Disclosure Policy Design and Regulatory Agent Behavior

Makofske, Matthew

30 June 2022

DISCLOSURE POLICY DESIGN AND REGULATORY AGENT BEHAVIOR

Matthew Philip Makofske*

June 30, 2022

Abstract

Mandatory disclosure promotes self-regulation of product quality, but may also provide incentives for firms to manipulate disclosed information. Collection of quality information by inspectors protects against direct manipulation, but firms may attempt to unduly influence inspectors. In Las Vegas, Nevada, food-service health inspections are numerically scored, with health code violations from three categories carrying prescribed demerit amounts. For disclosure however, numeric scores are coarsened into letter grades, which may encourage lobbying of inspectors to under-report violations near threshold scores. Beginning in 2013, the demerit amount prescribed for each *good-practices* violation was reduced from 1 to 0; while the letter-grade scale and penalties prescribed for *major* and *critical* violations (3 and 5 demerits each) remained unchanged. Exploiting this removal of scoring implications from good-practices violations, coupled with the discontinuous punishment severity inherent in letter-grade disclosure, I find that inspectors significantly under-reported good-practices violations prior to 2013—by 31 to 90 percent in some cases—when those violations were likely to affect letter grades. Without careful design, disclosure policies supplementing inspection programs may inadvertently undermine the regulatory efforts they were meant to support.

*Department of Economics, Colgate University. 234 Persson Hall. Email: mmakofske@colgate.edu. Phone: (315) 228-7530. I thank Limin Fang and seminar participants at Colgate University, and the International Industrial Organization Conference 2022 for helpful comments. Any remaining errors are mine.

1 Introduction

Periodic inspections, in which regulatory violations are detected and penalized, are central to the provision of public health, environmental quality, and the safety of workplaces and many consumer products. Regulators in these settings confront a double-moral-hazard problem. Absent inspection, regulated entities would not comply to the extent desired; but also, inspectors may exert suboptimal detection effort, or even engage in behaviors like misreporting that undermine regulatory efforts. While perhaps less salient, the regulator-inspector relationship is more fundamentally important—inspector incentives must align with regulatory objectives if an inspection program is to achieve them. Yet, we know relatively little about the alignment of inspector and regulator incentives in practice.[1] Do inspectors sometimes behave counter to regulatory objectives; and if so, to what extent, and why? Using Southern Nevada Health District (SNHD) food establishment health inspections spanning nearly 9 years, I explore these questions and show that discontinuous punishment severity due to regulatory and disclosure-policy design likely drove misreporting by inspectors.

The SNHD conducts regular inspections of food establishments in which detected violations are recorded and potentially penalized. In addition to direct penalties for non-compliance such as fines or even temporary closures, the SNHD uses an increasingly popular supplement: mandatory disclosure of inspection performance. Mandatory disclosure promotes general deterrence by raising the expected cost of violations. Yet in general, the efficacy of disclosure policies can hinge significantly on details of their design or implementation; in many cases, because these details affect the incentives of firms.[2] A concern in this regard is that mandatory disclosure also creates incentives for firms to manipulate the disclosed information. If they are able, and the incentives to do so are sufficiently strong, manipulation may occur. Notably, Forbes et al. (2015) document

[1] Of note, two recent papers examine potential misalignment. With restaurant inspections from Florida, Jin and Lee (2018) find that in a second visit to an establishment, diminishing inspector attention explains the reporting of 13-18% fewer violations. Using Los Angeles County restaurant inspections, Makofske (2020b) finds that inspectors likely understated the severity of some reported violations when firms might otherwise face escalated penalties.

[2] On disclosure and firm behavior, see: Jin and Leslie (2003) on mandatory disclosure; Dai and Luca (2020), Johnson (2020), and Makofske (2020a) on the salience of disclosed information.

misreporting of arrival times by some US airlines in response to a disclosure-policy design feature. Thus, in addition to careful design, effective disclosure policies may require mechanisms protecting against potential manipulation.

When mandatory disclosure supplements inspection programs, a safeguard against direct manipulation already exists. Firms can't directly manipulate the disclosed information because it is collected and reported by the inspector. However, firms *can* lobby inspectors for beneficial treatment like, e.g., under-reporting the true extent of their non-compliance. If successful, not only is this indirect style of manipulation detrimental to the inspection program, it may become entrenched and difficult to remedy by its very nature. For one, inspectors who under-report and the firms that benefit have a shared interest in concealing it from the regulator. Moreover, under-reporting could easily be mistaken by the regulator as evidence of disclosure-induced improvements. Finally, because the regulated entities in this setting can't directly manipulate the disclosed information, regulators may simply overlook the possibility of this less direct form. As such, when designing disclosure policies, carefully considering potential manipulation incentives may actually be more important when said policies will supplement regulatory inspection.

SNHD health code violations incur prescribed demerit amounts which aggregate into inspection scores, and establishments must prominently display placards revealing information from their most recent inspection. Yet rather than their numeric score, the placards disclose only letter grades which pool all inspection scores within wide intervals,[3] and create relatively strong incentives for manipulation near letter-grade boundaries. Further, each reduction in letter grade below an A incurs fines. This policy design creates sharp jumps in punishment severity as demerit totals cross arbitrary grade thresholds. If inspectors harbor sympathy, favoritism, or some other interest toward establishments they inspect, these sharp discontinuities in penalty may induce inspectors to report fewer violations than detected—especially when full reporting might just trigger penalty escalation. Yet, assessing whether under-reporting actually occurs is fraught with difficulty.

Inability to observe the true set of violations present in inspections means detection

[3] A for 10 or fewer demerits, B for 11 to 20 demerits, C for 21 to 40 demerits, and X for more than 40 demerits.

of under-reporting is often not feasible, even when inspectors are found to act favorably to firms in other ways. In Los Angeles (LA) County food-service health inspections, numeric scores are also coarsened to letter grades for disclosure. Jin and Leslie (2003) find that the 1998 adoption of grade-card disclosure in LA County caused significant hygiene quality improvements.[4] Yet, they also notice dubious bunching at the *A*-grade threshold following adoption, pointing to possibly unintended responses by inspectors. Jin and Leslie (2005) mention interviews in which LA County inspectors felt great pressure from restaurateurs following grade card disclosure. Makofske (2020b) examines that issue using inspections from 2014-2016. Per LA County health code, some violations carry either a 2 or 4-point penalty based on severity. Makofske (2020b) finds that the lesser deduction was substantially more likely when deductions on these violations had the potential to affect an inspection's letter grade. That study finds that some inspectors likely understated the severity of *reported* violations due to the punishment schedule, but is still unable to assess whether under-reporting occurred. In this paper, I exploit a revision to SNHD scoring policy that provides a rare opportunity to detect evidence of under-reporting itself.

The SNHD defines—from most to least severe—*critical*, *major*, and *good-practices* violations,[5] which each initially carried 5, 3, and 1-demerit penalties, respectively. Beginning January 1, 2013, the scoring criteria were revised slightly. While each critical and major violation still incurs 5 and 3 demerits respectively, each good-practices violation—though still detected and reported—now carries 0 demerits. The letter-grade system and disclosure policy remained in place.

Before this scoring change, while the reported count of good-practices violations would always affect an inspection's score, it only carried letter-grade implications if it exceeded the margin between what I call *base demerits* (the demerits accrued on critical and major violations) in the inspection, and the next letter-grade threshold. Following the scoring

[4]They find substantial improvements in scores following staggered adoption of mandatory disclosure within the county, as well as evidence of declines in hospital admissions related to foodborne pathogens that are consistent with actual hygiene improvements.

[5]The least severe type were originally called *minor* violations, and later renamed violations of *good food management practices*. For expositional ease, I refer to this category throughout as *good-practices* violations (an abbreviation of the current name).

change, regardless of *base demerits*, the reported count of good-practices violations never carried letter-grade implications. With inspections spanning July 2011 to early March 2020, I exploit these features of the scoring change and test whether inspectors may have under-reported good-practices violations in response to letter-grade implications with difference-in-differences style approach. Controlling for various inspection-specific characteristics, the difference across scoring-regimes in reported good-practices violations is compared between inspections where the scoring change likely did, and likely did not, affect the presence of letter-grade implications.

Prior to 2013, while the reported count of good-practices violations would always affect an inspection's score, it only carried letter-grade implications if it exceeded the margin between what I call *base demerits* (the demerits accrued on critical and major violations) in the inspection, and the next letter-grade threshold. Following the scoring change, regardless of *base demerits*, the reported count of good-practices violations never carried letter-grade implications. Exploiting these features of the scoring change, I test for under-reporting using a difference-in-differences style approach. The difference across scoring-regimes in reported good-practices violations is compared between inspections where the scoring change likely did, and likely did not, affect letter-grade implications.

In the post-change period, a clear and positive linear relationship between base demerits and the reported count of good-practices violations is evident. Moreover, in pre-change inspections with 6 or fewer base demerits (where it is very unlikely that enough good-practices violations were present to cause a B grade), the reported count of good-practices violations exhibits a very similar linear slope in base demerits with a slightly higher intercept.[6] Beginning at 8 base demerits however,[7] reported counts of good-practices violations from pre-change inspections drop sharply as base demerits approach the 10-demerit threshold for an A grade. My primary estimates suggest that the true count of good-practices violations was understated 38 to 39 percent in pre-change inspections with 8 to 10 base demerits. Among pre-change inspections with 10 base demerits, estimates

[6]The change in intercept is consistent with the fact that the SNHD removed two violations from the good-practices category in 2013.

[7]Because critical and major violations carry 5 and 3 demerits each, 7 base demerits are not possible.

suggest that the true count good-practices violations was under-reported by 90 percent. Similar results are found as base demerits approach the B-grade threshold which, along with several other tests, suggests that my results likely reflect inspector behavior rather than an optimal response by establishments to scoring criteria. To my knowledge, this is the first study to provide empirical evidence of under-reporting by inspectors.

My finding reveal a subtle but important potential pitfall to mandatory disclosure in this setting. The adoption of mandatory disclosure strengthens incentives for self-regulation, and when it supplements existing inspection programs, the additional cost disclosure is relatively low. Yet, the design of disclosed information plays a very important role. Letter grades likely help consumers to broadly assess inspection performances in a that numeric scores alone might not; but they also introduce strong incentives for manipulation near cutoff scores if *only* the letter grade is disclosed. Disclosing numeric score along with letter grades could give broader context to consumers while also mitigating the manipulation incentives that letter-grade-only disclosure creates.[8]

In the space remaining, I discuss the SNHD inspection program and policies, the scoring revision, and other coincident changes. I then describe and summarize the data, and present some preliminary analyses that motivate my methodology. This is followed by a discussion of my empirical approach, tests of underlying assumptions, and presentation of the main results. I then conduct a battery of robustness tests supporting the interpretation of the main results, and conclude.

2 Policy Background

2.1 Scoring and Policy Regimes

The Southern Nevada Health District (SNHD) conducts routine health inspections of food-service establishments in the Las Vegas, Nevada, metropolitan area. It was established jointly by Clark County, and the cities of Boulder City, Henderson, Las Vegas, Mesquite, and North Las Vegas, as the public health authority within those entities.

[8]Notably, the Georgia Department of Public Health does just this, as seen here: https://ga.healthinspections.us/_templates/87/food_2015/_report_full.cfm?fsimID=57306665.

Establishments under the SNHD's jurisdiction receive at least one unannounced routine inspection each year. Through the SNHD website, inspection data going back to 2005 are made available. Since that time, a basic structure of scoring, grading, and disclosing inspection performance has been in place, but certain scoring and grading criteria have been changed along the way.

Throughout this period, detected health code violations carry prescribed demerit amounts. Demerits aggregate to form an inspection score which is then coarsened to a letter grade. Establishments are issued a placard displaying their inspection grade, which they must conspicuously post until their next inspection. The first policy regime spans 2005 until March 24, 2010. During that period, establishments were assigned: an A grade for 10 demerits or fewer; a B grade for 11 to 20 demerits; and a C grade for 21 to 40 demerits. Establishments incurring more than 40 demerits were shut down, fined, and assigned an X grade. This grade scale has been maintained ever since.

On March 25, 2010, the SNHD adopted revised *Regulations Governing the Sanitation of Food Establishments*. These revisions changed the health code's defined set of violations and demerit values in an effort to focus regulation on foodborne illness risk. The revised code established four categories of violation. The most severe, *imminent health hazards*, result in immediate closure of the establishment and fines. The SNHD defines the second-most severe category, *critical violations*, as "items directly related to the protection of the public from foodborne illness or injury". *Major violations* are described as items that "if left un-addressed may lead to a situation detrimental to public health". The least severe category were originally called *minor violations*, but later renamed violations of *good food management practices*. For expositional ease, I call this category *good-practices violations* throughout.

Under the revisions adopted in March 2010, each critical violation carried 5 demerits, each major violation carried 3 demerits, and each good-practices violations carried 1 demerit. Imminent health hazards did not carry demerits, but rather, forced immediate temporary closure an X grade.

An interim period—in which establishments were inspected against the revised code,

but did not incur demerits on violations of new regulations—was in place for the remainder of 2010 and "much of 2011" per Southern Nevada Health District (2013); after which the revised regulations were fully implemented. Assessing SNHD inspection data suggests that the revisions were fully implemented by July 1, 2011. Along with these changes, a downgrade rule was adopted: establishments committing a particular critical or major violation in consecutive inspections are downgraded one letter. The downgrade rule has remained in effect ever since.

Beginning January 1, 2013, revisions were made to some of the SNHD inspection and scoring criteria. Intending to focus greater attention on foodborne-illness risk, the SNHD stopped assessing demerits on good-practices violations (although they continued to cite these violations in inspection reports). Along with these changes, some establishments (those with potential for five different foodborne-illness risk factors) that would have received a grade below A in their first inspection of 2013 were allowed a grace period with a follow-up inspection 15-30 days later. During this period, they would continue to post their old grade until their follow-up inspection. Unfortunately, the data do not indicate which establishments were eligible for this grace period, or what the original inspection performances might have been. However, this particular program was discontinued in early November 2013;[9] and more importantly, my results are ultimately robust to excluding inspections from 2013.

2.2 Inspection Process, Downgrades, and Re-inspections

SNHD food-service inspections follow a two-page "Food Establishment Inspection Form". The first and second pages of the current form are presented in Appendix Figures A1 and A2, respectively. Appendix Figures A3 and A4 present the first and second pages of the form from the period before the 2013 scoring change. Comparing these forms reveals some other slight changes. Compliance with the Nevada Clean Indoor Air Act was listed on the inspection report as critical violation prior to 2013, and removed from the form

[9]See page 22 of the SNHD report found here: https://media.southernnevadahealthdistrict.org/download/boh13/112613-hosr1.pdf.

after.[10] Also beginning in 2013: one major violation, related to the use of approved food and warewashing equipment was added; and two good-practices violations were removed. The SNHD criteria for inspecting, scoring, and grading food-service establishments have not been altered since the changes implemented in 2013.

In Appendix Figures A1 and A2, notice that the inspection form provides a checklist for each inspection. Listed first are the imminent health hazards followed by: a section listing all critical violations; a section listing all major violations; and finally, a section listing all good-practices violations. While inspectors may not be required to check for violations in the order listed on the inspection form, it is worth noting that the good-practices violations are after all others on the inspection form.

In some cases, SNHD food-service regulation is enhanced by dynamic-enforcement mechanisms, which enhance regulatory scrutiny based on prior non-compliance.[11] One such mechanism is a downgrade rule: any inspection involving a consecutive identical critical or major violation—*i.e.*, if the establishment commits the same major or critical violation for a second straight inspection—will be downgraded one letter.[12] Establishments must also be re-inspected within 15 business days of receiving a B or C grade. Establishments assessed more than 10 demerits in re-inspections are: downgraded to a C if the re-inspection follows a B grade, and closed with fines if the re-inspection follows a C grade. Closed establishments require re-inspection and SNHD approval before re-opening.

In addition to the potential cost of consumers responding to a poor inspection performance, grades below A also carry direct and non-trivial financial penalties. During this paper's sample period, establishments were required to pay: \$118 following a B grade, \$477 following a C grade, and \$716 following an X grade.[13]

[10] This violation was cited in only 1.44% of 23,381 pre-scoring-change inspections.

[11] For dynamic enforcement's effect on compliance in other settings, see Blundell (2020) and Blundell et al. (2020).

[12] Thus, when a consecutive identical major or critical violation is detected: 10 demerits or fewer result in a B, 11 to 20 demerits result in a C, and 21 to 40 demerits force closure and assessment of the closure fee.

[13] These fines have subsequently increased. Beginning February 1, 2021, the respective penalties were changed to \$143, \$1,200, \$1,400.

3 Data and Preliminary Analyses

Data are from the Southern Nevada Health District (2020) website.[14] Observations correspond to inspections and report (among other things) the: inspection date, establishment, codes for violations cited, total assessed demerits, letter grade, and an inspector identifier. The sample period analyzed spans July 1, 2011 (when the regulatory revisions discussed in Section 2 took full effect) through to March 9, 2020 (just before disruptions due to the COVID-19 pandemic).

As seen in Appendix Table A2, the SNHD inspects many different establishment types, including some which are quite obscure. To focus on establishment types that are fairly common, the primary analyses are restricted to routine inspections of establishments classified by the SNHD as: bars/taverns, buffets, food trucks/mobile vendors, restaurants, and snack bars. Descriptive statistics summarizing these inspections in the pre-change (July 2011 through December 2012 scoring regime) and post-change (current scoring regime) periods are reported in Table 2.

In inspection j of establishment i, let: $Base_{i,j}$ denote base demerits (those assessed on critical and major violations); $GP_{i,j}$ denote the reported number of good-practices violations; and $Demerits_{i,j}$ denote the official demerit total (under the scoring system in place at the time of the inspection). Good-practices violations carried 1 demerit prior to 2013, and—though still inspected and reported—0 demerits thereafter. As such,

$$Demerits_{i,j} = \begin{cases} Base_{i,j} + GP_{i,j}, & Pre_{i,j} = 1 \\ Base_{i,j}, & Pre_{i,j} = 0 \ , \end{cases} \tag{1}$$

where $Pre_{i,j}$ is a binary variable indicating an inspection occurred before January 1, 2013.

Figure 1 compares two distributions of the same inspection outcome, $(Base_{i,j} + GP_{i,j})$, in pre-change (red dots) versus post-change (navy diamonds) inspections. Dashed lines at 10, 20, and 40 demerits mark letter-grade thresholds. In pre-change inspections, the threshold value in each letter-grade interval (worst score still sufficient for that grade) is

[14] https://www.southernnevadahealthdistrict.org/permits-and-regulations/restaurant-inspections/developers/.

also the modal score. By contrast, the post-change distribution of $(Base_{i,j} + GP_{i,j})$—when good-practices violations carried no demerits—is much smoother through letter-grade thresholds. The striking contrast in these two distributions provides an initial piece of evidence suggesting that letter-grade implications affected the reported number of good-practices violations. In light of the coincident policy changes however, the post-change distribution is probably not an appropriate counterfactual for the pre-change distribution. Thus, I use the underlying data generating process to inform further analyses.

In inspection j of establishment i, let $GP^T_{i,j}$ denote the true number of good-practices violations present, and $GP^D_{i,j}$ denote the detected number of good-practices violations. We observe only the reported count, $GP_{i,j}$. Because $GP_{i,j}$ may sometimes understate $GP^T_{i,j}$ due to detection errors that are unrelated to the punishment schedule, I define *under-reporting* as when $GP_{i,j}$ understates $GP^T_{i,j}$ because of letter-grade implications.

There are two possible modes for under-reporting in the pre-change period. One, inspectors may chose to report $GP_{i,j} < GP^D_{i,j}$ because the demerits incurred on $GP^D_{i,j}$ would result in a lower letter grade and additional fines. Two, inspectors may report $GP^D_{i,j}$ but reduce or stop detection effort on good-practices violations when $Base_{i,j}$ is at or near a letter-grade threshold. Either way, letter-grade considerations cause $GP_{i,j} < GP^T_{i,j}$. In the post-change period by contrast, to the extent that $GP_{i,j}$ and $GP^T_{i,j}$ may differ, letter-grade considerations can be ruled out as a cause.

Figure 2 compares the relationship between base demerits and reported good-practices violations before (red dots) and after (navy diamonds) the 2013 scoring change.[15] The post-change period reveals a positive and fairly stable linear relationship. Across most pre-change base-demerit levels, a similar linear relationship is apparent, with a greater intercept and perhaps slightly steeper positive slope.[16] However—consistent with under-reporting—pre-change averages tend to diverge from that linear relationship as base-demerit values approach letter-grade thresholds from the left.

[15]The figure displays inspections with 30 or fewer base demerits, as these account for 98.6 percent of sampled inspections over this time, and there are relatively few observations at base-demerit levels beyond that.

[16]In part, this may be explained by the fact that two additional good-practices violations were defined in the pre-change period.

Among inspections with 6 base demerits or fewer, the relationship between $GP_{i,j}$ and $Base_{i,j}$ appears linear, and strikingly similar under both scoring regimes. Figure 3 reproduces these plots among inspections with 10 or fewer base demerits, and adds prediction lines from two simple ordinary least squares (OLS) regressions. The black prediction line comes from regressing $GP_{i,j}$ on $Base_{i,j}$ from post-change routine inspections with 10 or fewer base demerits. The maroon prediction line comes from regressing $GP_{i,j}$ on $Base_{i,j}$ from pre-change routine inspections with 6 or fewer base demerits. Among inspections with 8 to 10 base demerits the red pre-change dots are hollowed out to indicate that they were excluded from the regression producing the maroon prediction line. My methodology draws on information from these preliminary analyses, and especially from the relationship evident in Figure 3.

4 Methodology

To detect evidence of under-reporting in $GP_{i,j}$, I exploit the facts that potential letter-grade implications: (i) *never* exist in post-change inspections, and (ii) only exist in *some* pre-change inspections. In pre-change inspections, under-reporting only has the potential to affect letter grade if $GP^T_{i,j}$ exceeds the margin between $Base_{i,j}$ and the next letter-grade threshold (which I denote $margin_{i,j}$).[17] Though $GP^T_{i,j}$ is not observable, $margin_{i,j}$ is, and can be used to account for the likely presence of letter-grade implications.

Figures 2 and 3 suggest that the expectation of $GP^T_{i,j}$ increases with $Base_{i,j}$. In post-change inspections, any difference between the true and reported counts of good-practices violations will be due to detection errors or other factors unrelated to letter-grade considerations. If the probability of missing the commission of any particular good-practices violation is small, then the post-change averages of $GP_{i,j}$ in Figure 2 suggest a positive and linear relationship between base demerits and $GP^T_{i,j}$. In pre-change inspections where letter-grade implications were unlikely, a similar marginal relationship is observed. Thus, after an intercept shift to account for the effect of coincident policy changes, the slope

[17] For example, if $Base_{i,j} = 12$ (which already exceeds the threshold for an A grade), then $margin_{i,j} = 20 - 12 = 8$.

of $GP_{i,j}$ in $Base_{i,j}$ from the post-change period can provide an estimate of $GP^T_{i,j}$ against which under-reporting in the pre-change period can be detected.

I focus initially on inspections with 10 or fewer base demerits, which account for 86.16% of all routine inspections. Among these inspections, a natural break exists between 6 and 8 base demerits. $Base_{i,j} = 7$ is not possible under the scoring system, and the probability that $GP^T_{i,j} > margin_{i,j}$ likely increases considerably as base demerits change from 6 to 8. Across the entire sample, 5 or more good-practices violations were reported in only 2.46 percent of all inspections, and only 1.11 percent of inspections with 10 or fewer base demerits. Whereas 3 to 4 good-practices violations were reported in 12.62 percent of all inspections, and 10.12 of inspections with 10 or fewer base demerits.

I refer to an establishment as being "on the bubble" in an inspection when $8 \leq Base_{i,j} \leq 10$, and construct binary variable, $Bubble_{i,j}$, indicating when this is the case. I test for evidence of under-reporting with a baseline specification of

$$GP_{i,j} = \beta_1 \left(Bubble_{i,j} \times Pre_{i,j}\right) + \beta_2 Pre_{i,j} + \beta_3 Base_{i,j} + \mathbf{X}'_{i,j}\boldsymbol{\omega} + \epsilon_{i,j}, \tag{2}$$

where $Pre_{i,j}$ is a binary variable indicating an inspection occurred before January 1, 2013. The inclusion of $Base_{i,j}$ accounts for,

$$\beta_3 = \frac{\partial \mathbb{E}\left(GP_{i,j} \mid Bubble_{i,j} \times Pre_{i,j} = 0, \mathbf{X}_{i,j}\right)}{\partial Base_{i,j}},$$

a common slope parameter across pre-change *off-bubble* inspections, and all post-change inspections. Under a full specification, the vector of controls, $\mathbf{X}_{i,j}$, contains fixed effects for the month of the year and day of the week in which an inspection occurred. Inspector fixed effects are also added in a more limited sample that permits their inclusion.[18]

The parameter of interest, β_1, is the difference between expected $GP_{i,j}$ in pre-change bubble inspections, and post-change bubble inspections; after accounting for base-demerit-invariant differences across the two time periods, and a common underlying slope. An

[18]This limited sample consists of 49,929 inspections involving inspectors who—among routine inspections with 10 or fewer base demerits—conducted at least: one pre-change bubble inspection, one pre-change off-bubble inspection, one post-change bubble inspection, and one post-change off-bubble inspection.

identifying assumption underlying this specification is that expected $GP^T_{i,j}$ exhibits a common slope in $Base_{i,j}$ across the pre and post-change periods, and that deviations of $GP_{i,j}$ from this slope will not occur absent letter-grade influences. If this assumption holds, then under a null hypothesis that letter-grade considerations do not affect the reporting of good-practices violations, $\beta_1 = 0$. I also estimate two alternative specifications which allow some relaxation of this assumption.

The first alternative specification allows bubble inspections to break from the common slope in both periods,

$$GP_{i,j} = \gamma_1 \left(Bubble_{i,j} \times Pre_{i,j}\right) + \gamma_2 Pre_{i,j} + \gamma_3 Bubble_{i,j} + \gamma_4 Base_{i,j} + \mathbf{X}'_{i,j}\boldsymbol{\omega} + \epsilon_{i,j}. \quad (3)$$

Further, if the underlying relationship between $Base_{i,j}$ and expected $GP_{i,j}$ changed in 2013 such that post-change inspections exhibited a steeper positive slope in base demerits, then estimates from equations (2) and (3) would overstate any potential under-reporting. Thus, I also consider a second alternative which allows for different slopes (in base demerits) in the pre-change and post-change periods,

$$\begin{aligned} GP_{i,j} = \delta_1 \left(Bubble_{i,j} \times Pre_{i,j}\right) + \delta_2 Pre_{i,j} + \delta_3 Bubble_{i,j} \\ + \delta_4 \left(Pre_{i,j} \times Base_{i,j}\right) + \delta_5 Base_{i,j} + \mathbf{X}'_{i,j}\boldsymbol{\omega} + \epsilon_{i,j}. \end{aligned} \quad (4)$$

If the scoring change and other revisions led to $\delta_4 < \delta_5$, this specification controls for such differences, preventing them from being projected on to $\widehat{\delta}_1$.

4.1 Testing the Common Slope Assumption

The methodology for detecting under-reporting adopted in equation (2) rests on the assumption that the relationship between the conditional expectation of $GP^T_{i,j}$ and $Base_{i,j}$ possesses a common slope. To assess the validity of this assumption, I estimate

$$GP_{i,j} = \alpha_1 \left(Base_{i,j} \times Pre_{i,j}\right) + \alpha_2 Pre_{i,j} + \alpha_3 Base_{i,j} + \mathbf{X}'_{i,j}\boldsymbol{\omega} + u_{i,j}, \quad (5)$$

using pre-change off-bubble inspections and post-change inspections. Because letter-grade implications are either highly unlikely or impossible among these observations, $GP_{i,j}$ can be used to test the underlying assumption. Under the null hypothesis of the common slope assumption, $\alpha_1 = 0$.

Table 3 reports estimates of equation (5) using pre-change inspections with 6 base demerits or fewer, and post-change inspections with 10 base demerits or fewer. Appendix Table A3 reports similar estimates using pre-change and post-change inspections with 6 or fewer base demerits only. In both tables, coefficients on $(Pre_{i,j} \times Base_{i,j})$ are very similar across the different sets of controls and estimating samples; and relatively small, ranging from 0.0207 to 0.0217. They are different from zero at the 90 percent significance levels; but more importantly, they are positive and suggest that the marginal effect of base demerits on expected $GP_{i,j}$ was slightly greater in the pre-change period. Thus, by assuming a common slope, equations (2) and (3) will provide smaller counterfactual predictions for $GP_{i,j}$ in pre-change bubble inspections; and, if anything, potentially understate the true extent of under-reporting.

5 Results

5.1 Primary Estimates

Table 4 presents estimates of equation (2) using inspections with 10 or fewer base demerits. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.[19] Column (1) reports estimates from a simple specification in which no additional controls are included. In column (2), fixed effects for the inspection's day of the week and month of the year are added. In columns (3) and (4), inspector fixed effects are added to the specifications reported in columns (1) and (2), respectively.

Across all four sets of controls, the coefficient of interest suggests that good-practices violations were significantly and substantially under-reported due to letter-grade considerations. Using only post-change inspections and pre-change off-bubble inspections,

[19]Two-way clustering of standard errors is performed with the `vcemway` command by Gu and Yoo (2019) in Stata 17 (StataCorp, 2021).

estimates of these four specifications yield average predictions of $GP_{i,j}$ for pre-change bubble inspections of: 2.0319, 2.0308, 1.9723, and 1.9721. As such, Table 4 estimates suggest that, in the pre-change bubble inspections, the number of good-practices violations reported were 37.5 to 39.3 percent lower than amounts consistent with inspections where letter-grade considerations are unlikely (pre-change off-bubble) or certain not to (post-change) affect reporting decisions.

Table 5 reports estimates of equation (2) with the same estimating samples and specifications, but separately interacts $Pre_{i,j}$ with binary variables indicating that $Base_{i,j}$ equals 8, 9, and 10. Consistent with letter-grade considerations driving the observed effect, the extent of under-reporting estimated increases considerably as base demerits approach the letter-grade threshold at 10 demerits. Reported good-practices violations understate predicted values by: 31.0 to 32.9 percent in pre-change inspections with 8 base demerits, 43.9 to 45.4 percent among pre-change inspections with 9 base demerits, and 90.0 to 90.3 percent among pre-change inspections with 10 base demerits.

Table 6 reports estimates of equation (3). Post-change deviations from the underlying relationship between base demerits and reported good-practices violations among bubble inspections are negative, but relatively very small and not significantly different from zero at conventional levels. After accounting for these slight differences, coefficients of interest suggest that in the pre-change bubble inspections, the number of good-practices violations reported were 36.8 to 39.0 percent lower than they would have been absent potential letter-grade implications.

Finally, recall that equation (4) would allow for the marginal effect of base demerits on expected $GP_{i,j}$ to differ across the two scoring regimes. However, the tests performed in Section 4.1 reveal that the marginal effect of base demerits was, if anything, slightly larger in the pre-change period. Thus, as seen in Appendix Table A4, estimates of equation (4) yield slightly larger coefficients of interest than the other two estimating equations.

5.2 Robustness Analyses

As discussed in Section 2, select establishments (those with potential for five different foodborne-illness risk factors) that would have received a grade below *A* in their first inspection of 2013 were allowed a grace period with a follow-up inspection 15-30 days later. Establishments would keep their old grade posted until the follow-up inspection. It is unclear which establishments qualified for this or ultimately used it. To test whether my primary results are robust to any coincident effects that the program may have had, Appendix Table A6 reports estimates of equations (2) and (3) excluding inspections from 2013. Across all specifications, coefficients of interest are *very* similar to their full-sample counterparts (ranging from 0.08 percent closer to 3.25 percent further from zero). Appendix Table A5 reports common-slopes tests for these specifications with the exclusion of inspections from 2013. In all cases, the marginal effect of base demerits on $GP_{i,j}$ is slightly greater in the pre-change period; meaning that estimates of equations (2) and (3) will, if anything, slightly understate the extent of under-reporting.

Appendix Table A8 reports estimates of equations (2) and (3) with the estimating sample expanded to include routine inspections from all establishment types. Coefficients of interest are similar to estimates from the preferred sample in sign, significance, and magnitude. Appendix Table A7 reports common-slope tests for these specifications with the inclusion of all establishment types. Again, the marginal effect of base demerits on $GP_{i,j}$ is slightly greater in the pre-change period across all specifications.

5.3 An Additional Approach to Detecting Under-reporting

My primary methodology exploits the fact that, as base demerits increase from 6 to 8, the probability that the true count of good-practices violations exceeds $margin_{i,j}$ likely jumps substantially. Thus, the effect of coincident changes on the expected $GP^T_{i,j}$ can accounted for by comparing $GP_{i,j}$ between off-bubble pre and post-change inspections. Yet, there may also be some off-bubble inspections in which $GP^T_{i,j} > margin_{i,j}$.

I provide additional evidence of under-reporting by utilizing two implications of the letter-grade system on pre-change inspections where $GP^T_{i,j} > margin_{i,j}$. First, under-

reporting in these inspections will only affect letter grades when reported $GP_{i,j} \leq margin_{i,j}$. Second, any under-reporting of good-practices violations beyond the margin (such that $GP_{i,j} < margin_{i,j}$) is inconsequential. Thus, if under-reporting occurs, it will most likely result in $GP_{i,j} = margin_{i,j}$. I construct a binary variable indicating this and estimate

$$\begin{aligned} \mathrm{I}\left(GP_{i,j} = margin_{i,j}\right) = &\left[\sum_{k=3}^{10} \zeta_k \times \mathrm{I}\left(Base_{i,j} = k\right) \times Pre_{i,j}\right] \\ &+ \left[\sum_{l=5}^{10} \eta_l \times \mathrm{I}\left(Base_{i,j} = l\right)\right] + \mathbf{X}_{i,j}^{'}\boldsymbol{\omega} + \epsilon_{i,j}, \end{aligned} \tag{6}$$

using routine inspections with at least 3,[20] and no more than 10, base demerits.

Estimates of equation (6) are reported in Table 9. Consistent with the primary estimates from Section 5.1, the probability that $GP_{i,j} = margin_{i,j}$ is substantially higher among pre-change bubble inspections than among observably similar post-change inspections. The coefficients in column (2) imply that $GP_{i,j} = margin_{i,j}$ was 134.4, 158.7, and 181.6 percent more likely in the pre-change inspections with 8, 9, and 10 base demerits respective, relative to post-change counterparts. Conditional on 5 or 6 base demerits, $GP_{i,j} = margin_{i,j}$ is also significantly more likely in the pre period, suggesting that there may have been some off-bubble pre-change inspections where letter-grade implications were present, and under-reporting might have occurred.[21]

Appendix Figure A6 graphically demonstrates this evidence. Solid red circles and navy diamonds mark simple pre and post-change averages of $\mathrm{I}\left(GP_{i,j} = margin_{i,j}\right)$. Pre-change averages are greater at most base demerit levels, with the difference in averages becoming larger as *margin* decreases. Further, note that every grade threshold is a multiple of 10 demerits, but there is no letter-grade threshold at 30 demerits. I generate a binary indicator, $\mathrm{I}\left(GP_{i,j} = [30 - Base_{i,j}]\right)$, for inspections where $Base_{i,j} \in [22, 30]$. Hollow red circles and navy diamonds mark pre and post-period averages of this indicator variable among those inspections. A light-blue dashed line at 30 demerits marks this

[20] Only 9 good-practices violations defined in the post-change period, hence the restriction to inspections with at least 3 base demerits.

[21] Using all inspections with fewer than 41 base demerits and $margin_{i,j} \leq 8$, Appendix Table A9 reports similar estimates but without interacting $Pre_{i,j}$ with base-demerit indicators. Across all base demerit levels, $GP_{i,j} = margin_{i,j}$ was significantly and substantially more likely.

10-demerit multiple that is not a grade threshold. Notably, the pattern observed between pre and post inspections in leading to the other 10-demerit multiples is not seen here.

A potential issue with this approach is that $GP_{i,j} = margin_{i,j}$ may be more likely pre-change because there were more good-practices violations defined. If that—independent of letter-grade considerations—is partially driving these results, then *cet. par.*, outcomes where $GP_{i,j} > margin_{i,j}$ should be more likely as well. Appendix Table A10 reports estimates analogous to those in Appendix Table A9, but with a dependent variable of $\mathrm{I}\left(GP_{i,j} > margin_{i,j}\right)$, and shows that this was significantly *less* likely prior to 2013.

5.4 Inspector-level Analyses

The results thus far show a common relationship between $GP_{i,j}$ and $Base_{i,j}$ in both periods when letter-grade implications are absent, or very likely absent; and that in reported $GP_{i,j}$ diverges distinctly from that relationship in pre-change bubble inspections. An interesting question though is how heterogeneous the responses to letter-grade implications might be across individual inspectors. Are results driven by a small set of inspectors, or were most inspectors reluctant to have good-practices violations cause significant increases in punishment severity?

To analyze responses at the inspector level, I include all establishment types in the interest of individual-inspector sample size. Focusing on routine inspections with 10 or fewer base demerits, there are 40 inspectors who conducted at least 25 inspections in each of the following categories: pre-change off-bubble inspections, pre-change bubble inspections, post-change off-bubble inspections, and post-change bubble inspections. Among these 40 inspectors, I estimate the equation

$$GP_{i,j} = \theta_{k,1}\left(Bubble_{i,j} \times Pre_{i,j}\right) + \theta_{k,2} Pre_{i,j} + \theta_{k,3} Base_{i,j} + \mathbf{X}_{i,j}^{'}\boldsymbol{\omega_k} + \epsilon_{i,j}, \qquad (7)$$

separately for each inspector $k = 1, \ldots, 40$. The vector, $\mathbf{X}_{i,j}$, contains fixed effects for day of the week and month of the year.

For inspector k, $\widehat{\theta}_{k,1}$ estimates how sensitive their reporting of $GP_{i,j}$ is to the likely

presence of letter-grade implications. By estimating equation (7) separately for each inspector, $\widehat{\theta}_{k,1}$ provides a test of under-reporting that is based on internal consistency. All else the same, does ($Bubble_{i,j} \times Pre_{i,j}$) affect the conditional expectation of $GP_{i,j}$ reported by inspector k? Under a null hypothesis of internally consistent reporting by inspector k, independent of likely letter-grade implications, $\theta_{k,1} = 0$.

The frequency distribution of $\widehat{\theta}_{k,1}$, is shown in Figure 4. Bin width is 0.05, and beige bars display the frequency $\widehat{\theta}_{k,1}$ among inspectors where the null hypothesis is rejected at the 95-percent significance level. Black drop-lines show the frequency $\widehat{\theta}_{k,1}$ when estimates fail to reject the null hypothesis at the 95-percent significance level. Inference is based on standard errors clustered two-way on the establishment inspected and the month of the sample in which an inspection occurred. Of the 40 inspectors evaluated, 33 reject the null hypothesis of no under-reporting.

Next, I group inspectors based on whether their individual tests rejected the null hypothesis. Figure 5 compares the relationship between base demerits and $GP_{i,j}$ between each group. OLS prediction lines come from regressing $GP_{i,j}$ on $Base_{i,j}$ and $Pre_{i,j}$ within each group, using pre-change off-bubble and post-change inspections. It appears that both groups tend to under-report when an establishment has 10 base demerits.[22] However among the group where individual tests fail to reject the null, pre-period reporting of $GP_{i,j}$ at 8 and 9 base demerits is much more consistent with the relationship seen elsewhere.

Interestingly, these findings reveal under-reporting in the pre-change period was quite prevalent among individual inspectors, and that internally consistent reporting of $GP_{i,j}$ across both scoring regimes was something of an exception. While the deliberations behind the scoring change are unknown, it is noteworthy that a large majority of the evaluated inspectors exhibit under-reporting in the pre-change period, and the scoring change implemented in 2013 served to better align SNHD regulation with these inspectors' apparent discomfort with a good-practices violation triggering significant penalty escalation.

[22]Note however that there are only 6 pre-period observations where $Base_{i,j} = 10$ among the fail-to-reject group.

6 Addressing a Potential Alternative Explanation

The results presented in Section 5 have been interpreted as evidence of letter-grade implications affecting inspector reporting decisions. Another possibility is that optimization by some establishments explains the sharp decline in good-practices violations as pre-period base demerits approach the A-grade threshold. In the pre-change period, perhaps certain combinations of critical and/or major violations that put establishments on the bubble, coupled with relatively few good-practices violations would achieve an A grade at minimum compliance cost? In this section, I consider this possible alternative explanation, and present several pieces of evidence suggesting that it is unlikely.

One point in favor of the under-reporting interpretation is an appeal to plausibility. For an establishment, successfully targeting an inspection score on the bubble requires both detailed knowledge of the health code, as well as fine control of overall compliance at any time. In the pre-change period, establishments incurred demerits for committing any of 34 different violations (10 critical, 13 major, and 11 good-practices). Note also that an establishment targeting an A grade on the bubble would receive a B grade by unintentionally committing any of the 23 critical and major violations specified at that time. Thus, to contribute to the effects found in Section 5, a meaningful number of establishments would need to not only possess fine control of their overall compliance and very detailed knowledge of the health code and scoring system, but they would also need to be comfortable with the risk inherent to this approach.

Another detractor from the alternative explanation's plausibility is the SNHD downgrade rule. An establishment committing the same critical or major violation in consecutive inspections is downgraded one letter. As such, for an individual establishment, the cost-minimizing combination of violations sufficient for an A grade will change with each inspection. Not only does this add to the complexity of making cost-minimizing compliance decisions, it also reduces the potential gains of doing so because an establishment's cost-minimizing set of violations sufficient for an A grade, only suffices for an A grade in every other routine inspection.

Setting aside the discussion of plausibility, there exist several pieces of empirical evi-

dence which are inconsistent with this alternative explanation. First, within the primary estimating sample there are 1,670 establishments that made an A grade on the bubble in the pre-change period, and are observed in more than one pre-change routine inspection. Among them: 1,431 (about 85.69 percent) made an A grade on the bubble only once in the pre-change period; 229 made an A did so twice in the pre-change period; and 10 did so three times. If on-bubble A grades were optimal, and some establishments possessed the ability to successfully target them, one would expect them to want—and be able—to repeat such performances, yet this is relatively uncommon.

Moreover, did establishments that made pre-period A grades on the bubble achieve better grades in the post period than observably similar counterparts? I compute average pre-change scores for primary-sample establishments with more than one pre-change inspection and at least one post-change inspection. Restricting consideration to pre-change averages that allow comparison: there are 2,250 establishments that made pre-change A grades on the bubble, and 4,847 establishments that did not. Using these establishments, I estimate

$$\mathrm{I}\left(Grade_{i,j} < A\right) = \kappa_1 PreBubbleA_i + \mathbf{X}_{i,j}^{'}\boldsymbol{\omega} + \epsilon_{i,j}. \tag{8}$$

The dependent variable indicates that a letter grade of B or worse was made in inspection j of establishment i; $PreBubbleA_i$ indicates that establishment i made an A grade on the bubble in at least one pre-period inspection; and fixed effects for pre-change average score are included as controls.

Estimates of equation (8) are reported in Appendix Table A11. The estimates of interest use post-change observations, and are reported columns (1) and (2). Holding pre-change average score fixed, establishments that made pre-change A grades on the bubble were not significantly less likely to make sub-A grades in the post-change period. This counters what one would expect if the primary results stemmed from establishments able to target A grades at minimum compliance cost. Estimates in columns (3) and (4) use pre-change observations and show that, holding pre-change average score fixed, establishments that made pre-change bubble A grades, were significantly less likely to make sub-A grades in the pre-change period. This unsurprising result is reported simply

to verify that, conditional on pre-change average score fixed effects, *PreBubbleA* is not inadvertently selecting on propensity for sub-A grades. Given that A grades were slightly easier to attain in the post-change period; the findings in columns (1) and (2) further suggest that this paper's main results reflect inspector responses to inspection-specific conditions, rather than the strategic responses of a select group of food establishments.

Finally, results from Table A9 suggest under-reporting even near the B and C-grade thresholds, where establishment optimization can likely be ruled out. As an extension, I apply the original methodology to inspections with 30 base demerits or fewer. I treat inspections where $17 \leq Base_{i,j} \leq 20$ as on the bubble, and estimate

$$\begin{aligned} GP_{i,j} = \lambda_1 \left(BubbleA_{i,j} \times Pre_{i,j}\right) + \lambda_2 \left(BubbleB_{i,j} \times Pre_{i,j}\right) \\ + \lambda_3 Pre_{i,j} + \lambda_4 Base_{i,j} + \mathbf{X}_{i,j}^{'}\boldsymbol{\omega} + \epsilon_{i,j}, \end{aligned} \tag{9}$$

where $BubbleA_{i,j}$ indicates that $8 \leq Base_{i,j} \leq 10$, and $BubbleB_{i,j}$ indicates $17 \leq Base_{i,j} \leq 20$. Estimating equation (5) using post-change inspections with 30 or fewer base demerits, and off-bubble pre-change inspections with 30 or fewer base demerits; tests for a common underlying slope in this wider range. These estimates, reported in Appendix Table A12, reject a common slope at the 99 percent significance level; but as before, this rejection is due to a *steeper* slope in pre-change inspections, and estimates of equation (9) will—if anything—understate the extent of under-reporting.

Estimates of equation (9) are reported in columns (1) and (3) of Table 8. Columns (2) and (4) report estimates of this equation with the inclusion of *BubbleB* and *BubbleA* to allow for divergence from the common slope on the bubble under both scoring regimes. Both sets of estimates suggest significant and substantial under-reporting on the bubble of a B grade—albeit to an extent absolutely and relatively smaller than on the A-grade bubble—and provide conservative estimates in light of results from Table A12. It is quite unlikely that declines in pre-change $GP_{i,j}$ as base demerits approach the B-grade threshold reflect deliberate efforts by establishments.

7 Concluding Remarks

Mandatory disclosure of inspection performance is an increasingly popular supplement to food-service health inspection programs. These policies indirectly raise the penalty and expected cost of regulatory violations, as the information provided may cause consumers to substitute away from less compliant firms. In general, if disclosed information is capable of inducing an intended response by regulated entities, it may also create incentives among them to manipulate the information. When these incentives are strong and regulated entities are able, manipulation can occur. When disclosing inspection-performance information, this concern is mitigated by the fact that inspectors report the performance, and firms can't directly manipulate the information; but inspectors can manipulate this information, and firms may try to influence inspectors if stakes are sufficiently high.

The Southern Nevada Health District (SNHD) supplements its food establishment inspection program with mandatory disclosure. Detected violations carry prescribed demerits which aggregate into an inspection score, but only letter grades (which pool scores within wide intervals) are disclosed. The SNHD also fines establishments for non-compliance, but these fines are incurred only with changes in letter grade. As such, marginal punishment on an additional violation is zero, except when crossing letter-grade thresholds. The sharp discontinuities in punishment severity that result could induce under-reporting of violations by inspectors who hold sympathy or other interests toward establishments they inspect. Coarsening of numeric inspection scores into letter grades for disclosure is very common in food-service hygiene regulation. Yet testing for inspector under-reporting in response to these designs is—by nature—often prohibitively difficult. A small revision to SNHD scoring criteria provides a rare chance to test for and detect evidence of under-reporting.

Prior to 2013, the SNHD assigned 5, 3, 1 demerits each on critical, major, good-practices violations respectively. Beginning in 2013, though still inspected, reported, and corrected, the demerit penalty on good-practices violations has been removed. After accounting for coincident changes affecting all inspections and conditioning on controls, reported counts of good-practices violations very similar across pre and post-change in-

spections; except when demerits accrued on critical and major violations are at or near the next letter-grade threshold, where reported counts understate projected true counts by 30 to 90 percent. Several findings suggest that these sharp pre-change declines are due to under-reporting on the part of inspectors.

Penalty schedules that sharply increase punishment severity about arbitrary cutoff points can lead to selective reporting of violations by inspectors. In particular, mandatory disclosure policies that coarsen a quality signal, like the letter-grade-only disclosure policies common in restaurant hygiene regulation, can potentially undermine the regulatory programs they are meant to strengthen.

References

Blundell, W. (2020). When threats become credible: A natural experiment of environmental enforcement from Florida. *Journal of Environmental Economics and Management 101*, 102288.

Blundell, W., G. Gowrisankaran, and A. Langer (2020). Escalation of scrutiny: The gains from dynamic enforcement of environmental regulations. *American Economic Review 110*(8), 2558–2585.

Dai, W. and M. Luca (2020). Digitizing disclosure: The case of restaurant hygiene scores. *American Economic Journal: Microeconomics 12*(2), 41–59.

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.

Gu, A. and H. I. Yoo (2019). vcemway: A one-stop solution for robust inference with multiway clustering. *Stata Journal 19*(4), 900–912.

Jin, G. Z. and J. Lee (2018). A tale of repetition: Lessons from Florida restaurant inspections. *Journal of Law and Economics 61*(1), 159–188.

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.

Jin, G. Z. and P. Leslie (2005). The case in support of restaurant hygiene grade cards. *Choices: The Magazine of Food, Farm, and Resource Issues 20*(2), 93482.

Johnson, M. S. (2020). Regulation by shaming: Deterrence effects of publicizing violations of workplace safety and health laws. *American Economic Review 110*(6), 1866–1904.

Makofske, M. P. (2020a). The effect of information salience on product quality: Louisville restaurant hygiene and Yelp.com. *The Journal of Industrial Economics 68*(1), 52–92.

Makofske, M. P. (2020b). Mandatory disclosure, letter-grade systems, and corruption: The case of Los Angeles County restaurant inspections. *Journal of Economic Behavior and Organization 172*(C), 292–313.

Southern Nevada Health District (2013). 2013 Think Risk initiative and audit update. Technical report. Retrieved from http://media.southernnevadahealthdistrict.org/download/boh13/112613-hosr1.pdf.

Southern Nevada Health District (2020). Entire restaurant/food establishment database. Retrieved from https://www.southernnevadahealthdistrict.org/permits-and-regulations/restaurant-inspections/developers/.

StataCorp (2021). Stata statistical software: Release 17.

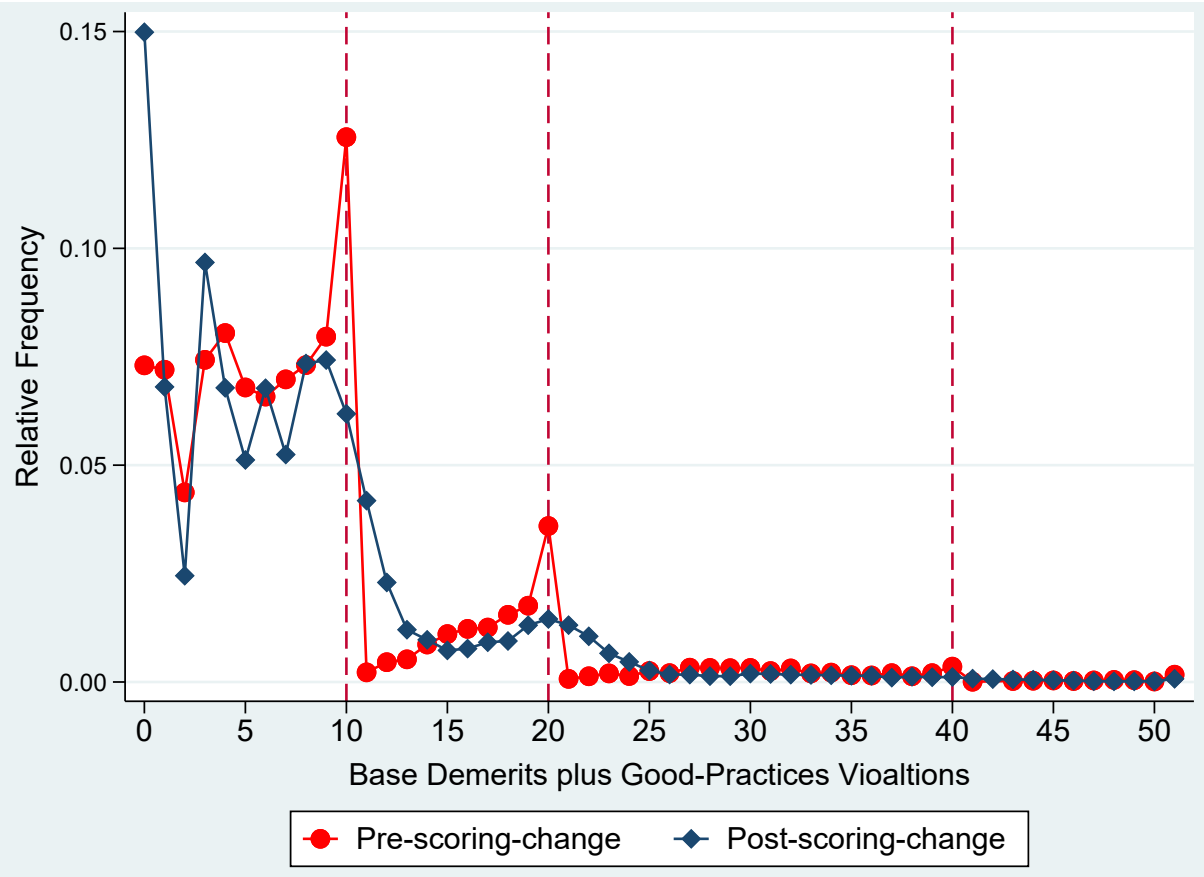


Figure 1: Inspection Totals Under Initial Scoring Rules

Distribution of demerit totals under initial scoring rules from 16,353 routine inspections in the pre-scoring-change period, and 92,063 routine inspections in the post-scoring-change period. Dashed lines at 10, 20, and 40 demerits mark the threshold values for A, B, and C grades, respectively. All totals greater than 50 are included at 51.

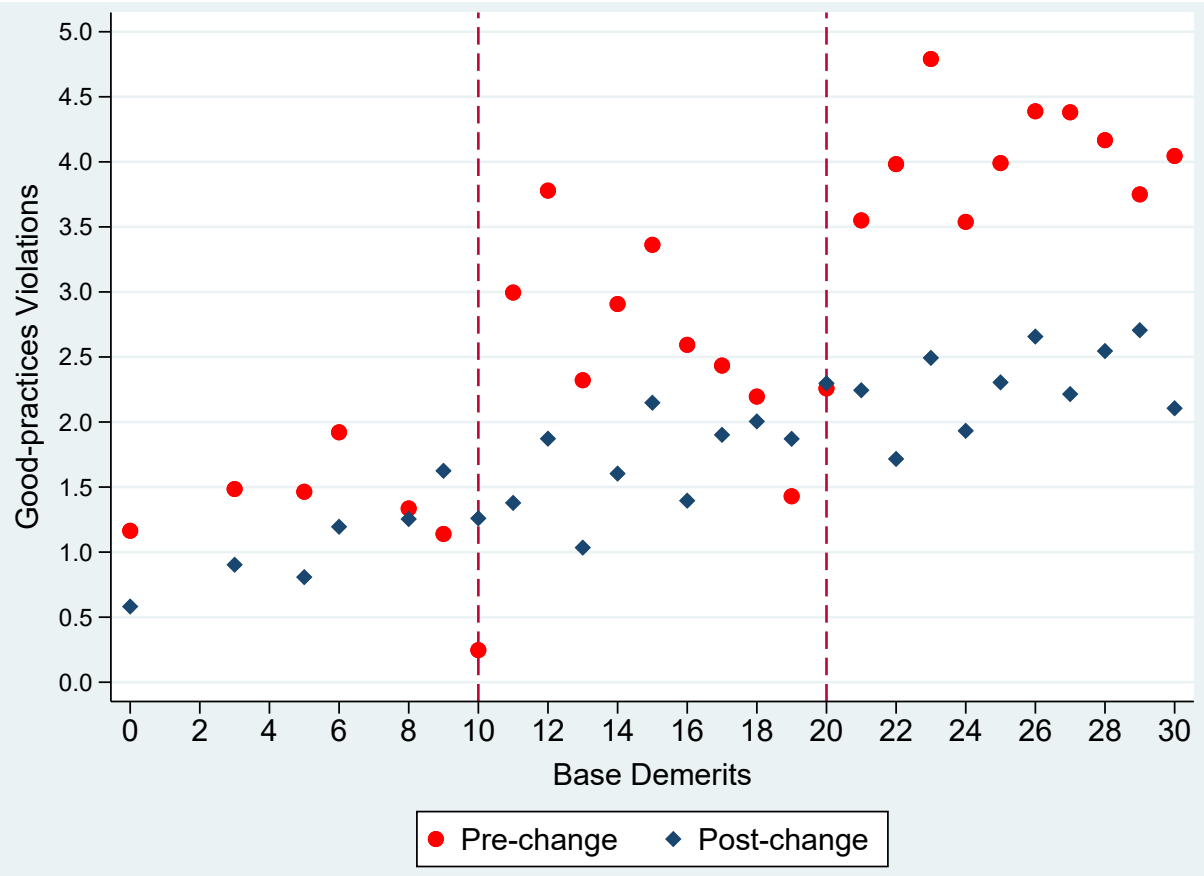


Figure 2: Reported Good-practices Violations by Base Demerits

Red dots mark average good-practices violations from 16,398 routine inspections with 30 base demerits or fewer from the pre-change period. Navy diamonds mark average good-practices violations from 90,775 routine inspections with 30 base demerits or fewer from the post-change period. Dashed lines at 10 and 20 demerits mark the threshold values for A and B grades, respectively.

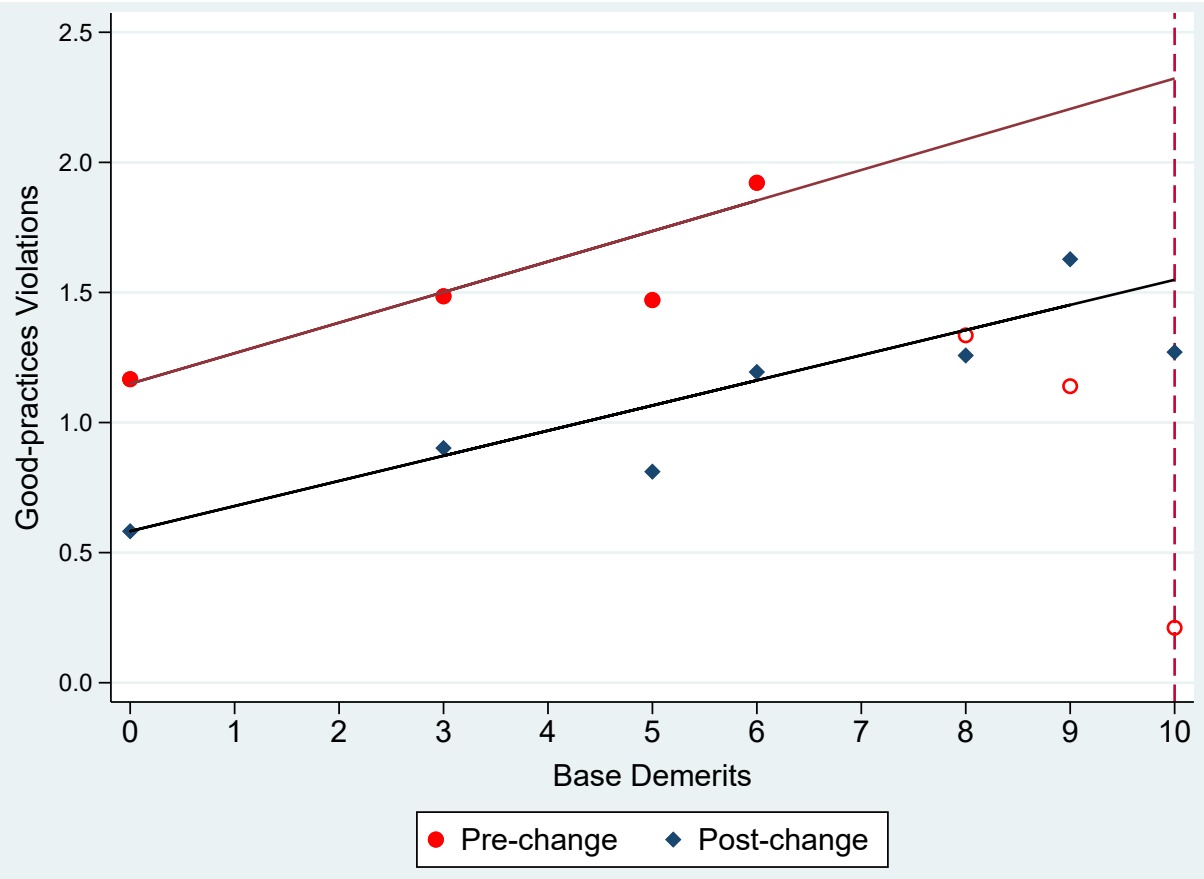


Figure 3: Reported Good-practices Violations by Base Demerits with OLS Predictions

The maroon line marks predicted values from regressing good-practices violations on base demerits from 11,027 pre-change routine inspections in which base demerits were less than or equal to 6. The navy line marks predicted values from regressing good-practices violations on base demerits from 79,544 post-change routine inspections in which base demerits were less than or equal to 10. The dashed line at 10 demerits marks the threshold value for an *A* grade. Red dots mark averages from the pre-change period. Navy diamonds mark averages from the post-change period. Red dots from pre-change bubble inspections are hollowed out to indicate these observations were excluded from the regression producing the maroon prediction line.

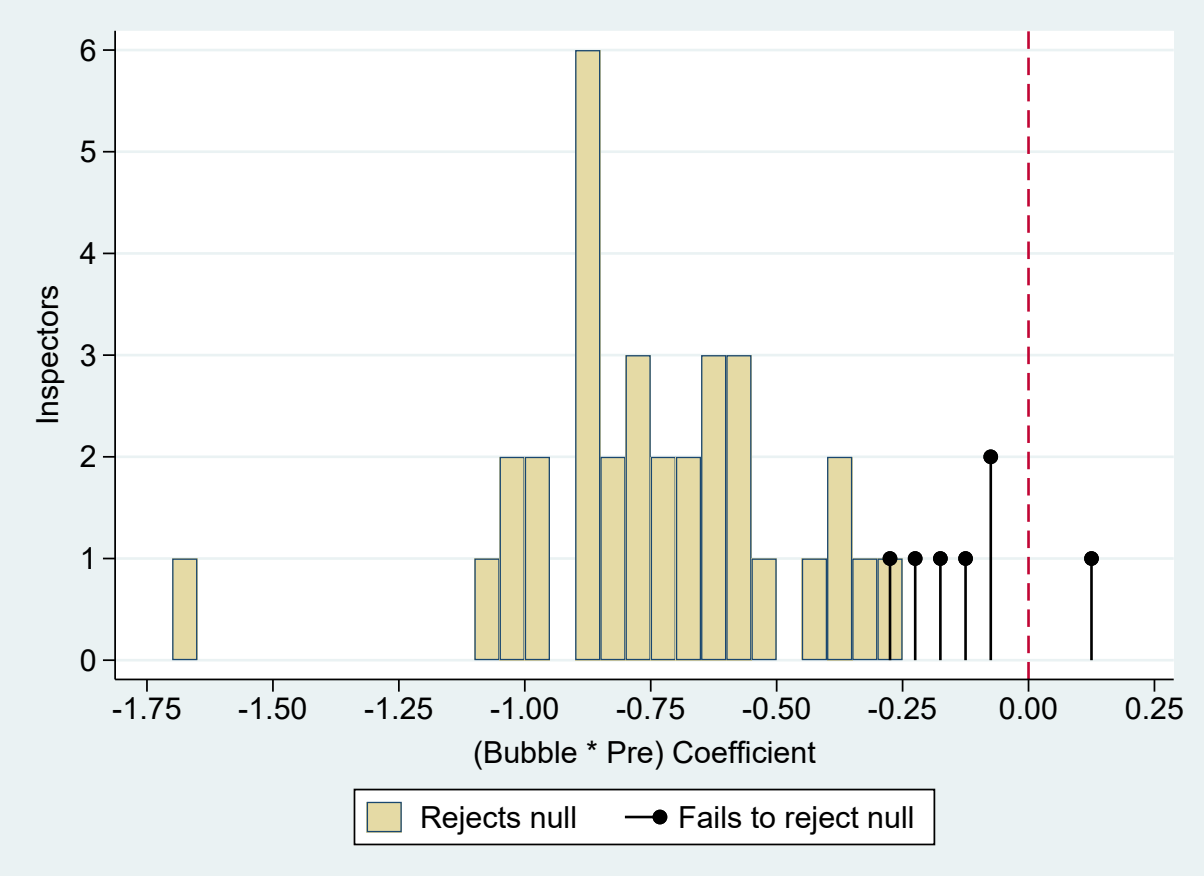


Figure 4: Frequency of Inspector Coefficients

Frequency distribution of $\widehat{\theta}_{k,1}$ from separately estimating equation (7) for each of 40 inspectors sampled. Bin width is 0.05. Solid beige bars show the frequency of $\widehat{\theta}_{k,1}$ when estimates reject the null hypothesis ($\theta_{k,1} = 0$) at the 95% significance level. Black drop-lines show the frequency of $\widehat{\theta}_{k,1}$ when estimates fail reject the null hypothesis at the 95% significance level. Hypothesis tests used standard errors clustered multiway on the establishment inspected and the month of the sample.

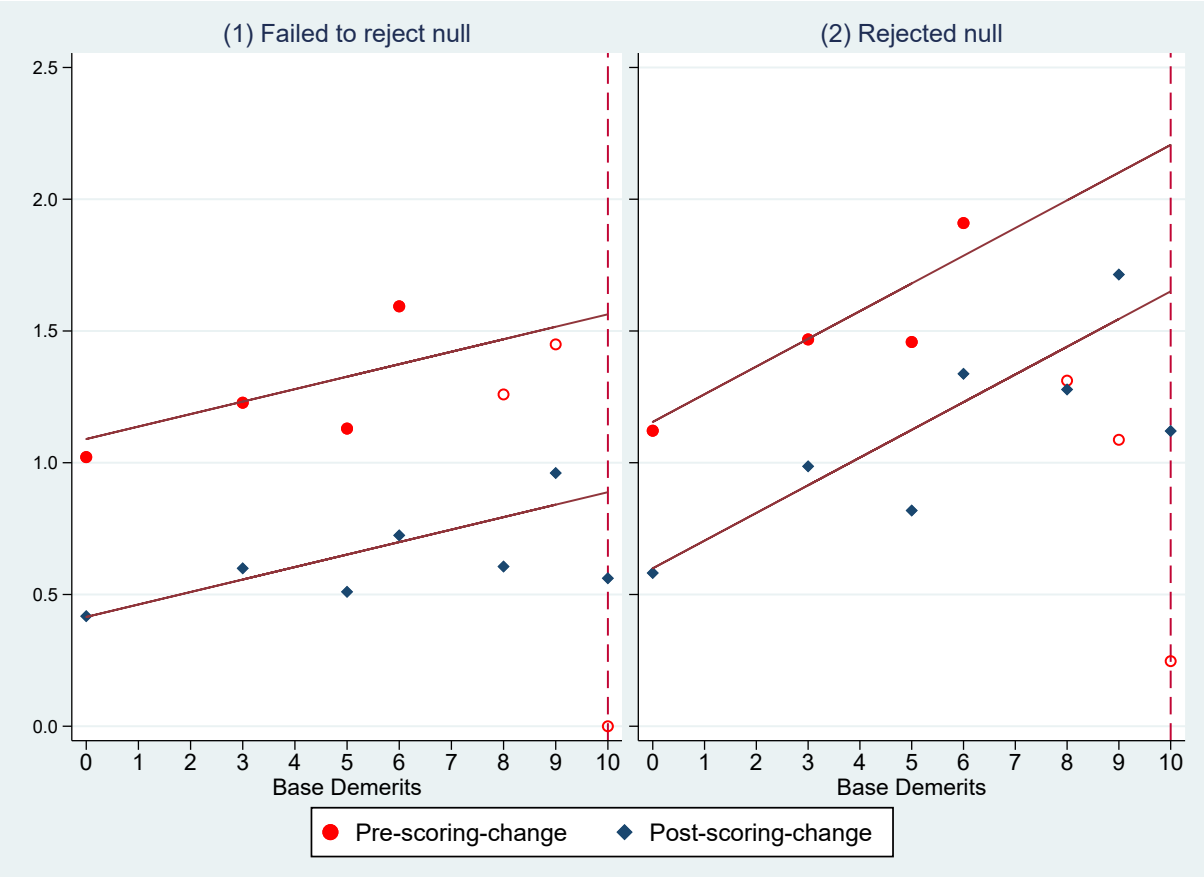


Figure 5: Reported Good Practices Violations by Base Demerits: Inspector Group Comparison

Graph (1) shows averages from 12,921 inspections conducted by 7 inspectors for whom $\widehat{\theta}_{k,1}$ fails to reject the null hypothesis at the 95% significance level. Graph (2) shows averages from 46,530 inspections conducted by 33 inspectors for $\widehat{\theta}_{k,1}$ rejects the null at the 95% significance level. Maroon lines marks predicted values from regressing good-practices violations on base demerits and $Pre_{i,j}$, using pre-change off-bubble, and post-change inspections. Red dots mark averages from the pre-change period. Navy diamonds mark averages from the post-change period. Red dots from pre-change bubble inspections are hollowed out to indicate these observations were excluded from the regression producing the maroon prediction line.

Table 1: SNHD Violation Scoring

	7/1/2011-12/31/2012	1/1/2013–Present
Imminent Health Hazards	No demerits, temporary closure	No demerits, temporary closure
Critical Violations	5 demerits	5 demerits
Major Violations	3 demerits	3 demerits
Good-practices Violations	1 demerits	0 demerits

Table 2: Summary Statistics

Variable	N	Mean	Std. Dev.	Min.	Max.
Demerits | $Pre = 1$	16,651	8.2573	(7.6543)	0	70
Demerits | $Pre = 0$	92,064	6.3484	(6.7637)	0	64
Base Demerits | $Pre = 1$	16,651	6.5399	(6.9025)	0	65
Base Demerits | $Pre = 0$	92,064	6.3484	(6.7637)	0	64
Good-practices Violations | $Pre = 1$	16,651	1.7174	(1.4654)	0	11
Good-practices Violations | $Pre = 0$	92,064	1.1275	(1.2675)	0	9
Grade points | $Pre = 1$	16,651	2.7536	(0.5769)	0	3
Grade points | $Pre = 0$	92,064	2.8158	(0.4903)	0	3

Summary statistics from routine inspections of establishments classified as: bars/taverns, buffets, food trucks/mobile vendors, restaurants, and snack bars. *Pre* is a binary variable indicating that an inspection occurred prior to 2013 (in the pre-change period). To summarize letter grades, I assigned 3 grade points for an A, 2 grade points for a B, 1 grade point for a C, and 0 grade points for an X. In the pre-change period, 16,651 routine inspections were conducted on 9,872 different establishments. In the post-change period 92,064 routine inspections were conducted on 15,826 different establishments.

Table 3: Testing for Common Slope: Baseline Specification

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
Pre × *Base*	0.0207*	0.0207*	0.0217*	0.0216*
	(0.0123)	(0.0123)	(0.0111)	(0.0110)
Base	0.0966***	0.0957***	0.0759***	0.0752***
	(0.0062)	(0.0062)	(0.0079)	(0.0079)
Pre	0.5667***	0.5644***	0.4930***	0.4926***
	(0.0684)	(0.0681)	(0.0606)	(0.0615)
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.0955	0.0976	0.2249	0.2266
N	89,332	89,332	47,223	47,223

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from pre-change routine inspections with 6 base demerits or fewer, and post-change inspections with 10 base demerits or fewer. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table 4: Testing for Under-reporting: Baseline Specification

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
Pre × *Bubble*	-0.7980***	-0.7969***	-0.7395***	-0.7392***
	(0.0531)	(0.0534)	(0.0633)	(0.0634)
Base	0.0976***	0.0967***	0.0788***	0.0780***
	(0.0060)	(0.0060)	(0.0074)	(0.0073)
Pre	0.6285***	0.6262***	0.5630***	0.5635***
	(0.0525)	(0.0522)	(0.0558)	(0.0562)
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.0935	0.0957	0.2140	0.2156
N	92,142	92,142	49,929	49,929

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from routine inspections with 10 base demerits or fewer. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table 5: Baseline Specification: Under-Reporting by Base-Demerit Level

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
$Pre \times \mathrm{I}(Base = 10)$	-1.9742***	-1.9591***	-1.8890***	-1.8786***
	(0.1098)	(0.1094)	(0.1517)	(0.1525)
$Pre \times \mathrm{I}(Base = 9)$	-0.9478***	-0.9456***	-0.8910***	-0.8906***
	(0.0731)	(0.0731)	(0.0859)	(0.0858)
$Pre \times \mathrm{I}(Base = 8)$	-0.6544***	-0.6545***	-0.5986***	-0.5988***
	(0.0502)	(0.0505)	(0.0610)	(0.0611)
$Base$	0.0980***	0.0970***	0.0794***	0.0787***
	(0.0060)	(0.0060)	(0.0074)	(0.0073)
Pre	0.6290***	0.6266***	0.5637***	0.5642***
	(0.0524)	(0.0522)	(0.0557)	(0.0561)
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.0948	0.0969	0.2161	0.2176
N	92,142	92,142	49,929	49,929

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from routine inspections with 10 base demerits or fewer. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table 6: Testing for Under-reporting: First Alternative Specification

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
Pre × *Bubble*	-0.7887***	-0.7880***	-0.7165***	-0.7169***
	(0.0552)	(0.0554)	(0.0681)	(0.0684)
Bubble	-0.0202	-0.0194	-0.0453	-0.0437
	(0.0342)	(0.0340)	(0.0309)	(0.0310)
Base	0.0996***	0.0986***	0.0828***	0.0820***
	(0.0070)	(0.0070)	(0.0077)	(0.0077)
Pre	0.6255***	0.6232***	0.5558***	0.5565***
	(0.0525)	(0.0522)	(0.0551)	(0.0555)
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.0935	0.0957	0.2141	0.2157
N	92,142	92,142	49,929	49,929

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from routine inspections with 10 base demerits or fewer. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table 7: First Alternative Specification: Under-Reporting by Base-Demerit Level

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
$Pre \times \mathrm{I}(Base = 10)$	-1.6853***	-1.6719***	-1.5769***	-1.5666***
	(0.1184)	(0.1178)	(0.1466)	(0.1478)
$Pre \times \mathrm{I}(Base = 9)$	-1.1139***	-1.1120***	-1.0034***	-1.0038***
	(0.0814)	(0.0814)	(0.0895)	(0.0895)
$Pre \times \mathrm{I}(Base = 8)$	-0.5486***	-0.5491***	-0.4869***	-0.4878***
	(0.0491)	(0.0494)	(0.0616)	(0.0620)
$\mathrm{I}(Base = 10)$	-0.2916***	-0.2894***	-0.3405***	-0.3394***
	(0.0670)	(0.0667)	(0.0665)	(0.0673)
$\mathrm{I}(Base = 9)$	0.1638***	0.1646***	0.0915**	0.0932**
	(0.0400)	(0.0399)	(0.0384)	(0.0381)
$\mathrm{I}(Base = 8)$	-0.1077***	-0.1069***	-0.1308***	-0.1293***
	(0.0357)	(0.0355)	(0.0324)	(0.0325)
$Base$	0.0983***	0.0973***	0.0830***	0.0822***
	(0.0069)	(0.0069)	(0.0077)	(0.0076)
Pre	0.6260***	0.6237***	0.5559***	0.5566***
	(0.0525)	(0.0522)	(0.0551)	(0.0554)
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.0985	0.1006	0.2183	0.2198
N	92,142	92,142	49,929	49,929

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from routine inspections with 10 base demerits or fewer. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table 8: Testing for Under-reporting: A and B Range

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
$Pre \times BubbleB$	-1.2092***	-1.2219***	-1.2020***	-1.2253***
	(0.0890)	(0.0928)	(0.0886)	(0.1000)
$BubbleB$		0.0127		0.0235
		(0.0507)		(0.0614)
$Pre \times BubbleA$	-0.9531***	-1.0969***	-0.9755***	-1.0400***
	(0.0435)	(0.0491)	(0.0442)	(0.0608)
$BubbleA$		0.1438***		0.0650*
		(0.0267)		(0.0340)
$Pre \times Base$	0.0456***	0.0487***	0.0483***	0.0502***
	(0.0069)	(0.0066)	(0.0063)	(0.0063)
$Base$	0.0753***	0.0722***	0.0674***	0.0656***
	(0.0050)	(0.0046)	(0.0057)	(0.0059)
Pre	0.5023***	0.5193***	0.4315***	0.4371***
	(0.0558)	(0.0552)	(0.0610)	(0.0606)
Inspector FE	N	N	Y	Y
Day-of-Week FE	Y	Y	Y	Y
Month-of-Year FE	Y	Y	Y	Y
R-squared	0.1629	0.1646	0.2813	0.2815
N	105,634	105,634	54,718	54,718

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from inspections with 30 or fewer base demerits. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

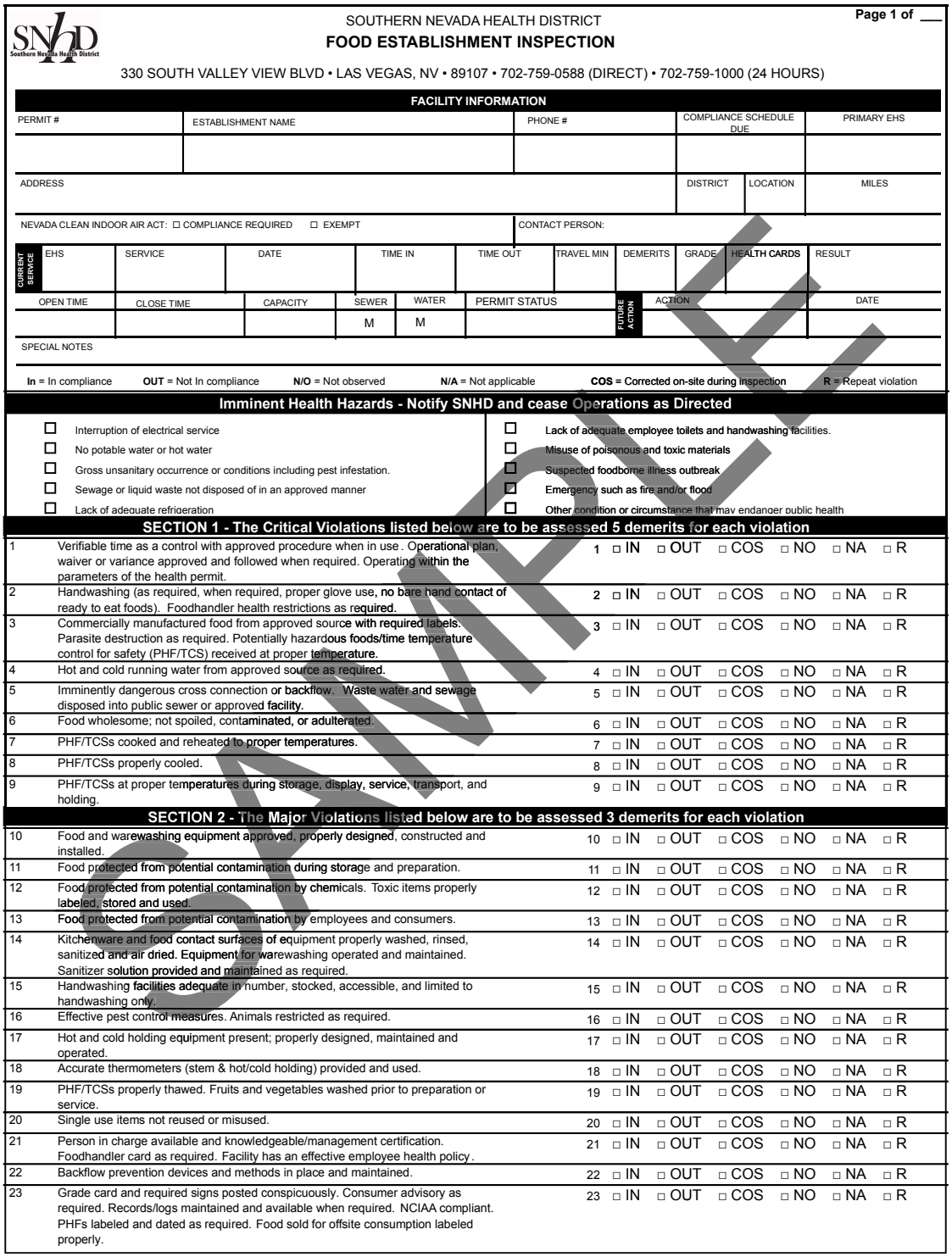
Table 9: Probability of *A*-grade Threshold Outcomes

Variable	(1) I(*GP* = *margin*)	(2) I(*GP* = *margin*)	(3) I(*GP* = *margin*)	(4) I(*GP* = *margin*)
I(*Base* = 10) × *Pre*	0.5809***	0.5821***	0.5453***	0.5457***
	(0.0451)	(0.0455)	(0.0557)	(0.0560)
I(*Base* = 9) × *Pre*	0.4641***	0.4636***	0.4769***	0.4765***
	(0.0236)	(0.0235)	(0.0282)	(0.0280)
I(*Base* = 8) × *Pre*	0.2728***	0.2725***	0.2822***	0.2822***
	(0.0197)	(0.0197)	(0.0197)	(0.0196)
I(*Base* = 6) × *Pre*	0.1032***	0.1028***	0.0941***	0.0940***
	(0.0104)	(0.0104)	(0.0095)	(0.0093)
I(*Base* = 5) × *Pre*	0.0213***	0.0208***	0.0181**	0.0178**
	(0.0074)	(0.0072)	(0.0081)	(0.0079)
I(*Base* = 3) × *Pre*	0.0004	-0.0001	-0.0047	-0.0046
	(0.0003)	(0.0007)	(0.0039)	(0.0040)
Base FE	Y	Y	Y	Y
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.2271	0.2275	0.2756	0.2763
N	65,431	65,431	37,433	37,433

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from inspections with at least 3, and no more than 10, base demerits. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

A1 Appendix



SOUTHERN NEVADA HEALTH DISTRICT

FOOD ESTABLISHMENT INSPECTION

Page 1 of ___

330 SOUTH VALLEY VIEW BLVD • LAS VEGAS, NV • 89107 • 702-759-0588 (DIRECT) • 702-759-1000 (24 HOURS)

FACILITY INFORMATION

PERMIT #	ESTABLISHMENT NAME	PHONE #	COMPLIANCE SCHEDULE DUE	PRIMARY EHS

ADDRESS	DISTRICT	LOCATION	MILES

NEVADA CLEAN INDOOR AIR ACT: ☐ COMPLIANCE REQUIRED ☐ EXEMPT — CONTACT PERSON:

CURRENT SERVICE	EHS	SERVICE	DATE	TIME IN	TIME OUT	TRAVEL MIN	DEMERITS	GRADE	HEALTH CARDS	RESULT

OPEN TIME	CLOSE TIME	CAPACITY	SEWER	WATER	PERMIT STATUS	FUTURE ACTION	ACTION	DATE
			M	M				

SPECIAL NOTES

In = In compliance **OUT** = Not In compliance **N/O** = Not observed **N/A** = Not applicable **COS** = Corrected on-site during inspection **R** = Repeat violation

Imminent Health Hazards - Notify SNHD and cease Operations as Directed

- ☐ Interruption of electrical service
- ☐ No potable water or hot water
- ☐ Gross unsanitary occurrence or conditions including pest infestation.
- ☐ Sewage or liquid waste not disposed of in an approved manner
- ☐ Lack of adequate refrigeration
- ☐ Lack of adequate employee toilets and handwashing facilities.
- ☐ Misuse of poisonous and toxic materials
- ☐ Suspected foodborne illness outbreak
- ☑ Emergency such as fire and/or flood
- ☐ Other condition or circumstance that may endanger public health

SECTION 1 - The Critical Violations listed below are to be assessed 5 demerits for each violation

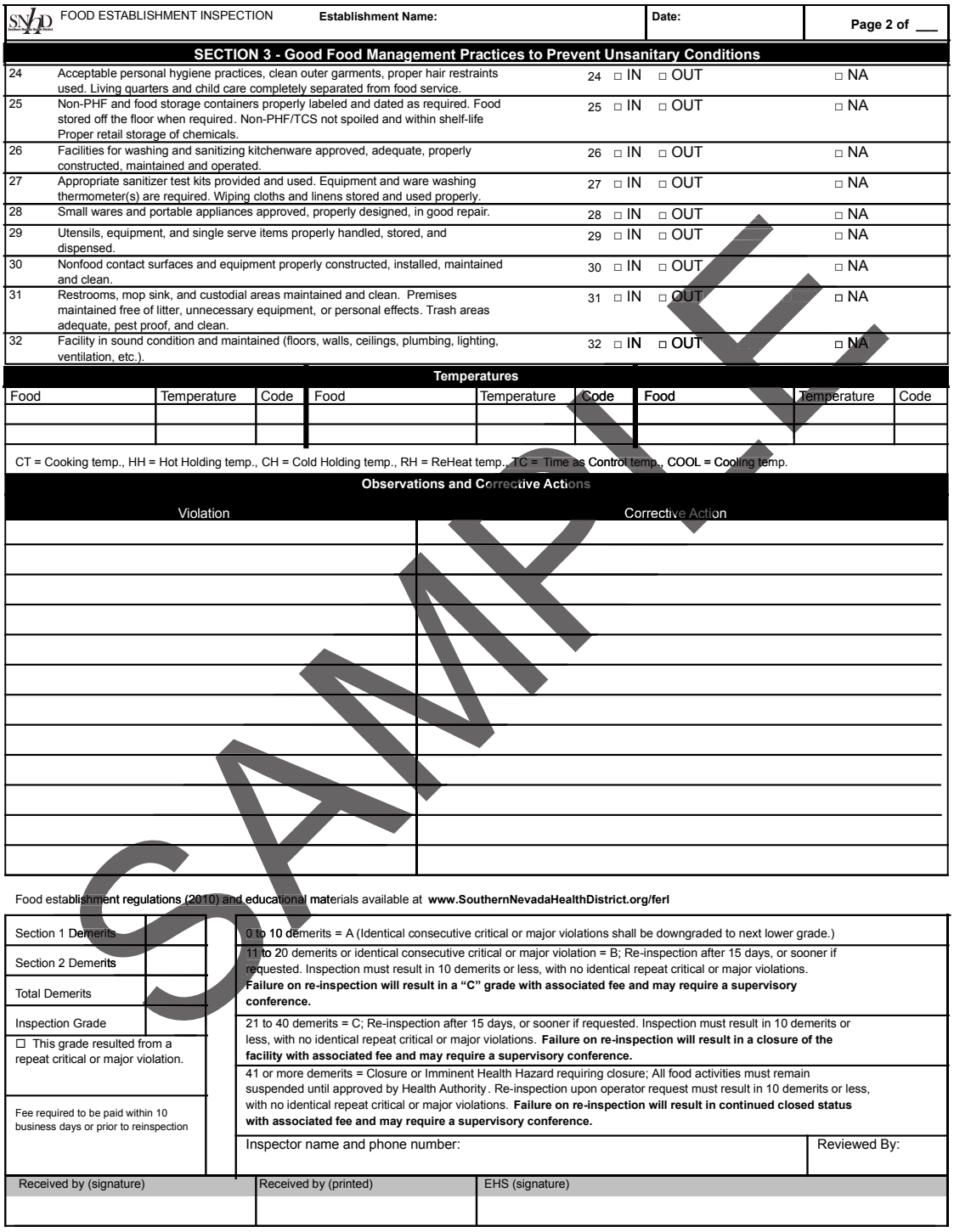
1	Verifiable time as a control with approved procedure when in use. Operational plan, waiver or variance approved and followed when required. Operating within the parameters of the health permit.	1	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
2	Handwashing (as required, when required, proper glove use, no bare hand contact of ready to eat foods). Foodhandler health restrictions as required.	2	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
3	Commercially manufactured food from approved source with required labels. Parasite destruction as required. Potentially hazardous foods/time temperature control for safety (PHF/TCS) received at proper temperature.	3	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
4	Hot and cold running water from approved source as required.	4	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
5	Imminently dangerous cross connection or backflow. Waste water and sewage disposed into public sewer or approved facility.	5	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
6	Food wholesome; not spoiled, contaminated, or adulterated.	6	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
7	PHF/TCSs cooked and reheated to proper temperatures.	7	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
8	PHF/TCSs properly cooled.	8	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
9	PHF/TCSs at proper temperatures during storage, display, service, transport, and holding.	9	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R

SECTION 2 - The Major Violations listed below are to be assessed 3 demerits for each violation

10	Food and warewashing equipment approved, properly designed, constructed and installed.	10	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
11	Food protected from potential contamination during storage and preparation.	11	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
12	Food protected from potential contamination by chemicals. Toxic items properly labeled, stored and used.	12	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
13	Food protected from potential contamination by employees and consumers.	13	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
14	Kitchenware and food contact surfaces of equipment properly washed, rinsed, sanitized and air dried. Equipment for warewashing operated and maintained. Sanitizer solution provided and maintained as required.	14	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
15	Handwashing facilities adequate in number, stocked, accessible, and limited to handwashing only.	15	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
16	Effective pest control measures. Animals restricted as required.	16	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
17	Hot and cold holding equipment present; properly designed, maintained and operated.	17	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
18	Accurate thermometers (stem & hot/cold holding) provided and used.	18	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
19	PHF/TCSs properly thawed. Fruits and vegetables washed prior to preparation or service.	19	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
20	Single use items not reused or misused.	20	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
21	Person in charge available and knowledgeable/management certification. Foodhandler card as required. Facility has an effective employee health policy.	21	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
22	Backflow prevention devices and methods in place and maintained.	22	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R
23	Grade card and required signs posted conspicuously. Consumer advisory as required. Records/logs maintained and available when required. NCIAA compliant. PHFs labeled and dated as required. Food sold for offsite consumption labeled properly.	23	☐ IN	☐ OUT	☐ COS	☐ NO	☐ NA	☐ R

Figure A1: Current SNHD Food Establishment Inspection Report: Page 1

Form was retrieved 9/15/2020 from https://www.southernnevadahealthdistrict.org/download/eh/fe-inspection-report.pdf.



FOOD ESTABLISHMENT INSPECTION **Establishment Name:** **Date:** **Page 2 of ___**

SECTION 3 - Good Food Management Practices to Prevent Unsanitary Conditions

24	Acceptable personal hygiene practices, clean outer garments, proper hair restraints used. Living quarters and child care completely separated from food service.	24 □ IN	□ OUT	□ NA
25	Non-PHF and food storage containers properly labeled and dated as required. Food stored off the floor when required. Non-PHF/TCS not spoiled and within shelf-life Proper retail storage of chemicals.	25 □ IN	□ OUT	□ NA
26	Facilities for washing and sanitizing kitchenware approved, adequate, properly constructed, maintained and operated.	26 □ IN	□ OUT	□ NA
27	Appropriate sanitizer test kits provided and used. Equipment and ware washing thermometer(s) are required. Wiping cloths and linens stored and used properly.	27 □ IN	□ OUT	□ NA
28	Small wares and portable appliances approved, properly designed, in good repair.	28 □ IN	□ OUT	□ NA
29	Utensils, equipment, and single serve items properly handled, stored, and dispensed.	29 □ IN	□ OUT	□ NA
30	Nonfood contact surfaces and equipment properly constructed, installed, maintained and clean.	30 □ IN	□ OUT	□ NA
31	Restrooms, mop sink, and custodial areas maintained and clean. Premises maintained free of litter, unnecessary equipment, or personal effects. Trash areas adequate, pest proof, and clean.	31 □ IN	□ OUT	□ NA
32	Facility in sound condition and maintained (floors, walls, ceilings, plumbing, lighting, ventilation, etc.).	32 □ IN	□ OUT	□ NA

Temperatures

Food	Temperature	Code	Food	Temperature	Code	Food	Temperature	Code

CT = Cooking temp., HH = Hot Holding temp., CH = Cold Holding temp., RH = ReHeat temp., TC = Time as Control temp., COOL = Cooling temp.

Observations and Corrective Actions

Violation	Corrective Action

Food establishment regulations (2010) and educational materials available at **www.SouthernNevadaHealthDistrict.org/ferl**

Section 1 Demerits	0 to 10 demerits = A (Identical consecutive critical or major violations shall be downgraded to next lower grade.)
Section 2 Demerits	11 to 20 demerits or identical consecutive critical or major violation = B; Re-inspection after 15 days, or sooner if requested. Inspection must result in 10 demerits or less, with no identical repeat critical or major violations. **Failure on re-inspection will result in a "C" grade with associated fee and may require a supervisory conference.**
Total Demerits	
Inspection Grade	21 to 40 demerits = C; Re-inspection after 15 days, or sooner if requested. Inspection must result in 10 demerits or less, with no identical repeat critical or major violations. **Failure on re-inspection will result in a closure of the facility with associated fee and may require a supervisory conference.**
☐ This grade resulted from a repeat critical or major violation.	41 or more demerits = Closure or Imminent Health Hazard requiring closure; All food activities must remain suspended until approved by Health Authority. Re-inspection upon operator request must result in 10 demerits or less, with no identical repeat critical or major violations. **Failure on re-inspection will result in continued closed status with associated fee and may require a supervisory conference.**
Fee required to be paid within 10 business days or prior to reinspection	Inspector name and phone number: / Reviewed By:

Received by (signature)	Received by (printed)	EHS (signature)

Figure A2: Current SNHD Food Establishment Inspection Report: Page 2

Form was retrieved 9/15/2020 from https://www.southernnevadahealthdistrict.org/download/eh/fe-inspection-report.pdf.

PAGE 1 OF _____

SOUTHERN NEVADA HEALTH DISTRICT

FOOD ESTABLISHMENT INSPECTION REPORT

P.O. Box 3902 • LAS VEGAS, NV • 702-759-0588 (DIRECT) • 702-759-1000 (24 HOURS)

FACILITY INFORMATION

PERMIT #	ESTABLISHMENT NAME	PHONE #	COMPLIANCE SCHEDULE DUE	PRIMARY EHS

ADDRESS	PREVIOUS ACTION	EHS	ACTION	DATE	DEMERITS

NEVADA CLEAN INDOOR AIR ACT: ☐COMPLIANCE REQUIRED ☐EXEMPT — CONTACT PERSON:

CURRENT ACTION	EHS	ACTION	DATE	HEALTH CARDS	DEMERITS	GRADE	PERMIT STATUS	SEWER	WATER	TIME IN	TIME OUT	MILES	FUTURE ACTION	ACTION	DATE

SPECIAL NOTES

Imminent Health Hazards

	Interruption of electrical service		Lack of adequate employee toilets and handwashing facilities
	No potable water or hot water		Misuse of poisonous or toxic materials
	Gross unsanitary occurrences or conditions including pest infestation		Suspected foodborne illness outbreak
	Sewage or liquid waste not disposed of in an approved manner		Emergency such as fire and/or flood
	Lack of adequate refrigeration		Other condition or circumstance that may endanger public health

SECTION 1
The Critical Violations listed below are to be assessed 5 demerits for each violation

#	Condition	C	NC	COS	NO	NA	Previous Inspection History 'Date 1'	'Date 2'	'Date 3'
1	Verifiable time as a control with approved procedure when in use. Operational plan, HACCP plan, waiver or variance approved and followed when required. Operating within the parameters of the health permit.								
2	Handwashing (as required, when required, proper glove use, no bare hand contact of ready to eat foods). Foodhandler health restrictions as required.								
3	Commercially manufactured food from approved source with required labels. Parasite destruction as required. Potentially hazardous foods/time temperature control for safety (PHF/TCS) received at proper temperature.								
4	Hot and cold running water from approved source as required.								
5	Imminently dangerous cross connection or backflow. Waste water and sewage disposed into public sewer or approved facility.								
6	Food wholesome; not spoiled, contaminated, or adulterated.								
7	PHF/TCSs cooked and reheated to proper temperatures.								
8	PHF/TCSs properly cooled.								
9	PHF/TCSs at proper temperatures during storage, display, service, transport, and holding.								
10	Nevada Clean Indoor Air Act compliant.								
Total Demerits									

SECTION 2
The Major Violations listed below are to be assessed 3 demerits for each violation

#	Condition	C	NC	COS	NO	NA	Previous Inspection History 'Date 1'	'Date 2'	'Date 3'
11	Food protected from potential contamination during storage and preparation.								
12	Food protected from potential contamination by chemicals. Toxic items properly labeled, stored and used.								
13	Food protected from potential contamination by employees and consumers.								
14	Kitchenware and food contact surfaces of equipment properly washed, rinsed, sanitized and air dried. Sanitizer solution provided and maintained as required.								
15	Handwashing facilities adequate in number, stocked, accessible, and limited to handwashing only.								
16	Effective pest control measures. Animals restricted as required.								
17	Hot and cold holding equipment present; properly designed, maintained and operated.								
18	Accurate thermometers (stem & hot/cold holding) provided and used.								
19	PHF/TCSs properly thawed.								
20	Single use items not reused or misused.								
21	Person in charge available and knowledgeable/management certification.								
22	Backflow prevention devices and methods in place and maintained.								
23	"B" or "C" grade card and required signs posted conspicuously. Consumer advisory as required. Records/logs maintained and available when required.								
Total Demerits									

Updated Nov. 9, 2010

Figure A3: Pre-Period Inspection Report: Page 1

This form is available from the Wayback Machine at https://web.archive.org/web/20120617034842/http://www.southernnevadahealthdistrict.org:80/download/eh/fe-inspection-report.pdf. It was captured by that organization from the SNHD website on June 17, 2012.

SOUTHERN NEVADA HEALTH DISTRICT FOOD ESTABLISHMENT INSPECTION REPORT PAGE 2 OF_____, PERMIT #

SECTION 3 The Minor Violations listed below are to be assessed 1 demerit for each violation								Previous Inspection History		
#	Condition	C	NC	COS	NO	NA		'Date 1'	'Date 2'	'Date 3'
24	Acceptable personal hygiene practices, clean outer garments, proper hair restraints used. Living quarters and child care completely separated from food service.									
25	Food and food storage containers properly labeled and dated as required. Food stored off the floor when required. Non-PHF/TCS not spoiled and within shelf-life. Proper retail storage of chemicals.									
26	Facilities for washing and sanitizing kitchenware approved, adequate, properly constructed, maintained and operated.									
27	Appropriate sanitizer test kits provided and used. Ware washing thermometer(s) as required. Wiping cloths & linens stored and used properly.									
28	Food contact surfaces and equipment approved, food grade material, smooth, easily cleanable, properly constructed and installed.									
29	Utensils, equipment, and single serve items properly handled, stored, and dispensed.									
30	Nonfood contact surfaces and equipment properly constructed, installed, maintained and clean.									
31	Health cards as required. Facility has an effective employee health policy. "A" grade card posted conspicuously.									
32	Restrooms, mop sink, and custodial areas maintained and clean. Premises maintained free of litter, unnecessary equipment, or personal effects. Trash areas adequate, pest proof, and clean.									
33	Facility in sound condition and maintained (floors, walls, ceilings, plumbing, lighting, ventilation, etc.).									
34	Fruits and vegetables washed prior to preparation or service.									
	Total Demerits									

C = In Compliance NC = Not in Compliance COS = Corrected On Site
NO = Not Observed at time of inspection NA = Not Applicable

Temperatures								
Food	Temperature	Code	Food	Temperature	Code	Food	Temperature	Code

CT = Cooking Temperature, HH = Hot Holding temperature, CH = Cold Holding temperature, RH = ReHeat temperature, TC = Time as Control temperature

Observations and Corrective Actions			
#	Violation and Location	Corrective Action	Correct by date

Section 1 Demerits	
Section 2 Demerits	
Section 3 Demerits	
Total Demerits	
Inspection Grade	
☐ This grade resulted from a repeat critical or major violation.	

0 to 10 demerits = A (Identical consecutive critical or major violations shall be downgraded to next lower grade.)

11 to 20 demerits or identical consecutive critical or major violation = B; Re-inspection after 15 days, or sooner if requested. Inspection must result in 10 demerits or less, with no identical repeat critical or major violations. **Failure on re-inspection will result in a "C" grade with associated fee and may require a supervisory conference.**

21 to 40 demerits = C; Re-inspection after 15 days, or sooner if requested. Inspection must result in 10 demerits or less, with no identical repeat critical or major violations. **Failure on re-inspection will result in a closure of the facility with associated fee and may require a supervisory conference.**

41 or more demerits = Closure or Imminent Health Hazard requiring closure; All food activities must remain suspended until approved by Health Authority. Re-inspection upon operator request must result in 10 demerits or less, with no identical repeat critical or major violations. **Failure on re-inspection will result in continued closed status with associated fee and may require a supervisory conference.**

Fee required to be paid within 10 business days or prior to reinspection — $

Inspector name and phone number:

Received by (signature)	Received by (printed)	EHS (signature)	Reviewed by

Figure A4: Pre-Period Inspection Report: Page 2

This form is available from the Wayback Machine at https://web.archive.org/web/20120617034842/http://www.southernnevadahealthdistrict.org:80/download/eh/fe-inspection-report.pdf. It was captured by that organization from the SNHD website on June 17, 2012.

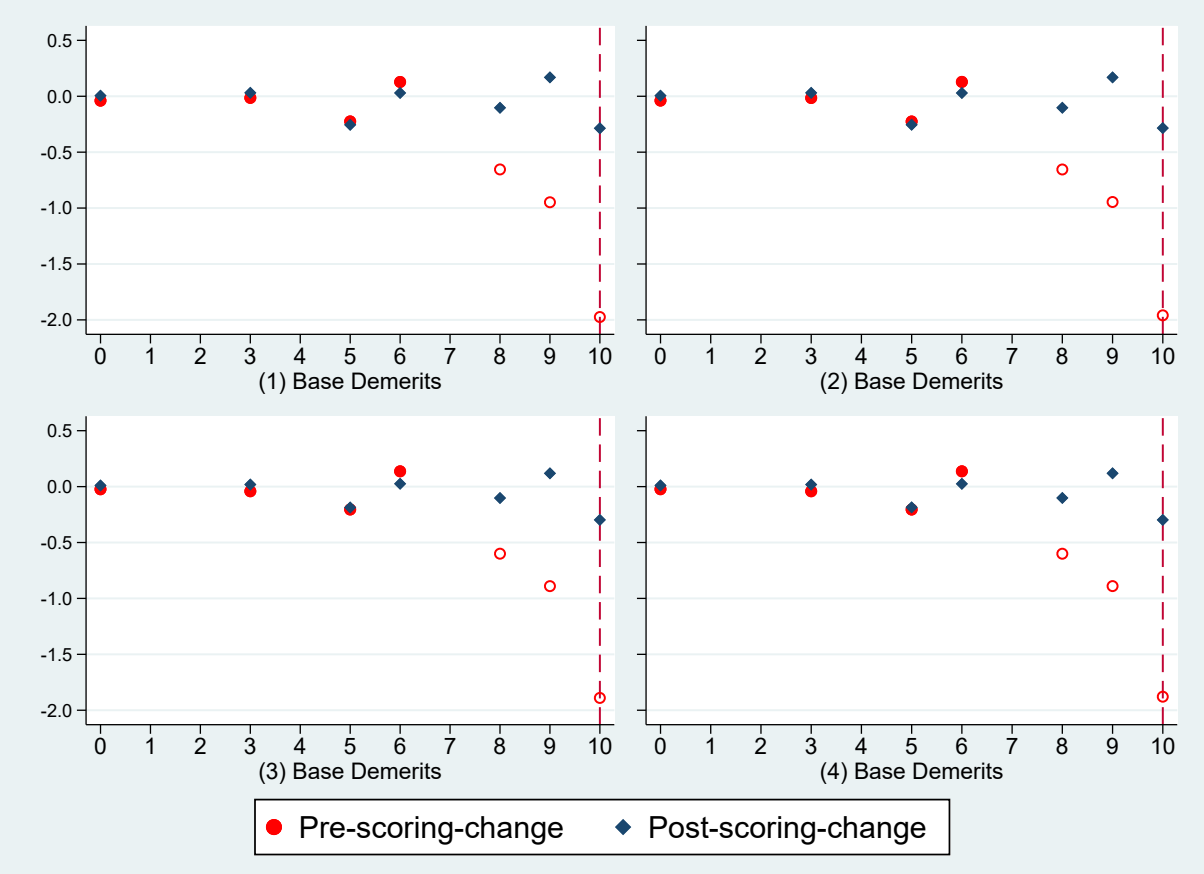


Figure A5: Residual Good-practices Violations by Base Demerits

Using the estimating samples in Table 4, with pre-change bubble inspections excluded, $GP_{i,j}$ is regressed on the controls reported in each column (excluding $Pre \times Bubble$). The full-sample residuals are then averaged by base demerits within scoring regime. Average residuals plotted in the upper left, upper right, lower left, and lower right panels; corresponds to Table 4 columns (1), (2), (3), (4), respectively. Red dots mark averages from the pre-change period. Navy diamonds mark averages from the post-change period. Red dots from pre-change bubble inspections are hollowed out to indicate these observations were excluded from the regression producing the residuals.

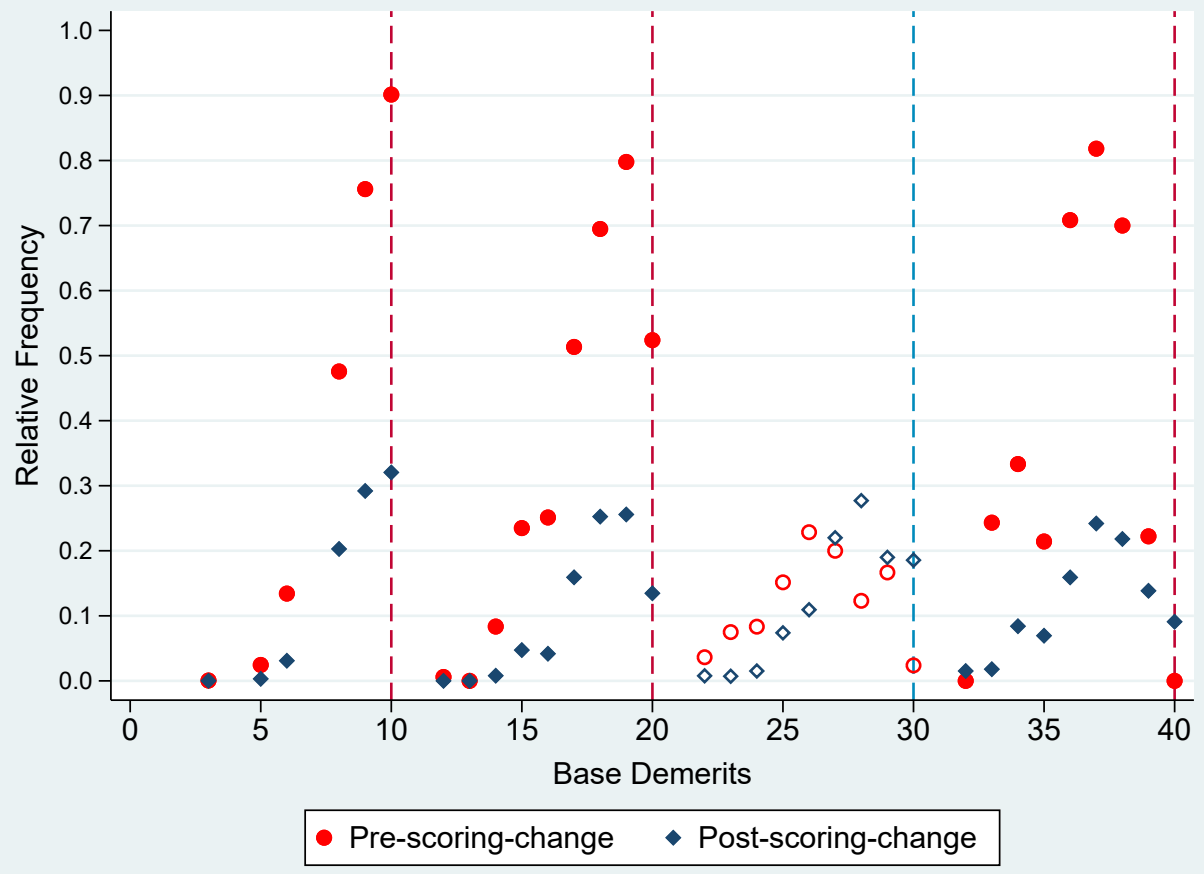


Figure A6: Relative Frequency of Reported Good-practices Violations Equal to the Margin

Red dots mark averages of $\mathrm{I}\left(GP_{i,j} = margin_{i,j}\right)$ from 12,088 pre-change routine inspections where: $Base_{i,j} \leq 40$, there were no IHH violations to trigger an X grade, and no downgrades occur. Navy diamonds mark averages of $\mathrm{I}\left(GP_{i,j} = margin_{i,j}\right)$ from 65,469 post-change inspections meeting the same criteria. When $Base_{i,j} \in [22, 30]$, hollow red circles and navy diamonds mark pre and post-period averages of $\mathrm{I}\left(GP_{i,j} = [30 - Base_{i,j}]\right)$. The dashed line at 30 is colored light blue because it is not an actual letter-grade threshold.

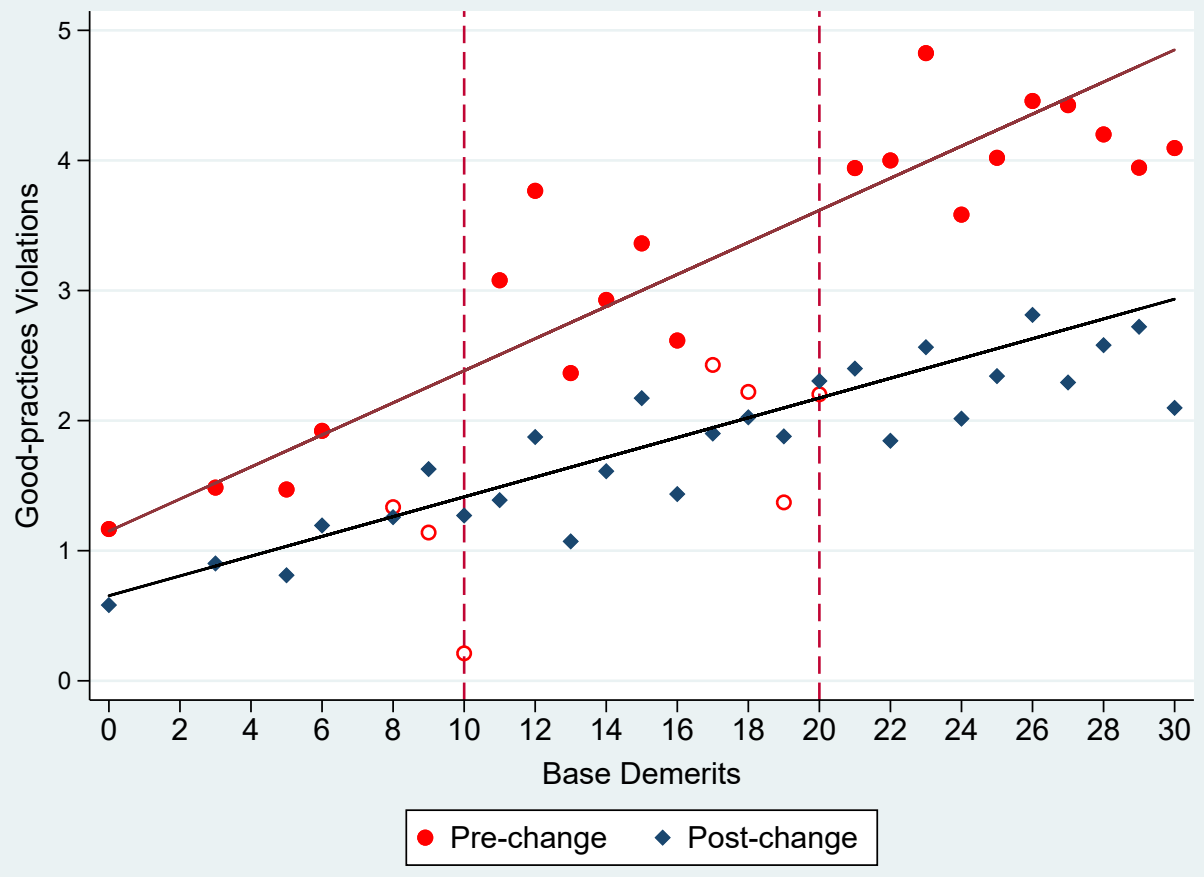


Figure A7: Reported Good-practices Violations by Base Demerits with OLS Predictions

The maroon line marks projected good-practices violations from regressing good-practices violations on base demerits from 13,166 pre-change inspections in which: $Base_{i,j} \leq 6$, $10 < Base_{i,j} \leq 16$, or $20 < Base_{i,j} \leq 30$. The navy line marks projected good-practices violations from regressing good-practices violations on base demerits from 90,775 post-change inspections in which, $Base_{i,j} \leq 30$. Red dots from pre-change bubble inspections are hollowed out to indicate these observations were excluded from the regression producing the maroon prediction line. Dashed lines at 10 and 20 demerits mark the threshold values for A and B grades, respectively.

Table A1: Good-Practices Violations from Pre-Change Period

* Acceptable personal hygiene practices, clean outer garments, proper hair restraints used. Living quarters and child care completely separated from food service.
* Food and food storage containers properly labeled and dated as required. Food stored off the floor when required. Non-PHF/TCS not spoiled and within shelf-life. Proper retail storage of chemicals.
* Facilities for washing and sanitizing kitchenware approved, adequate, properly constructed, maintained and operated.
* Appropriate sanitizer test kits provided and used. Ware washing thermometer(s) as required. Wiping cloths & linens stored and used properly.
Food contact surfaces and equipment approved, food grade material, smooth, easily cleanable, properly constructed and installed.
* Utensils, equipment, and single serve items properly handled, stored, and dispensed.
* Nonfood contact surfaces and equipment properly constructed, installed, maintained and clean.
Health cards as required. Facility has an effective employee health policy. "A" grade card posted conspicuously.
* Restrooms, mop sink, and custodial areas maintained and clean. Premises maintained free of litter, unnecessary equipment, or personal effects. Trash areas adequate, pest proof, and clean.
* Facility in sound condition and maintained (floors, walls, ceilings, plumbing, lighting, ventilation, etc.).
Fruits and vegetables washed prior to preparation or service.

* Indicates a good-practices violation in pre-change and post-change periods.

Table A2: Establishment Types in Raw Data

Establishment Type	Number of Establishments
Bakery Sales	101
Banquet Kitchen	98
Banquet Support	38
Bar/Tavern	4,613
Barbeque	142
Beer Bar	14
Buffet	468
Caterer	424
Childcare Kitchens	54
Concessions	82
Confection	55
Elementary School Kitchen	264
Farmer's Market	54
Food Trucks/Mobile Vendor	995
Frozen Meat Sales	35
Garde Manger	90
Grocery Store Sampling	91
Institutional Food Service	142
Kitchen Bakery	103
Main Kitchen	13
Meat/Poultry/Seafood	219
Pantry	460
Portable Bar	85
Portable Unit	1,239
Produce Market	106
Restaurant	8,624
Self-Service Food Truck	66
Snack Bar	2,474
Special Kitchen	2,214
Vegetable Prep	53

Since 2005, the number of establishments in each of the SNHD classifications.

Table A3: Testing for Common Slope: First Alternative Specification

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
Pre × *Base*	0.0226*	0.0226*	0.0216*	0.0218*
	(0.0122)	(0.0122)	(0.0121)	(0.0119)
Base	0.0947***	0.0937***	0.0766***	0.0757***
	(0.0072)	(0.0072)	(0.0094)	(0.0092)
Pre	0.5612***	0.5598***	0.5036***	0.5030***
	(0.0686)	(0.0680)	(0.0603)	(0.0613)
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.0968	0.0994	0.2251	0.2271
N	68,090	68,090	38,112	38,112

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

OLS estimates from pre-change and post-change routine inspections with 6 base demerits or fewer. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table A4: Testing for Under-reporting: Second Alternative Specification

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
Pre × *Bubble*	-0.8759***	-0.8749***	-0.8173***	-0.8170***
	(0.0616)	(0.0618)	(0.0547)	(0.0552)
Pre × *Base*	0.0152	0.0152	0.0164	0.0164
	(0.0124)	(0.0123)	(0.0110)	(0.0110)
Base	0.0966***	0.0957***	0.0767***	0.0759***
	(0.0062)	(0.0062)	(0.0080)	(0.0080)
Pre	0.5830***	0.5805***	0.5118***	0.5122***
	(0.0684)	(0.0680)	(0.0596)	(0.0602)
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.0936	0.0958	0.2142	0.2158
N	92,142	92,142	49,929	49,929

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

OLS estimates from routine inspections with 10 base demerits or fewer. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table A5: Testing for Common Slope: Excluding 2013

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$	(5) $GP_{i,j}$	(6) $GP_{i,j}$
Pre × *Base*	0.0210	0.0209	0.0209*	0.0237*	0.0237*	0.0207
	(0.0130)	(0.0129)	(0.0121)	(0.0130)	(0.0129)	(0.0134)
Base	0.0964***	0.0955***	0.0749***	0.0936***	0.0926***	0.0760***
	(0.0065)	(0.0065)	(0.0092)	(0.0071)	(0.0071)	(0.0110)
Pre	0.5974***	0.5953***	0.5703***	0.5910***	0.5901***	0.5849***
	(0.0734)	(0.0731)	(0.0674)	(0.0737)	(0.0731)	(0.0669)
Inspector FE	N	N	Y	N	N	Y
Day-of-Week FE	N	Y	Y	N	Y	Y
Month-of-Year FE	N	Y	Y	N	Y	Y
R-squared	0.1017	0.1038	0.2451	0.1080	0.1104	0.2479
N	78,993	78,993	36,491	60,556	60,556	29,797

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from routine inspections, excluding inspections from 2013. Columns (1) through (3) use post-change inspections with 10 or fewer base demerits, and pre-change inspections with 6 or fewer base demerits. Estimates in columns (4) through (6) use pre-change and post-change inspections with 6 or fewer base demerits. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table A6: Testing for Under-reporting: Excluding Year 2013

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$	(5) $GP_{i,j}$	(6) $GP_{i,j}$
Pre × *Bubble*	-0.7974***	-0.7966***	-0.7621***	-0.7922***	-0.7918***	-0.7402***
	(0.0547)	(0.0550)	(0.0699)	(0.0584)	(0.0586)	(0.0759)
Bubble				-0.0114	-0.0103	-0.0413
				(0.0344)	(0.0342)	(0.0348)
Base	0.0975***	0.0966***	0.0779***	0.0986***	0.0976***	0.0815***
	(0.0062)	(0.0062)	(0.0084)	(0.0068)	(0.0068)	(0.0088)
Pre	0.6601***	0.6579***	0.6431***	0.6583***	0.6563***	0.6360***
	(0.0565)	(0.0563)	(0.0593)	(0.0565)	(0.0564)	(0.0584)
Inspector FE	N	N	Y	N	N	Y
Day-of-Week FE	N	Y	Y	N	Y	Y
Month-of-Year FE	N	Y	Y	N	Y	Y
R-squared	0.0995	0.1016	0.2321	0.0995	0.1016	0.2322
N	81,803	81,803	38,987	81,803	81,803	38,987

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from routine inspections with 10 base demerits or fewer, excluding inspections from 2013. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table A7: Testing for Common Slope: All Establishment Types

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$	(5) $GP_{i,j}$	(6) $GP_{i,j}$
Pre × *Base*	0.0374***	0.0373***	0.0374***	0.0326***	0.0326***	0.0347***
	(0.0105)	(0.0105)	(0.0102)	(0.0103)	(0.0103)	(0.0104)
Base	0.1008***	0.1000***	0.0796***	0.1056***	0.1044***	0.0827***
	(0.0063)	(0.0063)	(0.0080)	(0.0075)	(0.0075)	(0.0088)
Pre	0.4755***	0.4804***	0.4187***	0.4801***	0.4869***	0.4275***
	(0.0595)	(0.0590)	(0.0516)	(0.0594)	(0.0586)	(0.0516)
Inspector FE	N	N	Y	N	N	Y
Day-of-Week FE	N	Y	Y	N	Y	Y
Month-of-Year FE	N	Y	Y	N	Y	Y
R-squared	0.1096	0.1122	0.2384	0.1069	0.1100	0.2410
N	126,877	126,877	70,477	100,808	100,808	58,621

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

OLS estimates from routine inspections of all establishment types. Columns (1) through (3) use post-change inspections with 10 or fewer base demerits, and pre-change inspections with 6 or fewer base demerits. Estimates in columns (4) through (6) use pre-change and post-change inspections with 6 or fewer base demerits. Standard errors, clustered two-way at the inspector and establishment levels, are reported in parentheses.

Table A8: Testing for Under-reporting: All Establishment Types

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$	(5) $GP_{i,j}$	(6) $GP_{i,j}$
Pre × *Bubble*	-0.7114***	-0.7105***	-0.6701***	-0.6698***	-0.6694***	-0.6213***
	(0.0528)	(0.0527)	(0.0633)	(0.0540)	(0.0540)	(0.0684)
Bubble				-0.0916***	-0.0902***	-0.0968***
				(0.0335)	(0.0334)	(0.0315)
Base	0.1030***	0.1021***	0.0844***	0.1115***	0.1105***	0.0927***
	(0.0062)	(0.0062)	(0.0074)	(0.0073)	(0.0072)	(0.0076)
Pre	0.5731***	0.5778***	0.5235***	0.5612***	0.5658***	0.5100***
	(0.0510)	(0.0504)	(0.0476)	(0.0510)	(0.0504)	(0.0475)
Inspector FE	N	N	Y	N	N	Y
Day-of-Week FE	N	Y	Y	N	Y	Y
Month-of-Year FE	N	Y	Y	N	Y	Y
R-squared	0.1079	0.1104	0.2295	0.1083	0.1108	0.2299
N	130,362	130,362	73,828	130,362	130,362	73,828

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

OLS estimates from routine inspections of all establishment types, with 10 base demerits or fewer. Standard errors are clustered two-way on inspector and establishment and reported in parentheses.

Table A9: Bunching at Thresholds

Variable	(1) I(*GP* = *margin*)	(2) I(*GP* = *margin*)	(3) I(*GP* = *margin*)
Pre	0.1509***	0.1552***	0.1557***
	(0.0085)	(0.0100)	(0.0100)
Base FE	Y	Y	Y
Inspector FE	N	N	Y
Day-of-Week FE	N	Y	N
Month-of-Year FE	N	Y	N
R-squared	0.1851	0.2199	0.2205
N	77,557	44,109	44,109

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from inspections where: $margin_{i,j} \leq 8$, $Base_{i,j} \leq 40$, and no IHH violations or downgrades occurred. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses. In all specifications, the dependent variable indicates that the reported number of good-practices violations was exactly equal to the margin between base demerits and the next letter-grade threshold.

Table A10: Probability of Passing Threshold

Variable	(1) I(*GP* > *margin*)	(2) I(*GP* > *margin*)	(3) I(*GP* > *margin*)
Pre	-0.0792***	-0.0900***	-0.0903***
	(0.0079)	(0.0120)	(0.0117)
Base FE	Y	Y	Y
Inspector FE	N	N	Y
Day-of-Week FE	N	Y	N
Month-of-Year FE	N	Y	N
R-squared	0.3731	0.3789	0.3793
N	77,557	44,109	44,109

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from inspections where: $margin_{i,j} \leq 8$, $Base_{i,j} \leq 40$, and no IHH violations or downgrades occurred. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses. In all specifications, the dependent variable indicates that the reported number of good-practices violations exceeds the margin between base demerits and the next letter-grade threshold.

Table A11: Probability of Sub-A Grades: Establishments with and without Pre-Period Bubble A's

	Pre-Change Period		Post-change Period	
Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
PreBubbleA	-0.0041	-0.0043	-0.0976***	-0.0979***
	(0.0061)	(0.0061)	(0.0091)	(0.0094)
Pre-Period Avg. Score FE	Y	Y	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.0484	0.0497	0.4734	0.4722
N	51,212	51,212	13,369	13,369

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from routine inspections of primary sample establishments with more than one pre-period inspection and at least one post-period inspection. Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.

Table A12: Testing for Common Slope: A and B Range

Variable	(1) $GP_{i,j}$	(2) $GP_{i,j}$	(3) $GP_{i,j}$	(4) $GP_{i,j}$
$Pre \times Base$	0.0473***	0.0474***	0.0504***	0.0504***
	(0.0069)	(0.0069)	(0.0064)	(0.0064)
Base	0.0760***	0.0753***	0.0676***	0.0670***
	(0.0050)	(0.0050)	(0.0057)	(0.0056)
Pre	0.4967***	0.4934***	0.4210***	0.4178***
	(0.0563)	(0.0558)	(0.0609)	(0.0616)
Inspector FE	N	N	Y	Y
Day-of-Week FE	N	Y	N	Y
Month-of-Year FE	N	Y	N	Y
R-squared	0.1624	0.1644	0.2911	0.2926
N	102,051	102,051	51,330	51,330

$^{***}p < 0.01$, $^{**}p < 0.05$,$^{*}p < 0.1$

OLS estimates from post-change inspections with 30 or fewer base demerits, and off-bubble pre-change inspections with 30 or fewer base demerits (i.e., excluding pre-change inspections where $8 \leq Base_{i,j} \leq 10$ or $17 \leq Base_{i,j} \leq 20$). Standard errors are clustered two-way on inspector and establishment, and reported in parentheses.