restaurants receive a numeric score out of 100 possible points. While scores are measured in one-point increments, restaurants post letter grades based on ten-point intervals in their windows.<sup>7</sup> Letter-grade disclosure distorts restaurant incentives on both sides of the 70, 80, and 90-point thresholds. Compared to numeric-score disclosure, it provides relatively weak incentives for hygiene improvement at (and just above) letter-grade thresholds, and relatively strong incentives for manipulation just below those thresholds.

Examining all routine LA County restaurant inspections involving violations from October 1, 2014 to September 30, 2016, I find disproportionate masses of scores at letter-grade cutoffs, especially at 90 points (the modal score in the sample). While dubious, this is not necessarily evidence of manipulation.<sup>8</sup> I test for evidence of manipulation by exploiting an institutional feature of the LA County scoring criteria. Most health code violations carry a single prescribed deduction of 1, 2, or 4 points. However, there are eleven violations which allow for a 2 or 4-point deduction based on the observed severity of the infraction. I refer to these as *discretionary* violations because, once detected, the inspector must discern which deduction is appropriate. Inspections involving discretionary violations provide a very useful margin for evaluating inspector behavior: given the set of violations detected, point-deduction assessments on discretionary violations either will, or will not, have letter-grade implications (for ease of exposition, I describe a restaurant as *on the margin* if those as-

---

<sup>7</sup>The county has also shared this information with Yelp.com since December 2013. A restaurant’s most recent letter grade is posted at the top of their Yelp profile page.

<sup>8</sup>In light of the letter-grade system, some restaurants may choose the least costly provision of hygiene quality sufficient for a particular grade.

assessments will have letter-grade implications). Moreover, because point deductions are only flexible for this subset of violations, across inspections exhibiting very similar hygiene quality, I observe restaurants that are on the margin, as well as restaurants that are not. A testable hypothesis follows: all else equal, in the absence of manipulation, a restaurant on the margin should be no more likely to receive the lesser deduction on discretionary violations.<sup>9</sup>

Using discretionary violations, I compare point-deduction assessments in inspections where restaurants are on the margin, with deductions in inspections where restaurants exhibiting slightly better hygiene quality are not on the margin (if anything, cleaner restaurants should be more likely to warrant the lesser deduction). Controlling for a variety of other inspection, restaurant, and violation-specific features, I find that restaurants are anywhere from 28 to 40 percent more likely to receive the lesser deduction when they are on the margin of a higher letter grade.

Several characteristics of the data suggest that these results reflect manipulation rather than the behavior of restaurants producing *A* grades at minimum cost. For one, being on the margin increases the probability of the lesser deduction by statistically significant and substantial amounts across all eleven discretionary violation types. This includes the violation of codes which bear little if any cost to obey, and could be easily hidden during inspections (*e.g.*, proper hand washing). Also, the inspection scores that I flag as likely manipulated are concentrated within particular areas of the county, suggesting the involvement of a common set of inspectors.<sup>10</sup> Additionally, restaurants on the margin of *B* grades are significantly and substantially more likely to receive the lesser deduction, and it is far less plausible that restaurants made *B* grades on the margin by design. Finally, of the 13,656 restaurants that received *A* grades on the margin, 67.68 percent did so only once in the sample. If these restaurants were truly optimizing, they should be capable of producing similar inspection

---

<sup>9</sup>Technically, in the absence of manipulation *via* favorable deduction decisions (*i.e.*, assessing a 2-point deduction when the 4-point deduction is warranted). Under the null, manipulation could still be practiced by not reporting detected violations.

<sup>10</sup>There are even twelve instances where four likely manipulated inspection scores occurred within the same zip code in a single day. The maximum distance between any two of the four restaurants involved is never more than 2.70 miles, and is less than one mile in four cases.

performances multiple times.<sup>11</sup> Moreover, even after excluding restaurants that made multiple *A* grades on the margin, the remaining sample suggests that restaurants on the margin are still 25.36 percent more likely to receive the lesser deduction.

My results highlight considerable drawbacks of disclosing broad quality measure intervals, rather than the underlying measure itself. It limits the extent to which consumers are informed, and more importantly, it distorts producer incentives in a manner conducive to manipulation. In my sample, I find that as many as 5,921 inspection scores were likely manipulated to improve letter grades.<sup>12</sup> Also, note that these results may actually understate the extent of such activity as they focus on only one possible mode of score manipulation; they don't address potential non-reporting of detected violations. As governments consider the adoption and revision of such policies, careful attention should be paid to details of the signal being disclosed. Reporting letter grades rather than underlying scores may seem like a trivial distinction (because either will substantially increase information provision relative to a state of no disclosure), but it can substantially limit disclosure's ability to induce product quality improvements, and can even compromise the integrity of regulatory inspections. Given that foodborne illness is a persistent yet preventable affliction in the US, these implications seem especially significant in the context of restaurant hygiene.

In the space remaining, I review the restaurant inspection process and scoring criteria used in LA County, as well as the data used in this paper. This is followed by a discussion of my empirical methodology and a presentation of my main results. I then assess the spatial distribution of likely manipulated inspection scores, and conclude with a series of tests which largely rule out elective bunching by restaurants as a plausible explanation of my results.

---

<sup>11</sup>Also, the restaurants that received *A* grades on the margin more than once seldom committed the same type(s) of discretionary violation(s) across these inspections.

<sup>12</sup>This amounts to: 4.2 percent of all the inspections in my sample, 8.19 percent of inspections involving discretionary violations, and 26.56 percent of all inspections where restaurants were on the margin (where manipulation *via* favorable deductions was possible).

## 2 Los Angeles County Restaurant Inspections

The Los Angeles County Department of Public Health (DPH) conducts unannounced restaurant hygiene inspections. Restaurants receive a numeric score and are issued a placard to post in their window until their next routine inspection. Restaurants receiving a score below 70 must post their numeric score, and restaurants scoring at or above 70 points post a letter grade based on the scale shown in Table 1.

Table 1: Los Angeles County: Restaurant Hygiene Grade Scale

Inspection Score	Grade Card Display
Score $\geq 90$	<i>A</i>
$90 > \text{Score} \geq 80$	<i>B</i>
$80 > \text{Score} \geq 70$	<i>C</i>
$70 > \text{Score}$	Score

The LA County DPH inspects all restaurants within unincorporated areas of the county, and 85 of the 88 municipalities in the county.<sup>13</sup> County Ordinance 97-0071 requires all restaurants in unincorporated areas of LA County, and in municipalities which have adopted the ordinance, to post their grade card until their next routine inspection. To date, all but seven municipalities under the jurisdiction of the LA County DPH have adopted the ordinance.<sup>14</sup> Throughout my sample however, the county has shared all inspections results with Yelp.com, and Yelp posts the most recent letter grade on the profiles of all restaurants regardless of whether or not their municipality has adopted the ordinance. Thus, restaurants in those seven municipalities likely face similar incentives for score manipulation as other restaurants in the county.

Violations of the LA County health code are categorized as either critical or non-critical, and critical violations are further categorized as major or minor.<sup>15</sup> The DPH classifies some critical violations as *either* major or minor, but there are eleven critical violations for

<sup>13</sup>Pasadena, Long Beach, and Vernon have their own health departments and codes.

<sup>14</sup>The municipalities which have not yet adopted the grade-posting ordinance are: Avalon, Bradbury, Hidden Hills, La Habra Heights, San Marino, Sierra Madre, and Signal Hill.

<sup>15</sup>Minor critical violations receive 2-point deductions, and major critical violations carry a 4-point penalty.



which the DPH distinguishes a minor and major violation of the health code (the so-called *discretionary* violations). The DPH specifies what constitutes a minor or major violation of these health codes, and the distinction is typically clear.<sup>16</sup> For example, the temperature at which hot or cold food are held, the concentration of cleaner used on surfaces, specific behavior of employees, *et cetera*. So in principle, discretion needn't be exercised in scoring these violations.

Finally, note that LA County reassessed its scoring criteria during this sample period. Following a review, the county elected to keep their letter-grade system, but adopt a scoring change: restaurants receiving two major critical violations would incur an additional 3-point deduction. Restaurants forced to shut down incur an additional 7-point penalty.<sup>17</sup> These changes don't apply to this paper's sample however, as they were not implemented until January 1, 2017.<sup>18</sup>

### 3 Data Description

The data used in this paper come from the open data portal of Los Angeles County.<sup>19</sup> Each observation corresponds to a violation and includes: the date of the inspection, the name and address of the restaurant, the health code pertaining to the violation, the number of points deducted for the violation, and the restaurant's ultimate score and letter grade on the inspection.

I aggregate point deductions within inspections, subtract from 100, and identify 239 inspections where the computed score does not match the reported score. Because it is unknown whether detected violations were mistakenly omitted from the dataset, or whether the inspection scores were entered incorrectly, I drop all observations from these inspections.

---

<sup>16</sup>An explanation of the inspection process and health codes is found at <http://publichealth.lacounty.gov/eh/docs/RefGuideFoodInspectionReport.pdf>.

<sup>17</sup>The letter sent to restaurants explaining the scoring change is found at <http://publichealth.lacounty.gov/eh/docs/LettertoFoodIndustry.pdf>.

<sup>18</sup>Note also, that the revisions do nothing to affect the incentive distortions around letter grade thresholds.

<sup>19</sup>See <https://data.lacounty.gov/>.

This leaves a final sample of 140,163 inspections conducted on 44,516 different restaurants from October 1, 2014 to September 30, 2016. I also identified and fixed the entries for 119 observations where the entered zip codes were incorrect.<sup>20</sup> Summary statistics are presented in Table 3.

The distribution of all inspection scores in the final sample is shown in Figure 1. Notice that there is an increase in relative frequency going from inspection scores of 79 to 80 points. Notice especially that relative frequency is decreasing in score over the interval from 86 to 89 points. This is followed by a sharp and substantial increase at 90 points, and then a sharp and immediate decrease at 91 points. There is a clearly disproportionate mass of scores at 90 points, which is actually the modal score in the sample.

Figure 2 plots separate score distributions based on whether or not discretionary violations were reported. Navy dots mark the relative frequency of each score among the 67,835 inspections in which no discretionary violations were reported. Orange triangles mark the relative frequency of each score among the 72,328 inspections in which one or more discretionary violations were reported. When some point deductions are flexible to the inspector, 22.42 percent of inspections result in scores of 90 points. The second most frequent score among this subgroup is 91 points, which occurred in 12.78 percent of inspections. Thus, among the 72,328 inspections in which one or more discretionary violations were reported, about 35.21 percent resulted in scores at, or one point above, the 90-point cutoff.<sup>21</sup>

The bunching observed in Figures 1 and 2, while dubious, is not necessarily evidence of score manipulation. In light of the letter-grade system, restaurants may choose to provide the least costly level of hygiene quality needed to receive an *A*, which would also produce a disproportionate mass of inspection scores at 90 points. To determine if manipulation occurred, I test whether scoring decisions were influenced by letter-grade implications.

---

<sup>20</sup>I identified these errors by looking up addresses in reported zip codes which had abnormally few inspections in the sample. In some cases, zip codes were incorrect due to transposition of digits. There were also several zip codes in LA County which, during the sample period, were changed to “P.O. box only” zip codes. Earlier entries pertaining these addresses were updated to match the current zip code.

<sup>21</sup>The relevance of being at, or one point above the letter-grade threshold, is that the point deduction assessment on a discretionary violation makes a 2-point difference in the inspection score.

## 4 Empirical Strategy

### 4.1 Detecting Evidence of Score Manipulation

My empirical approach exploits the fact that when discretionary violations are committed, conditional on the entire set of detected violations, the inspector’s point-deduction assessments either will, or will not, determine the letter grade the restaurant receives. If manipulation does not occur, point-deduction assessments should be independent of whether or not a restaurant is on the margin of a higher letter grade.

Let  $Score_i$  denote the numeric score given in inspection  $i$  of the sample. This score depends on the set of violations detected, and the point deductions that they carry. If any discretionary violations are detected, then multiple possible outcomes exist for  $Score_i$ . Given the set of detected violations, I denote the lowest possible score for inspection  $i$  as  $ScoreMin_i$ , and  $ScoreMax_i$  denotes the highest possible score.<sup>22</sup> A restaurant is on the margin in inspection  $i$  if  $ScoreMin_i$  and  $ScoreMax_i$  belong to different letter-grade intervals.

Manipulation *via* favorable deduction decisions occurs if an inspector should deduct 4 points for an observed discretionary violation, but assesses the 2-point deduction instead. I am unable to directly observe this, as appropriate deductions are known only to the inspector who detects the violations and perhaps the restaurant manager. However, indirect evidence of such behavior can be found by recognizing that manipulation is, in effect, only beneficial to restaurants if it produces a higher letter grade. Thus, I test for manipulation by estimating the probability that the lesser point deduction is assigned on discretionary violations, conditional on a variety of inspection and violation-specific features.

Each inspection,  $i$ , is a set of detected violations, which are indexed by  $j$ . To estimate the probability of the lesser (2-point) deduction, I specify the following linear model:

$$Y_{j(i)} = \beta_1 Margin_i + \mathbf{X}_i' \boldsymbol{\beta} + \mathbf{Z}_j' \boldsymbol{\gamma} + \epsilon_{j(i)}. \quad (1)$$

---

<sup>22</sup> $ScoreMin_i$  is the inspection score if 4-point deductions are assessed on all discretionary violations, and  $ScoreMax_i$  is the inspection score if 2-point deductions are assessed on all discretionary violations.

In inspection  $i$ , conditional on violation  $j$  being a discretionary violation,  $Y_{j(i)}$  is an indicator equal to 1 if the 2-point deduction is assessed, and equal to 0 if the 4-point deduction is assessed.  $Margin_i$  is an indicator equal to 1 if the restaurant is on the margin of a higher letter grade in inspection  $i$ , and equal to 0 otherwise. The violation-specific vector  $\mathbf{Z}_j$  contains fixed effects for the health code that  $j$  violates. The vector  $\mathbf{X}_i$  contains several inspection-specific controls including indicators for the month, year, day-of-week, and zip code in which the inspection occurred. Most important however are controls related to overall inspection performance, as these controls account for the hygiene quality of the restaurant being inspected, and thereby establish appropriate comparison groups for the inspections on the margin.

## 4.2 Comparison Group Construction

Presumably, restaurants exhibiting poorer overall hygiene quality are more likely to warrant the 4-point deduction on discretionary violations. This makes it important to compare deduction outcomes from inspections on the margin, with outcomes from inspections that are not on the margin and which exhibit equivalent or better hygiene quality. Ideally, comparisons would be made across inspections of equivalent hygiene quality and thereby capture the effect on scoring decisions of *Margin* only, but that is not feasible. If two inspections involve a discretionary violation (where deduction decisions may be observed and compared), and exhibit equivalent hygiene quality over all other detected violations, then both inspections will either be on the margin of a higher letter grade, or not on the margin.<sup>23</sup>

Instead, I compare deduction decisions in inspections on the margin, with decisions in inspections that are not on the margin, but which exhibit better overall hygiene quality. To enable these comparisons I construct a variable,  $Group_i$ , as defined in Table 2. Notice that within each group, the inspections which are on the margin exhibit lower overall hygiene quality (as measured by  $ScoreMin_i$ ) than the inspections which are not on the margin. I

---

<sup>23</sup>Section A1.1 of the Appendix illustrates why this is the case.

estimate equation (1) with fixed effects for each *Group* outcome included.<sup>24</sup> Because cleaner restaurants are more likely to warrant the lesser deduction, this approach errs on the side of potentially understating the effect of *Margin* on point-deduction decisions.

Table 2: Comparison Groups

Group Variable	If $Margin_i = 1$	If $Margin_i = 0$
$Group_i = 1$	$ScoreMin_i \in [86, 89]$ ( $N_i = 19,567, N_j = 33,517$ )	$ScoreMin_i \geq 90$ ( $N_i = 45,418, N_j = 49,217$ )
$Group_i = 2$	$ScoreMin_i \in [82, 85]$ ( $N_i = 947, N_j = 2,925$ )	$ScoreMin_i \in [86, 89]$ ( $N_i = 384, N_j = 384$ )
$Group_i = 3$	$ScoreMin_i \in [78, 81]$ ( $N_i = 963, N_j = 2,895$ )	$ScoreMin_i \in [82, 85]$ ( $N_i = 2,808, N_j = 5,524$ )
$Group_i = 4$	$ScoreMin_i \in [72, 77]$ ( $N_i = 727, N_j = 2,659$ )	$ScoreMin_i \in [78, 81]$ ( $N_i = 1,329, N_j = 3,390$ )
$Group_i = 5$	$ScoreMin_i \leq 71$ ( $N_i = 89, N_j = 430$ )	$ScoreMin_i \in [72, 77]$ ( $N_i = 52, N_j = 136$ )

Inspections are assigned to groups based on  $ScoreMin_i$ . The conditions for group assignment depend on the value  $Margin_i$  takes. The number of inspections,  $N_i$ , and the number of discretionary violations,  $N_j$ , in each subgroup are reported in parentheses.

## 5 Results

### 5.1 Evidence of Favorable Deductions on the Margin

Table 4 presents estimates from equation (1) under four specifications. The first two use all 101,077 discretionary violations. In column (1), indicators of the health code violated and value of  $Group_i$  are the only controls included. In column (2), fixed effects for the day of the week, month of the year, year, and restaurant zip code, are also included. While serial correlation in the error likely exists within zip codes, where restaurants are subject

<sup>24</sup>There are 44 inspections involving discretionary violations where  $ScoreMin_i \leq 71$  and  $Margin_i = 0$ . Observations from these inspections are dropped in estimation so that scoring decisions on the margin will be compared to decisions in inspections exhibiting better overall hygiene quality.

to similar inspectors; there is also potential serial correlation across zip codes relating to the health code that has been violated. To account for this, standard errors are clustered two-way, following Cameron et al. (2011), at the levels of restaurant zip code and health code violated.

The estimates presented in columns (1) and (2) both suggest a statistically significant and very large increase in the probability of the lesser point deduction when restaurants are on the margin of a higher letter grade. Evaluating all other covariates at their means, the two specifications predict that 2-point deductions are assigned at probabilities between 0.5734 and 0.5754 in inspections where restaurants are not on the margin.<sup>25</sup> Relative to those predictions, restaurants on the margin appear 28.61 to 29.59 percent more likely to receive the lesser deduction. This increase in probability is especially striking, not only for its relative magnitude, but given that comparison groups are constructed such that inspections on the margin are compared to inspections that exhibit better overall hygiene quality.<sup>26</sup>

Within my sample, there are 10,590 restaurants that committed at least one discretionary violation in an inspection where they were on the margin, and at least one discretionary violation in a inspection where they were not on the margin. Columns (3) and (4) of Table 4 report estimates from this subsample including restaurant fixed effects. These estimates compare how scoring decisions at individual restaurants vary when they are, and are not, on the margin of a higher letter grade. Standard errors in columns (3) and (4) are also clustered two-way at the levels of restaurant zip code and health code violated.<sup>27</sup> These estimates suggest an even larger effect than the full sample estimates reported in columns (1) and (2). Relative to the predicted probabilities when  $Margin_i = 0$ , they suggest that restaurants are roughly 40 percent more likely to receive the lesser deduction when they are

---

<sup>25</sup>For a simple average, when  $Margin_i = 0$ , the 2-point deduction is assessed at a relative frequency of about 0.5729.

<sup>26</sup>To test whether these results are sensitive to the way  $Group_i$  is constructed, I estimate of equation (1) restricting the sample to discretionary violations in inspections where  $Margin_i = 1$ , or  $ScoreMin_i \geq 90$ . These estimates are found in Table A1 of the Appendix. The estimated effect of  $Margin$  is still positive, statistically significant, and quite large.

<sup>27</sup>In both specifications, two-way clustering at the levels of the individual restaurant and health code violated produces slightly smaller standard errors on the coefficient of interest.

on the margin of a higher letter grade.

## 5.2 How Often Were Scores Manipulated?

The results presented in Table 4 provide strong evidence suggesting favorable deduction decisions when restaurants are on the margin. To better interpret these results, it will help to consider how many inspection scores were likely manipulated. To accomplish this, I estimate equation (1) with the sample restricted to discretionary violations where  $Margin_i = 0$ .<sup>28</sup> Parameter estimates are then used to generate predicted values,  $\hat{Y}_{j(i)}$ , for the full sample of discretionary violations.

These predicted values estimate conditional probabilities of the lesser point deduction based on observations in which scoring decisions were not likely corrupted (because the decisions had no letter-grade implications). In order to identify inspections that were likely manipulated, I project the inspection scores that these predicted probabilities imply. If  $\hat{Y}_{j(i)} \geq 0.5$ , I deduct 2 points, and 4 points are deducted otherwise. The out-of-sample projected scores suggest that manipulation very common.

A total of 19,804 inspections in which restaurants were on the margin resulted in *A* grades. Among those, 5,307 inspections (about 26.80 percent) were projected to result in *B* grades based on the manner in which deductions are assessed when restaurants are not on the margin. Overall, these projections imply that favorable deduction decisions were used to manipulate 5,921 inspection scores and produce higher letter grades. This amounts to 26.56 percent of the 22,293 inspections in which restaurants were on the margin of a higher letter grade. That is, manipulation *via* favorable deduction decisions occurred in about one of every four inspections where it was possible.

Among inspections where  $Margin_i = 1$ , Table 5 and Figure 3 compare the frequencies of letter grades corresponding to these projected scores against the actual letter grades that

---

<sup>28</sup>Thus, the variable  $Margin_i$  is dropped from the estimating equation. Included in estimation are indicators for the health code violated, the value of  $Group_i$ , and the day of the week, month of the year, year, and zip code in which the inspection occurred.

were received. Figure A1 of the Appendix contrasts the distribution of scores from all inspections in the sample, with the full sample distribution of projected scores.<sup>29</sup>

### 5.3 Where Were Inspection Scores Manipulated?

Expanding on the projections presented in section 5.2, I consider where in LA County the inspection scores flagged as likely manipulated occurred. My data do not indicate the inspectors conducting inspections. However, the District Surveillance and Enforcement Branch of the DPH conducts restaurant inspections out of twenty-nine district offices each covering different areas throughout the county. Thus, within each zip code, and across zip codes in proximity to each other, the individual inspectors conducting inspections should be highly recurrent. Assessing the spatial distribution of the flagged inspection scores can reveal information regarding the nature of corruption in the inspection process.

Are the suspect inspections distributed across county zip codes with a fairly even frequency, or are many of them concentrated in certain geographic areas? If the flagged inspections are evenly distributed, it would suggest that many inspectors, on rare occasions, made questionable deduction assessments. Whereas, geographic concentration would suggest that a relatively small group of inspectors issued suspect deductions with some regularity.

There are 283 different zip codes in the full sample. A group of 5 zip codes account for 10.89 percent of the flagged inspections, while only accounting for only 3.71 percent of the inspections in the sample. Additionally, a group of 40 zip codes involve about 25.61 percent the sample's inspections, and yet account for half of the flagged inspection scores. Using the full sample of 140,193 inspections, I compute the relative frequency with which the flagged inspections occur within each zip code. Figure 4 is a map of zip codes in LA County. Zip codes in which the relative frequency of flagged inspections was greater than or equal to the 90<sup>th</sup> percentile are shaded red.<sup>30</sup> Of those 26 zip codes, there are 17 which border at least

---

<sup>29</sup>In inspections which did not involve discretionary violations, the actual inspection score is used as the projected score as well.

<sup>30</sup>That is, the relative frequency of flagged inspections in those zip codes is greater than or equal to the relative frequency in at least 90 percent of zip codes.



one other zip code in that group.

Further evidence of such concentration is found by assessing the frequency of flagged inspections within zip codes on a single day. There are 12 separate instances where four flagged inspections occurred within the same zip code in one day. Table A2 in the Appendix reports the zip codes and dates of these instances, as well as the straight-line distance between the two furthest inspections involved in each instance. In all of these cases, the distance between the two furthest flagged inspections is less than or equal to 2.7 miles. In four instances, the furthest distance between two flagged inspections was less than one mile, and the average distance between the two furthest flagged inspections 1.28 miles.<sup>31</sup> The remarkable proximity of flagged inspections on these days strongly suggests a single inspector was responsible for all four of the flagged inspections in each. There were also 58 instances where three flagged inspections occurred in the same zip code on the same day, and 499 instances where two flagged inspections occurred in the same zip code, on the same day. Altogether, 1,220 (about 20.6 percent) of the flagged inspection scores occurred on the same day as at least one other flagged inspection in the same zip code.

## 6 Score Manipulation or Optimization by Restaurants?

Letter grades do not distinguish between, *e.g.*, a restaurant with a 90-point inspection score and a restaurant which committed no violations. Thus, restaurants may elect to provide the least costly level of hygiene quality sufficient to maintain an *A* grade. Suppose there is a subset of discretionary violations for which minor (2-point deduction) forms of the violation will reduce costs more than most other violations will. Further, suppose that with any of these violations, if the firm were to move from the minor to the major form of the violation, the additional reduction in cost would be less than the reduction received from committing an additional minor discretionary violation from that subset. The restaurants best able to

---

<sup>31</sup>For each of these 12 instances, Google Maps images with indicators marking the establishment locations are given in section OA1.1 of the Online Appendix found [here](#).

optimize would commit these minor discretionary violations, receive the 2-point deduction at a greater relative frequency than others, and end up on the margin by virtue of their ability to optimally respond to the letter-grade system. This presents a possible alternative interpretation of the main results.

Perhaps the strongest point in support of the manipulation interpretation is simply a plausibility argument. First, consider the conditions just listed which are needed for the alternative interpretation to hold up (in light of the Table 4 estimates). Also, note that the inspector has far greater control than the restaurant over incremental changes in an inspection score. For a restaurant to willingly commit a discretionary violation and exercise control over whether 2 or 4 points are deducted, they must understand the specific health code and the distinction between minor and major forms of the violation, as well as be able to commit the minor form without crossing the line beyond which the violation becomes major. An inspector however can affect whether 2 or 4 points are deducted by simply choosing to report a violation as major or minor.<sup>32</sup> Finally, there are 24 non-critical violations which carry a 1-point deduction, many of which employees might commit at any time due to ignorance of the code or lapse in attention/effort.<sup>33</sup> If a restaurant were attempting to coordinate a 90-point inspection score, this would result in a *B* grade, which could be quite costly given that they are so rare.<sup>34</sup> Unless restaurant managers/owners are highly confident in employees' competence (and confident that employee incentives are aligned with their own), intentionally generating a 90-point expected score would seem to exhibit fairly low aversion to risk.

The number of conditions needed for restaurant optimization to play a significant role

---

<sup>32</sup>Of course, an entirely scrupulous inspector, who will act only in accordance with the LA County health code, can exercise no such control. But, if an inspector being willing to break with the health code in some instances, producing a 2-point difference in inspection scores is a matter of simply checking one box or another on the inspection report.

<sup>33</sup>Examples are codes requiring: wiping cloths be properly used, employees exhibit personal cleanliness and properly use hair restraints, garbage and refuse be properly disposed, toilet facilities be properly supplied and cleaned, non-food contact surfaces be cleaned, and equipment/utensils be cleaned.

<sup>34</sup>In my sample, 94.86 percent of inspections resulted in *A* grades, and these are only inspections which found violations (the sample does not include perfect inspection scores). Thus, a *B* grade would place a restaurant among select and undesirable company.

in the data make manipulation involving inspectors the far more plausible explanation of the main results. Still, empirical tests of this alternative explanation will provide stronger evidence of manipulation. In this section, I examine whether, and to what extent, the observed increase in the relative frequency of the 2-point deduction in inspections on the margin can be attributed to optimization by restaurants.

## 6.1 Deduction Decisions Across Violation Type

Are the main results driven primarily by deduction decisions on a few violation types, or did a restaurant’s position on the margin affect deduction decisions across a variety of violation types? There are eleven different discretionary violation types, but a restaurant can violate at most five while maintaining an *A* grade.<sup>35</sup> If the estimates in Table 4 merely reflect the chosen violations of restaurants that are best able to optimize, then the effect of *Margin* on scoring decisions should be concentrated among a few violation types (violations of whichever health codes are most costly to obey). This hypothesis can be tested without any judgments as to which health codes are most costly to obey. If cost minimization (subject to producing an *A* grade) is truly driving the results, then, unless there is considerable heterogeneity across restaurants in the costs of observing different regulations, the data should reveal which health codes are most costly to obey.

I augment equation (1) to assess how the effect of *Margin* on deduction decisions varies across the different health codes being violated. I construct a categorical variable, *HealthCode<sub>j</sub>*, which indicates the health code that violation *j* disobeys, and estimate the following equation:

$$Y_{j(i)} = \left( \sum_{h=1}^{11} \alpha_h [Margin_i \times I(HealthCode_j = h)] \right) + \mathbf{X}_i' \boldsymbol{\beta} + \mathbf{Z}_j' \boldsymbol{\gamma} + \epsilon_{j(i)}. \quad (2)$$

---

<sup>35</sup>There is only one inspection in the sample where a restaurant committed exactly five discretionary violations for a score of 90 points. There are 19,804 inspections in the sample where a restaurants received *A* grades while on the margin. More than 91 percent of those inspections involved the commission of one or two discretionary violations, and more than 99 percent involve the commission of three discretionary violations or fewer.

Table 6 reports estimates of equation (2) with the inclusion of indicator variables for: the health code violated, the value of  $Group_i$ , and the day of the week, month of the year, year, and zip code in which the inspection occurred.<sup>36</sup> Standard errors are clustered two-way at the levels of the restaurant zip code and the health code violated. The third column of the table reports the predicted probability of the lesser deduction when: the health code corresponding to that row is violated,  $Margin_i = 0$ , and all other covariates are evaluated at their means. When restaurants are on the margin, the lesser point deduction is significantly more likely to be assessed on every single discretionary violation type, suggesting that deduction decisions were likely made giving undue consideration to their letter-grade implications.

In Table 6, the health code requiring “hands clean and properly washed; gloves used properly”, stands out. Relative to the predicted probability when  $Margin = 0$ , the coefficient of 0.2790 suggests that being on the margin causes a 92.51 percent relative increase in the probability of the lesser point deduction. Beyond the immense magnitude of that effect, this particular violation is noteworthy because it can’t plausibly be an artifact of restaurant optimization. Proper hand washing and glove use bear little if any cost and detected violation of this code likely results from ignorance or an employee mistake. That is, a restaurant could regularly violate this code and still hide that fact by simply behaving appropriately during an inspection. However, other discretionary violations likely can’t be hidden during unannounced inspections.<sup>37</sup> If a health code’s regular violation can be hidden during unannounced inspections, then the detected commission of that violation can never be optimal, regardless of the cost of complying with that code. This is because the point-deduction it incurs could instead be allocated toward the commission of a violation which can’t be hidden during inspections.

I also estimate equation (2) restricting the sample to discretionary violations in inspec-

---

<sup>36</sup>Estimates of equation (2) under the parsimonious specification used in column (1) of Table 4 produce very similar estimates in sign, significance, and magnitude.

<sup>37</sup>*E.g.*, keeping a refrigerator warmer than mandated, which violates the “proper hot and cold holding temperatures” code, can’t be hidden during an unannounced inspection because, once the inspector has arrived, it would take too long for the refrigerator to reach the compliant temperature.

tions where  $ScoreMin_i \geq 80$ . That is, for inspections on the margin of a higher letter grade, I use only those that were on the margin of an *A*. To maintain an *A* grade, restaurants can commit, at most, five discretionary violations. Thus, among this subsample, if the lesser point deduction is significantly more likely across several violation types, this would strongly reject any notion that the main results merely characterize restaurant optimization. These estimates are reported in Table A3 of the Appendix. They suggest that restaurants on the margin of an *A* grade are significantly more likely to receive the lesser point deduction on ten of the eleven discretionary violation types, including the code for “proper hand washing/glove use”. The lone exception is the seldom cited violation of “proper reheating procedures for hot holding”. These results can’t be explained solely by restaurants providing the least costly level of hygiene quality sufficient for an *A* grade. Rather, they suggest that in many cases, point deduction assessments were heavily influenced by a restaurant’s position on the margin of a higher letter grade.

## 6.2 Favorable Deductions on the Margins of *B* and *C* Grades

The notion of restaurants electively bunching at 90 points assumes that they have considerable control over their hygiene quality provision at all times. That assumption is far less plausible regarding restaurants on the margins of *B* or *C* grades. For one, inspections on these two margins involve far more violations than those on the margin of an *A* grade.<sup>38</sup> This makes it more difficult for a restaurant to coordinate an exact 70 or 80-point inspection score involving minor discretionary violations. Second, if a restaurant is capable of coordinating an exact 70 or 80-point score, then they are certainly capable of coordinating a 90-point score. It is doubtful that differences in costs would be large enough to make such a restaurant opt for a *B* grade over an *A* grade, especially given the relative scarcity of *B* grades in LA County (refer back to footnote 34). Thus, by examining point-deduction decisions on the margins of *B* and *C* grades only, any large disparity in the probability of the

---

<sup>38</sup>For instance, restaurants committed an average of 11.88 violations (3.28 discretionary and 8.60 others) in inspections that produced *B* grades on the margin.

lesser deduction depending on  $Margin_i$  can be reasonably attributed to favorable deduction decisions.

I estimate equation (1) restricting observations to inspections where:  $Margin_i = 0$  and  $Group_i \geq 3$  ( $ScoreMin_i \leq 85$ ), or  $Margin_i = 1$  and  $ScoreMin_i < 80$ . I also construct two new variables and estimate the following equation using the same restricted sample:

$$Y_{j(i)} = \beta_1 MarginB_i + \beta_2 MarginC_i + \mathbf{X}_i' \boldsymbol{\beta} + \mathbf{Z}_j' \boldsymbol{\gamma} + \epsilon_{j(i)}. \quad (3)$$

$MarginB_i$  equals 1 if  $ScoreMin_i < 80$  and  $ScoreMax_i \geq 80$ , and equals 0 otherwise; and  $MarginC_i$  equals 1 if  $ScoreMin_i < 70$  and  $ScoreMax_i \geq 70$ , and equals 0 otherwise. Equation (3) simply allows the effects of being on the margin of a  $B$  grade, and being on the margin of a  $C$  grade, to be separately estimated.

Estimates under both specifications are reported in Table 7. Column (1) reports estimates of equation (1) using the restricted sample. Notice that even among restaurants with no chance of receiving an  $A$  grade, being on the margin still significantly and substantially increases the probability of the lesser deduction. This effect is driven primarily by restaurants on the margin of  $B$  grades, as seen in column (2) which reports the estimates of equation (3) using the restricted sample. Compared to restaurant inspections exhibiting slightly better hygiene quality, being on the margin of a  $B$  grade is estimated to significantly increase the probability of the lesser deduction by 0.1304. Moreover, relative to this sample's predicted probability when  $Margin_i = 0$ , restaurants are about 30.81 percent more likely to receive the lesser deduction when on the margin of  $B$  grades. That restaurants on the margin of  $B$  grades are significantly and substantially more likely to receive the lesser deduction, even when compared to restaurants exhibiting better overall hygiene quality, provides particularly strong evidence of manipulation.

### 6.3 Assessing Repetition by Restaurants on the Margin

If to some extent the main results reflect the type and manner of violations chosen by the restaurants best able to optimize, then among these restaurants, repetition over time in the type and manner of violations committed should be observed. That is, the set of committed violations that a restaurant finds optimal is unlikely to change much from one inspection to the next. Thus, the frequency with which individual restaurants on the margin repeat certain violations can reveal the extent to which the main estimates might reflect the behavior of restaurants that were optimally responding to the letter-grade system.

First, I simply evaluate how often individual restaurants secure an *A* grade while on the margin. That is, before even assessing whether restaurants that make *A* grades on the margin commit certain violation types repeatedly, I assess how many restaurants were even able to repeatedly make an *A* grade on the margin. If a restaurant finds that making an *A* grade on the margin is optimal *and* is able to coordinate this, then that restaurant should wish to, and be able to, repeat that performance. In the full sample there were 19,804 inspections in which 13,672 different restaurants received an *A* grade while on the margin. Of those restaurants, 9,253 (about 67.68 percent) made an *A* grade while on the margin only once.

Next, among the restaurants which made *A* grades on the margin more than once, I assess how often they committed the same discretionary violations across these inspections. I construct a 22-digit binary string which indicates the exact combination of discretionary violations in an inspection (this distinguishes between minor and major forms of the eleven violation types). The 4,419 restaurants which made *A* grades on the margin more than once, did so in 10,551 inspections. Within those inspections, only 1,456 (about 13.8 percent) involved a combination of discretionary violations which the restaurant committed more than once. Moreover, only 14 of the 10,551 inspections involved a combination of all violations which the restaurant committed in more than one of their *A* grades on the margin. The general lack of repetition over time exhibited by these restaurants suggests that optimal response to the letter-grade system is likely not a significant factor in my results. However,

I also test the extent to which my main results are robust to excluding observations which may result from restaurant optimization.

Table 8 reports estimates of equation (1) using several different restricted samples. Recall that there were 7 restaurants that twice made *A* grades on the margin with the same combination of all violations. Column (1) excludes all discretionary violations committed in those 14 inspections, and column (2) excludes all discretionary violations committed by those restaurants. There were also 1,456 inspections in which 700 restaurants made *A* grades on the margin by committing the same set of discretionary violations. Column (3) excludes discretionary violations committed in those 1,456 inspections, and column (4) excludes all discretionary violations committed by those restaurants. Finally, there were 10,551 inspections in which 4,419 restaurants made *A* grades on the margin more than once. Column (5) excludes discretionary violations committed in those inspections, and column (6) excludes all discretionary violations committed by those restaurants.

The coefficients reported in all columns of Table 8 come from estimating equation (1) with the inclusion of fixed effects for the health code violated, the value of  $Group_i$ , and the day of the week, month of the year, year, and restaurant zip code in which the inspection occurred. Notice in column (6) that, even while excluding all discretionary violations committed by restaurants that made multiple *A* grades on the margin (the most restrictive of the subsamples used), the effect of *Margin* on the probability of the lesser deduction is still statistically significant, substantial, and only less than the estimate reported in column (2) Table 4 by 0.0187. Relative to the probability of 0.5754 predicted from the unrestricted sample estimates, the coefficient reported in column (6) suggests that restaurants on the margin are still 25.36 percent more likely to receive the lesser deduction, even when all observations from restaurants that made multiple *A* grades on the margin are excluded. Note that these restaurants accounted for 29.89 percent of all minor discretionary violations in the sample. It indicates how extensive manipulation was, that the effect of *Margin* is still significant and large even after excluding so many observations where restaurants made *A* grades on the



margin. These estimates, together with the results presented in sections 6.1 and 6.2, show that little if any of the estimated effect of *Margin* on deduction decisions can be attributed to restaurants electively bunching at letter grade thresholds.

## 7 Concluding Remarks

In the regulation of restaurant hygiene, the growing popularity and considerable promise of mandatory disclosure policies, juxtaposed with the apparent heterogeneity across jurisdictions in their design and implementation,<sup>39</sup> raises questions regarding how different disclosure policy designs compare in their ability to induce quality improvements. This paper addresses the effects of disclosing broad intervals to which a quality measure belongs, rather than the measure itself; the significant feature being that producers with observed differences in product quality disclose a signal that does not reveal that difference. Such designs distort restaurant incentives on both sides of interval thresholds.

I show that these distortions have inhibited the effectiveness of hygiene quality disclosure in LA County. Specifically, the relatively strong incentives for score manipulation that exist below letter-grade thresholds have influenced scoring decisions by some inspectors. From October 2014 through September 2016, even when compared with restaurants exhibiting better hygiene quality, restaurants on the margin of a higher letter grade were 28 to 40 more likely to receive the lesser point deduction on discretionary violations. I find that 5,921 inspection scores were likely manipulated to produce higher letter grades, which amounts to 4.2 percent of the inspections in the sample, and 26.56 percent of inspections where restaurants were on the margin. Put differently, favorable deductions were used to improve letter grades in about one of every four inspections where that was possible. The inspection scores flagged as likely manipulated were largely concentrated within particular geographic areas

---

<sup>39</sup>Other LA County, nineteen city, county, and state governments presently share restaurant inspection data with Yelp in addition to requiring on-site disclosure of hygiene quality signals. Most of these jurisdictions score inspections in some manner, but only ten require numeric score disclosure. Five disclose letter-grade intervals similar to LA County, while Fort Worth, Sacramento, York (Canada), and the state of Florida merely disclose whether a restaurant passed or failed the inspection.

of the county suggesting that a similar set of inspectors were involved in many of the suspect scoring decisions.

It is doubtful that much, if any, of the increased probability of lesser deductions on the margin is an artifact of restaurants electively bunching at letter-grade thresholds. Being on the margin of a higher letter grade significantly and substantially increases the probability of the lesser deduction across all eleven discretionary violation types, including the violation of codes which could be hidden during inspections. Compared with inspections exhibiting slightly better hygiene quality, restaurants on the margin of *B* grades were 30.81 percent more likely to receive the lesser deduction, and there is little chance that restaurants would be able to coordinate *B* grades on the margin, let alone want to. Moreover, among restaurants that made *A* grades on the margin, there is very little repetition across inspections in the types of violations committed, and 67.68 percent of these restaurants made *A* grades on the margin only once in the sample. Finally, there were 4,419 restaurants that produced *A* grades on the margin more than once in the sample. Even after excluding observations from these restaurants, estimates suggest that restaurants on the margin are still 25.36 percent more likely to receive the lesser deduction.

These results have important implications for governments pondering the adoption of quality disclosure programs, or the revision of existing ones. When policymakers adopt mandatory disclosure policies, the distinction between reporting numeric scores or letter grades may seem rather insignificant given that either approach will substantially increase information provision relative to a state of no disclosure. However, this paper demonstrates that letter-grade disclosure is conducive to manipulation attempts which have the potential to undermine the inspection and regulation of restaurants. That is, beyond limiting the ability of mandatory disclosure to improve and regulate product quality, reporting broad quality measure intervals can also have the more detrimental effect of promoting manipulative and corrupt activity among producers and regulators.

## References

- Benneer, L. S. and S. M. Olmstead (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.
- Cameron, A., J. B. Gelbach, and D. Miller (2011). Robust inference with multiway clustering. *Journal of Business and Economic Statistics* 29(2), 238–249.
- Centers for Disease Control and Prevention (2013). Surveillance for foodborne disease outbreaks United States, 2013: Annual report. Retrieved from <http://www.cdc.gov/foodsafety/pdfs/foodborne-disease-outbreaks-annual-report-2013-508c.pdf>.
- Dranove, D. and G. Z. Jin (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature* 48(4), 935–963.
- Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite (2003). Is more information better? The effects of “report cards” on health care providers. *Journal of Political Economy* 111(3), 555–588.
- Figlio, D. N. and L. S. Getzler (2006). *Accountability, Ability and Disability: Gaming the System?*, pp. 35–49.
- Forbes, S., M. Lederman, and T. Tombe (2015). Quality disclosure programs and internal organizational practices: Evidence from airline flight delays. *American Economic Journal: Microeconomics* 7(2), 1–26.
- Hanushek, E. and M. E. Raymond (2004). The effect of school accountability systems on the level and distribution of student achievement. *Journal of the European Economic Association* 2(2/3), 406–415.
- Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics* 89(5-6), 761–796.

- Jacob, B. A. and S. Levitt (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics* 118(3), 843–877.
- Jin, G. Z. and P. Leslie (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics* 118(2), 409–451.
- Lu, S. F. (2012). Multitasking, information disclosure, and product quality: Evidence from nursing homes. *Journal of Economics and Management Strategy* 21(3), 673–705.
- Makofske, M. P. (2017). The effect of information salience on product quality: Louisville restaurant hygiene and Yelp.com. *MPRA Working Paper (79690)*.

Table 3: Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
<i>Per Inspection</i>					
Score	140,163	93.7271	(3.6527)	63	99
Violations	140,163	4.7323	(2.4338)	1	24
Critical Violations	140,163	0.9750	(0.9935)	0	9
Non-Critical Violations	140,163	3.6157	(1.9834)	0	16
Discretionary Violations*	140,163	0.7223	(0.8457)	0	7

\* Discretionary violations are critical violations for which 2 or 4 points may be deducted.

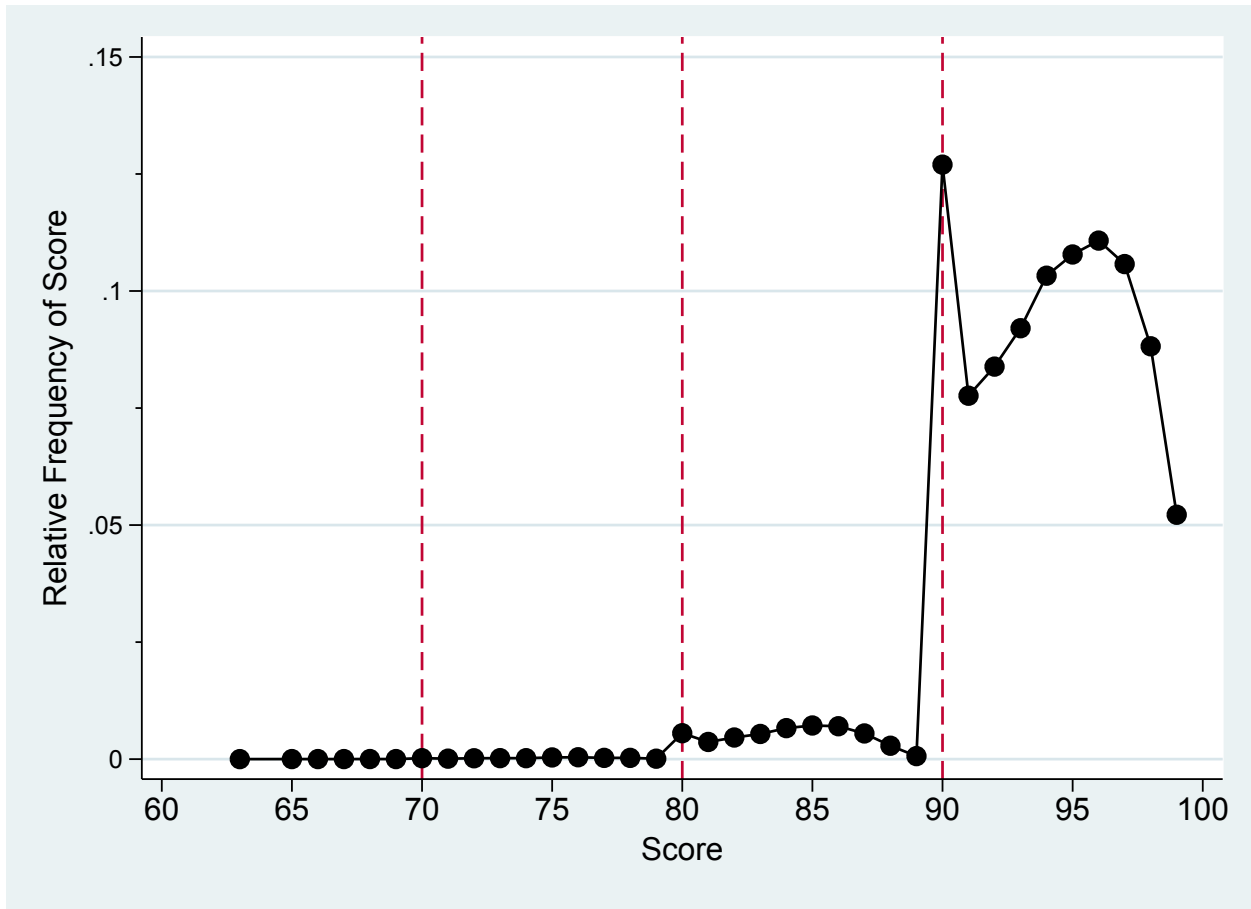


Figure 1: Distribution of Restaurant Health Inspection Scores

The plot is from 140,163 inspections conducted from October 1, 2014 to September 30, 2016. Black dots indicate the relative frequency of each score within that sample. Dashed lines at 90, 80, and 70 points, indicate the thresholds for A, B, and C letter-grades respectively.

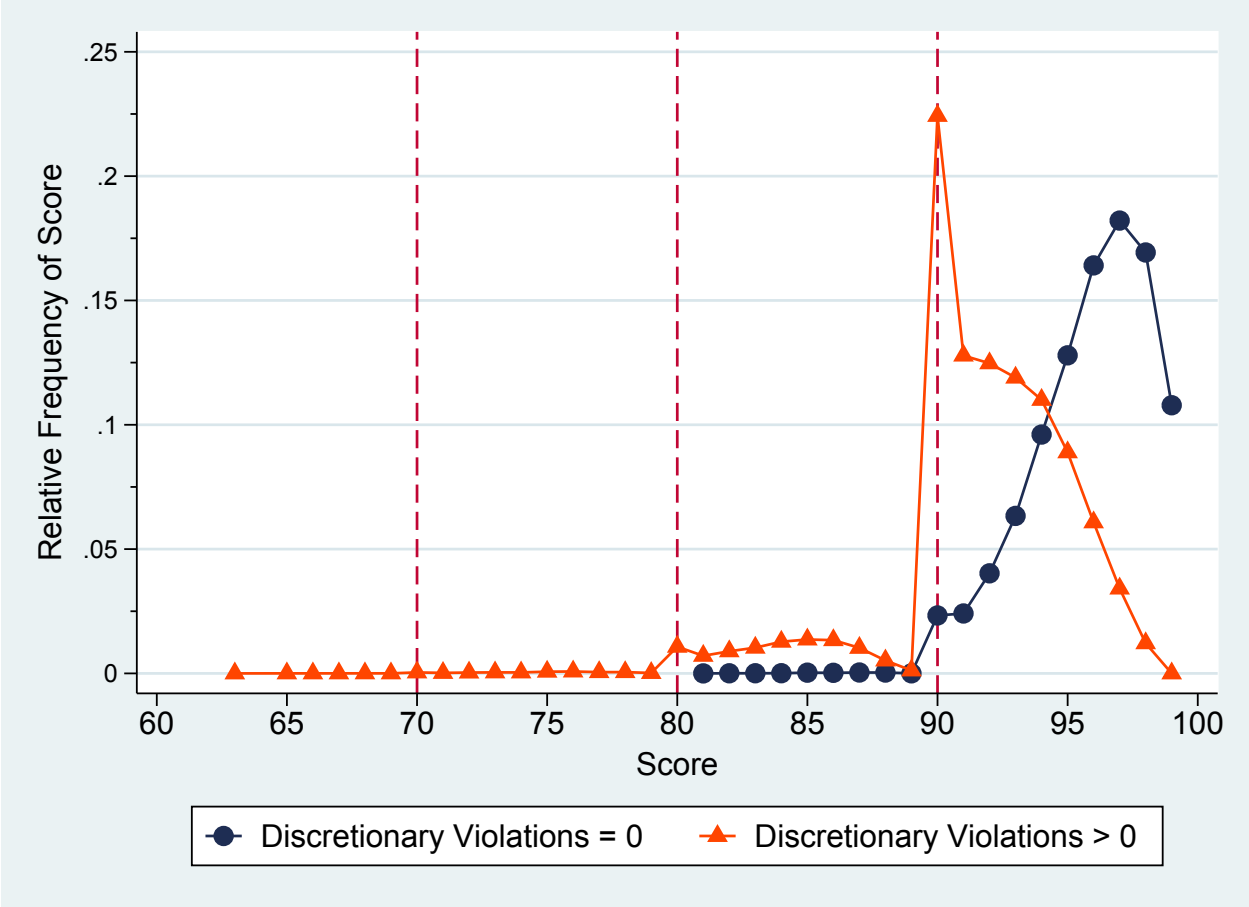


Figure 2: Distribution of Restaurant Health Inspection Scores by Discretionary Violations  
 Navy dots indicate the relative frequency of each score among the 67,835 inspections in which no discretionary violations were detected. Orange triangles indicate the relative frequency of each score among the 72,328 inspections in which one or more discretionary violations were detected. Dashed lines at 90, 80, and 70 points indicate the thresholds for *A*, *B*, and *C* letter-grades respectively.

Table 4: Deduction Decisions on the Margin

Variable	(1) $Y_{j(i)}$	(2) $Y_{j(i)}$	(3) $Y_{j(i)}$	(4) $Y_{j(i)}$
<i>Margin</i>	0.1697*** (0.0328)	0.1646*** (0.0318)	0.2062*** (0.0379)	0.2065*** (0.0390)
$\widehat{\Pr}(Y = 1   \text{Margin} = 0, \bar{\mathbf{X}}, \bar{\mathbf{Z}})$	0.5734	0.5754	0.5143	0.5142
Health Code FE	Y	Y	Y	Y
<i>Group</i> FE	Y	Y	Y	Y
Day of Week FE	N	Y	N	Y
Month FE	N	Y	N	Y
Year FE	N	Y	N	Y
Zip Code FE	N	Y	N	N
Restaurant FE	N	N	Y	Y
R-squared	0.1883	0.2111	0.3761	0.3767
N	101,077	101,077	54,760	54,760

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Results are OLS estimates. Columns (1) and (2) use 101,077 discretionary violations detected in 72,284 inspections of 32,871 restaurants. Columns (3) and (4) use 54,760 discretionary violations detected in 35,207 inspections of 10,590 restaurants with at least one inspection where  $\text{Margin}_i = 1$ , and at least one inspection with a discretionary violation where  $\text{Margin}_i = 0$ .  $Y_{j(i)}$  equals 1 when the lesser (2-point) deduction is assessed, and equals 0 otherwise. Standard errors, reported in parentheses, are clustered two-way at the levels of the restaurant zip code and the health code violated.

Table 5: Inspections on the Margin: Actual and Projected Letter Grades

	Actual Grade			
	<i>A</i>	<i>B</i>	<i>C</i>	<i>Below 70</i>
Projected Grade = <i>A</i>	14,497	214	0	0
Projected Grade = <i>B</i>	<b>5,307</b>	1,447	21	0
Projected Grade = <i>C</i>	0	<b>588</b>	182	2
Projected Score Below 70	0	0	<b>26</b>	9

Projected and actual letter grades in inspections where  $Margin_i = 1$ . **Bold** values correspond to inspections where the actual letter grade was higher than the projected letter grade.

Projections are based on OLS estimates of equation (1) with the sample restricted to discretionary violations where  $Margin_i = 0$  only. Indicators for the health code violated, the value of  $Group_i$ , and the day of the week, month of the year, year, and zip code in which the inspection occurred are included.



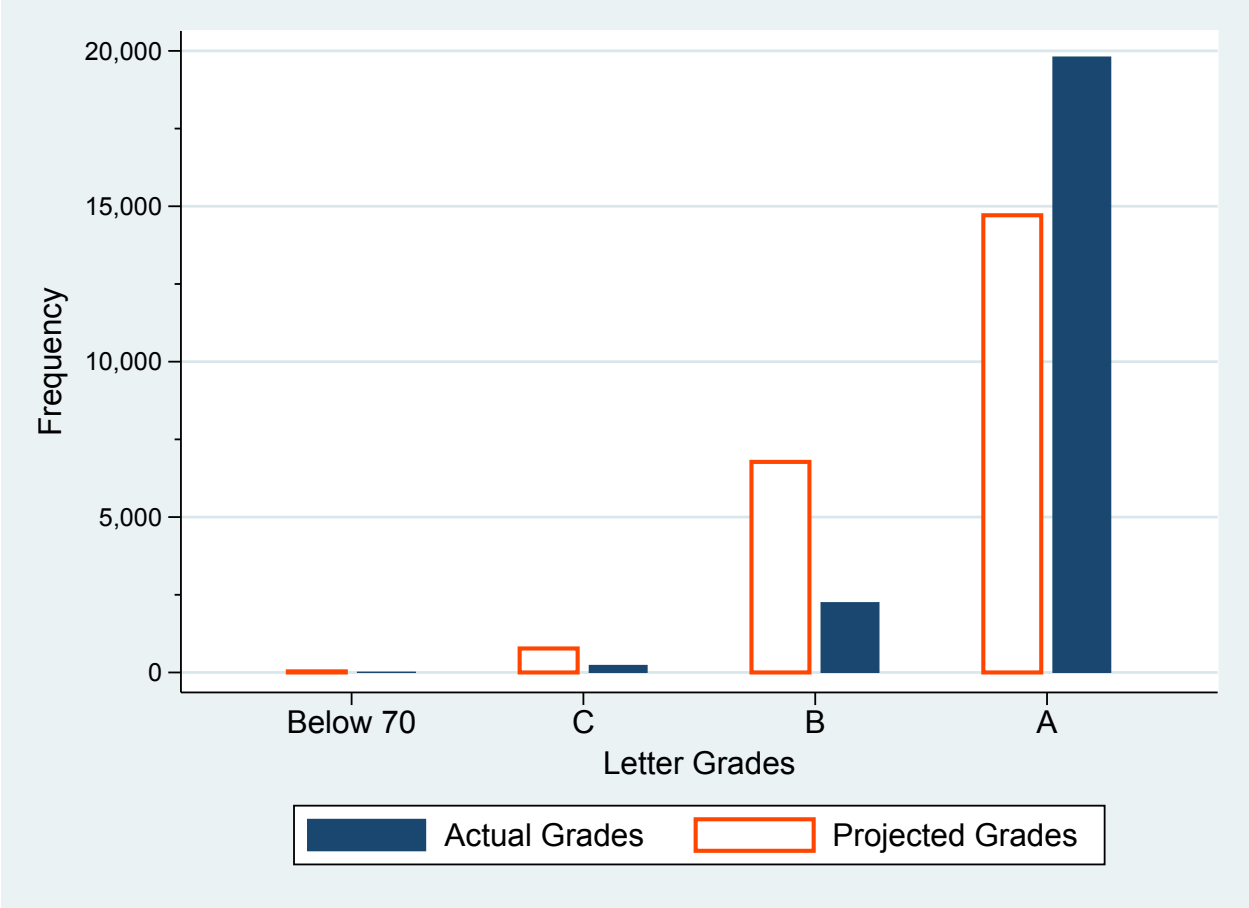


Figure 3: Inspections on the Margin: Actual and Projected Letter Grades

Navy bars show the frequency of each letter grade among the 22,293 inspections on the margin. Bars outlined in orange show the frequency of each projected letter grade among the 22,293 inspections on the margin.

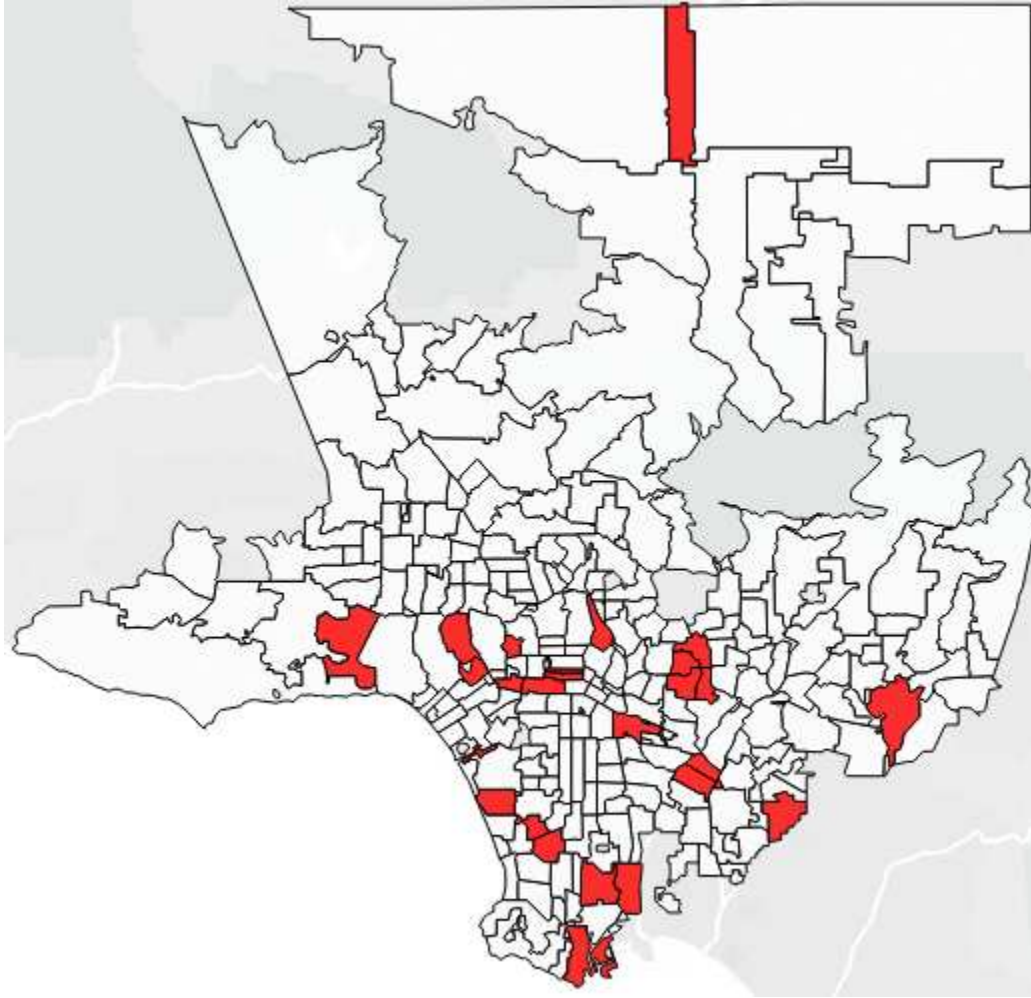


Figure 4: Relative Frequency of Flagged Inspections Across Zip Codes

The relative frequency (among all inspections) of inspection scores flagged as manipulated is computed for all zip codes in the sample. Percentiles were computed among all zip codes with at least 50 inspections in the sample. Zip codes in which the relative frequency of flagged inspections is greater than or equal to the 90<sup>th</sup> percentile are shaded red.

Table 6: Deduction Decisions on the Margin by Violation Type

Variable	$Y_{j(i)}$	$\widehat{\Pr}(Y = 1   \text{Margin} = 0)^\dagger$
<i>Margin</i> ×(Hands clean and properly washed; gloves used properly)	0.2790*** (0.0119)	0.3016
<i>Margin</i> ×(Proper hot and cold holding temperatures)	0.2122*** (0.0058)	0.4238
<i>Margin</i> ×(Time as a public health control; procedures & records)	0.2535*** (0.0116)	0.5141
<i>Margin</i> ×(Proper cooling methods)	0.1866*** (0.0073)	0.6437
<i>Margin</i> ×(Proper reheating procedures for hot holding)	0.0464*** (0.0087)	0.0000
<i>Margin</i> ×(Food in good condition, safe and unadulterated)	0.1274*** (0.0105)	0.6162
<i>Margin</i> ×(Food contact surfaces clean and sanitized)	0.2031*** (0.0093)	0.5188
<i>Margin</i> ×(Food obtained from approved source)	0.2325*** (0.0203)	0.2219
<i>Margin</i> ×(Hot and cold water available)	0.1327*** (0.0061)	0.7633
<i>Margin</i> ×(Sewage and wastewater properly disposed)	0.1043*** (0.0100)	0.7775
<i>Margin</i> ×(No rodents, insects, birds, or animals)	0.0424*** (0.0049)	0.8914
Health Code FE	Y	
Group FE	Y	
Day of Week FE	Y	
Month FE	Y	
Year FE	Y	
Zip Code FE	Y	
R-squared	0.2166	
N	101,077	

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

†Predicted probability given that the row's health code is violated,  $\text{Margin} = 0$ , and evaluating all other covariates at their means.

Results are OLS estimates from 101,077 discretionary violations detected in 72,284 inspections of 32,871 restaurants.  $Y_{j(i)}$  equals 1 when the lesser (2-point) deduction is assessed, and equals 0 otherwise. Standard errors, reported in parentheses, are clustered two-way at the levels of the restaurant zip code and the health code violated.

Table 7: Deduction Decisions on the Margins of  $B$  and  $C$  Grades

Variable	(1) $Y_{j(i)}$	(2) $Y_{j(i)}$
<i>Margin</i>	0.1288*** (0.0128)	— —
<i>B Grade Margin</i>	— —	0.1304*** (0.0128)
<i>C Grade Margin</i>	— —	0.0627* (0.0377)
$\widehat{\Pr}(Y = 1   \textit{Margin} = 0, \bar{\mathbf{X}}, \bar{\mathbf{Z}})$	0.4233	0.4234
Health Code FE	Y	Y
<i>Group</i> FE	Y	Y
Day of Week FE	Y	Y
Month FE	Y	Y
Year FE	Y	Y
Zip Code FE	Y	Y
R-squared	0.2014	0.2016
N	15,029	15,029

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Results are OLS estimates from discretionary violations detected in inspections of restaurants where  $\textit{Margin}_i = 0$  and  $\textit{Group}_i \geq 3$ , or  $\textit{Margin}_i = 1$  and  $\textit{ScoreMin}_i < 80$ .  $Y_{j(i)}$  equals 1 when the lesser (2-point) deduction is assessed, and equals 0 otherwise. *B Grade Margin* equals 1 if a restaurant is on the margin of a  $B$  grade, and *C Grade Margin* equals 1 if a restaurant is on the margin of a  $C$  grade. Standard errors, reported in parentheses, are clustered two-way at the levels of the restaurant zip code and the health code violated.

Table 8: Deduction Decisions on the Margins with Restricted Samples

Variable	(1) $Y_{j(i)}$	(2) $Y_{j(i)}$	(3) $Y_{j(i)}$	(4) $Y_{j(i)}$	(5) $Y_{j(i)}$	(6) $Y_{j(i)}$
<i>Margin</i>	0.1645*** (0.0318)	0.1645*** (0.0318)	0.1608*** (0.0310)	0.1599*** (0.0307)	0.1468*** (0.0291)	0.1459*** (0.0287)
$\widehat{\Pr}(Y = 1  $ <i>Margin=0, <math>\bar{X}</math>, <math>\bar{Z}</math>)</i>	0.5753	0.5753	0.5748	0.5754	0.5727	0.5820
R-squared	0.2111	0.2111	0.2084	0.2081	0.2037	0.2050
N	101,049	101,034	98,893	96,982	82,200	73,953

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Results are OLS estimates with fixed effects included for the health code violated, value of  $Group_i$ , the day of the week, month of the year, year, and restaurant zip code in which the inspection occurred. Standard errors, reported in parentheses, are clustered two-way at the levels of the restaurant zip code and the health code violated.

Each column reports estimates from a different restricted sample. (1) Excludes discretionary violations from inspections where restaurants made an  $A$  on the margin more than once with the same set of all violations. (2) Excludes all discretionary violations committed by restaurants that made an  $A$  on the margin more than once with the same set of all violations. (3) Excludes discretionary violations from inspections where restaurants made an  $A$  on the margin more than once with the same set of discretionary violations. (4) Excludes all discretionary violations committed by restaurants that made an  $A$  on the margin more than once with the same set of discretionary violations. (5) Excludes discretionary violations from inspections where restaurants made an  $A$  on the margin more than once. (6) Excludes all discretionary violations committed by restaurants that made an  $A$  on the margin more than once.

# A1 Appendix

## A1.1 Inspections with Identical Counts of Critical and Non-Critical Violations

This subsection demonstrates that an inspection where  $Margin_i = 0$  will exhibit slightly lower hygiene quality than an inspection involving identical counts of critical and non-critical violations where  $Margin_i = 1$ . Suppose two inspections both find four critical violations and zero non-critical violations, but one is on the margin of a higher letter grade, and the other is not. In *inspection 1*, one violation carries a 2-point deduction, one violation carries a 4-point deduction, and the other two violations are discretionary. *Inspection 1* is on the margin of a higher letter grade as  $ScoreMin_1 = 86$ , and  $ScoreMax_1 = 90$ . In *inspection 2*, one violation carries a 2-point deduction, two violations carry a 4-point deduction, and one violation is discretionary. *Inspection 2* is not on the margin of a higher letter grade as  $ScoreMin_2 = 86$ , and  $ScoreMax_2 = 88$ . At first glance, these two inspections might seem ideal for comparison as both involve the same number of critical violations, the same number of non-critical violations, and the same minimum score. However, *inspection 2* is not on the margin because it involves one less discretionary violation, and one more 4-point violation, than *inspection 1*. Thus, *inspection 2* is not on the margin due to exhibiting slightly lower hygiene quality. Restaurants of poorer hygiene quality may be more likely to warrant the 4-point deduction on discretionary violations, which will confound comparisons made within inspections that produce the same number of critical and non-critical violations.

Table A1: Deduction Decisions on the Margin: More Selective Comparisons

Variable	(1) $Y_{j(i)}$	(2) $Y_{j(i)}$
<i>Margin</i>	0.1407*** (0.0363)	0.1396*** (0.0356)
$\widehat{\Pr}(Y = 1   \textit{Margin} = 0, \bar{\mathbf{X}}, \bar{\mathbf{Z}})$	0.6039	0.6041
Health Code FE	Y	Y
Day of Week FE	N	Y
Month FE	N	Y
Year FE	N	Y
Zip Code FE	N	Y
R-squared	0.1579	0.1850
N	91,643	91,643

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Results are OLS estimates from 91,643 discretionary violations detected in 67,711 inspections of 32,249 restaurants in which either:  $\textit{Margin}_i = 1$ , or  $\textit{Margin}_i = 0$  and  $\textit{Group}_i = 1$ .  $Y_{j(i)}$  equals 1 when the lesser (2-point) deduction is assessed, and equals 0 otherwise. Standard errors, reported in parentheses, are clustered two-way at the levels of the restaurant zip code and the health code violated.

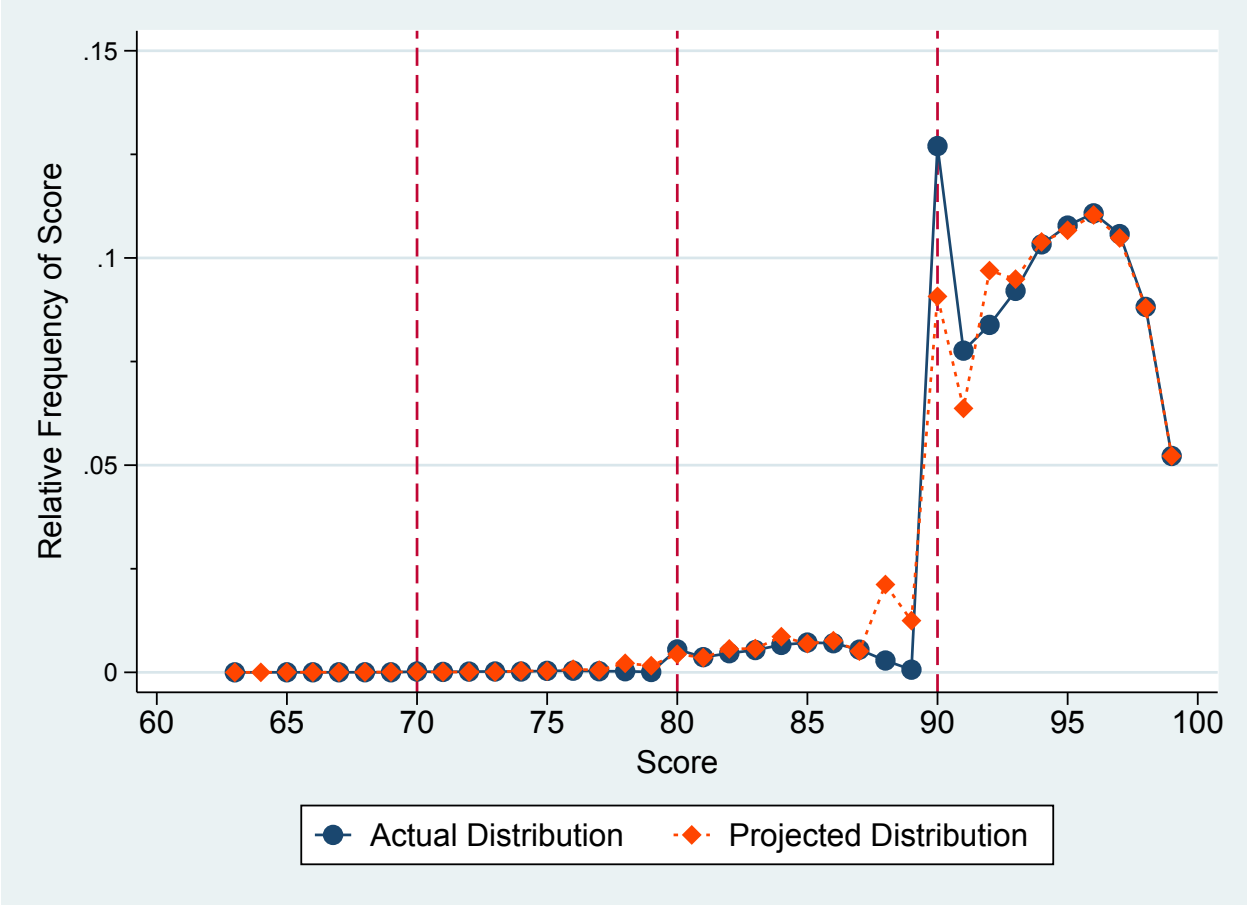


Figure A1: Actual and Projected Score Distributions

Orange dots indicate the relative frequency of each score within the full sample. Navy diamonds indicate the relative frequency of each predicted score within that sample. Predictions are from OLS estimates of equation (1) with observations restricted to inspections where  $Margin_i = 0$ . Controls correspond to those reported in columns (2) and (4) of Table 4. If  $\hat{Y}_{j(i)} \geq 0.5$ , a 2-point deduction is assessed. The 4-point deduction is assessed otherwise, and the projected scores follow.



Table A2: Zip Codes with Four Flagged Inspections in One Day

Zip Code	Date	Distance Between Furthest Inspections
90005	10/14/2015	1.06 miles
90019	10/28/2014	0.34 miles
90024	09/20/2016	0.49 miles
90027	05/10/2016	1.08 miles
90045	10/21/2015	1.98 miles
90069	05/25/2016	1.05 miles
90277	03/28/2016	0.46 miles
90277	05/20/2016	2.70 miles
90723	03/21/2016	2.07 miles
91801	05/28/2015	1.64 miles
91801	05/04/2016	0.94 miles
91801	07/28/2016	1.58 miles

Listed are twelve instances where four inspection scores flagged as manipulated occurred in the same zip code in a single day. Of the four establishments involved, the straight-line distance between the two which were furthest from each other is reported. Distances were measured using Google Maps.

For each of these instances, Google Maps images marking the locations of the establishments involved are available in the Online Appendix found [here](#).

Table A3: Deduction Decisions by Violation Type if  $ScoreMin_i \geq 80$

Variable	$Y_{j(i)}$	$\widehat{\Pr}(Y = 1   Margin = 0)$
<i>Margin</i> ×(Hands clean and properly washed; gloves used properly)	0.2576*** (0.0134)	0.3026
<i>Margin</i> ×(Proper hot and cold holding temperatures)	0.2305*** (0.0064)	0.4250
<i>Margin</i> ×(Time as a public health control; procedures & records)	0.2528*** (0.0132)	0.5150
<i>Margin</i> ×(Proper cooling methods)	0.1813*** (0.0080)	0.6442
<i>Margin</i> ×(Proper reheating procedures for hot holding)	-0.0015 (0.0071)	0.0004
<i>Margin</i> ×(Food in good condition, safe and unadulterated)	0.1338*** (0.0113)	0.6168
<i>Margin</i> ×(Food contact surfaces clean and sanitized)	0.2183*** (0.0094)	0.5203
<i>Margin</i> ×(Food obtained from approved source)	0.2431*** (0.0223)	0.2199
<i>Margin</i> ×(Hot and cold water available)	0.1209*** (0.0053)	0.7637
<i>Margin</i> ×(Sewage and wastewater properly disposed)	0.1086*** (0.0092)	0.7769
<i>Margin</i> ×(No rodents, insects, birds, or animals)	0.0383*** (0.0054)	0.8926
Health Code FE	Y	
Group FE	Y	
Day of Week FE	Y	
Month FE	Y	
Year FE	Y	
Zip Code FE	Y	
R-squared	0.2203	
N	94,962	

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Results are OLS estimates from 94,962 discretionary violations detected in 70,454 inspections of 32,684 restaurants, where  $ScoreMin_i \geq 80$ .  $Y_{j(i)}$  equals 1 when the lesser (2-point) deduction is assessed, and equals 0 otherwise. Standard errors, reported in parentheses, are clustered two-way at the levels of the restaurant zip code and the health code violated.