Is Public Grading Worth the Costs?

An Evaluation of New York City's Restaurant Grades Policy

Rachel Meltzer
Milano School, The New School

Michah W. Rothbart
Institute for Education and Social Policy
Wagner School PhD Student, New York University

Amy Ellen Schwartz
Institute for Education and Social Policy
Steinhardt and Wagner Schools, New York University
Maxwell School, Syracuse University

Thad Calabrese
Wagner School, New York University

Diana Silver
Steinhardt School, New York University

Tod Mijanovich
Wagner School, New York University

and

Meryle Weinstein
Institute for Education and Social Policy
Steinhardt School, New York University

Draft updated November 9, 2015

We gratefully acknowledge contributions from the NYC Department of Health and Mental Hygiene and the NYC Department of Finance, and financial support from the Robert Wood Johnson Foundation. Special thanks to Dan Kass, Wendy McKelvey, Bryan Jacobson, Melissa Wong, Corinne Schiff, Karen Schlain, Zachary Papper, and Steven Robinson.

Abstract

Grading schemes have become a popular way to accessibly and concisely convey the quality of public services. One recently popular grading mechanism applies to the sanitation of restaurants. Restaurant grading has an intuitive appeal: it makes information about restaurants' sanitary conditions more readily available and, therefore, may reduce the prevalence of foodborne illnesses. However, its reception has not been uniformly positive, especially on the part of restaurant operators, who are largely concerned about the policy's financial impacts. Most of what we know about both the food safety and economic impacts is anecdotal. Drawing from cross-agency municipal administrative records, we construct a rich dataset that tracks the food safety compliance and sales activity for the universe of graded restaurants in New York City over multiple years, both before and after the grading policy's implementation. We are then able to systematically test for the policy's impact on both food safety compliance and economic performance (on the part of the restaurants and municipality), providing a more holistic assessment of the grading policy than would have been achieved by relying on siloed administrative data. Results from our analysis indicate that New York City's grading policy has, after an initial period of adjustment, improved restaurants' sanitary conditions (as measured by inspection scores) and reduced fines levied. The impacts on sales revenues, and by association sales taxes, however are less clear and tend towards null in periods immediately following the policy's implementation. Altogether, the results suggest that while the policy induced restaurants to improve food safety compliance, it did not generate significant revenue for businesses through increased sales or for the City through fines and sales taxes levied.

I. Introduction

Grading schemes have become a popular way to accessibly and concisely convey the quality of public services. They are a good example of an information-based policy, which aims to influence, or “nudge” (Thaler and Sunstein 2009), the user’s behavior in a particular direction rather than mandating or directly incentivizing it. Municipalities across the U.S. grade the performance of public schools, street cleanliness is frequently scored and graded, and the Straphangers Campaign in New York City even produces a “report card” ranking the performance of each subway train line. One of the more recent grading mechanisms applies to the sanitation of restaurants, a policy that has taken hold in cities across the globe (Filion and Powell 2009). The intuition behind a grading policy is clear: the grades succinctly and conspicuously summarize information on sanitary conditions, reducing the chances of people consuming food in the places more likely to bear foodborne illnesses. However, its reception has not been uniformly positive, especially on the part of restaurant operators, who are largely concerned with the policy’s financial impacts. Does such a policy actually improve food safety? Does it help or harm the economic viability of the restaurants subject to the grading scheme? What are the costs and benefits for the municipality from such a grading scheme? Most of what we know about these issues is anecdotal; in this paper, we shed light on all of these questions by implementing a systematic analysis using large administrative datasets.

Theoretically, the economic and food safety impacts of the public grading policy are ambiguous. If given more information about the sanitary conditions of the restaurant (and a choice across dining options), we would expect consumers to internalize that information to make eating decisions that reduce the chances of foodborne illness. In the most optimistic scenario, the posting of grades will not only influence existing restaurant consumers to sort away from the establishments with lower grades towards those with the higher grades, but also induce consumers who did not patronize restaurants before to dine out (presumably more so at the establishments with higher grades). The restaurants, either in response or in expectation of this change in behavior on the part of the consumer, will adjust their sanitary practices to better comply and earn (and therefore post) higher grades. The municipality will also bear fiscal consequences, from both the inspection-based fines and the taxes levied off of the potentially increased sales revenues. We note, however, that patrons may end up sorting across restaurants (away from those with lower grades and towards those with higher grades) or minimally relying on the posted grades (in the case of establishments with long-standing and captive consumer bases), both of which could reduce the incentive for restaurants to improve their grades and dampen the overall improvement in food safety compliance and/or restaurant sales activity. On net, the aggregate effects from the grading regime are unknown and therefore require an empirical investigation—one that we conduct here.

While it is intuitive that such a grading policy would have economic implications, above and beyond those related to food safety and health, the evaluation of these effects presents challenges related to data coordination and assembly. Working closely with the New York City Department of Health and Mental Hygiene (the agency that administers the restaurant grading policy) and the Department of Finance (the agency that houses the business sales revenue data), we constructed a rich dataset that tracks the food safety compliance and sales activity for the universe of graded restaurants over multiple years both before and after the grading policy's implementation. We also overcome the challenges of confidentiality, as they pertain to individual business' sales revenues, by randomly "binning" restaurant-level data and conducting an analysis on grouped restaurant data that produces unbiased estimates of the policy's economic effects. In the end, these data compilation efforts produced a resource, and subsequently a policy analysis, that can be used to more holistically evaluate the effects of New York City's restaurant grading policy than would have been achieved relying on the siloed administrative data.

Results from our analysis indicate that New York City's grading policy has impacted restaurant sanitary conditions (as measured by inspection scores) and fines levied. Specifically, while there appears to be a period of adjustment, as observed by a slight increase in initial inspection scores, final inspection scores decline (i.e. sanitation improves) by about 4 points (about 17% of the pre-grading mean) upon policy implementation and then continue to decline at about ¼ point per quarter. Fines increase immediately after that start of the grading policy (by about $65 per inspection, or about 6% of percent of the mean fine before grading), but decline thereafter such that any gain is reversed by the second quarter post-implementation (and further reduced in each quarter after that). The impacts on sales revenues and -- by association -- sales taxes, however, are less clear. While pre-post analyses indicate positive revenue effects immediately after the policy's implementation and a continued positive trend during the time periods thereafter, it is unclear the extent to which this is a result of the policy or simply a reflection of broader upward trends in economic activity. For example, more inclusive retail trade and hotel tax forecasts during the same period also show similar increases in revenues. Furthermore, models that estimate effects during the policy's rollout year show no significant revenue change for restaurants that are first exposed to the grading regime, compared to those that are later exposed. Altogether, the results suggest that the policy induced restaurants to improve food safety compliance, but did not generate significant revenue for restaurants or for the City.

The rest of the paper is organized as follows. We continue most immediately with a brief history of the restaurant grading policy in NYC, and in section three, we review the relevant empirical literature. Section four presents the data and measures and section five, the empirical strategy. In section six, we discuss the results, and then conclude with a summary and policy implications in section seven.

II. Background

a. New York City's Restaurant Grading Policy

DOHMH has long inspected the City's restaurants to ensure proper food safety practices, fining restaurants for violations and closing restaurants with public health hazards. Inspections occur on a regular basis, but inspectors are randomly assigned, and the precise timing of inspections is randomly scheduled within a window of approximately two months. Starting July 2010, DOHMH began assigning each restaurant a letter grade (A, B, or C) based upon the inspection scores that restaurants were then required to post as a summary of food safety compliance in a conspicuous location near the restaurant's entrance. These letters are printed in large bold font and are required to be near eye-level and at the front entrance—passersby can easily discern them, even from across the street (a sample of a posted grade is shown in Appendix A). DOHMH also added the grades for each inspection to its website. While inspection scores were available for viewing online before the start of the grading policy, they were not observable at the point-of-purchase and understanding their meaning online was a more involved process of reviewing the particular fines and points applied. The stated goal of the public grading law was to improve restaurant sanitary practices and decrease the incidence of restaurant-attributable food borne illness in NYC.

Inspection scores are calculated as the sum of violation points assigned during inspections. The points for a particular violation depend on the health risk it poses to the public, and the level of public health risk falls into three categories:

(1) *public health hazards*, such as failing to keep food at the proper temperature, minimum of 7 points per violation,
(2) *critical violations*, such as serving salad without properly washing it, minimum of 5 points per violation,
(3) *general violations*, such as not properly sanitizing cooking utensils, minimum of 2 points per violation.[1]

Additional points are added to each violation to reflect the severity of the violation, and the most extreme public health hazard violation leads to a maximum of 11 points each. Points from violations are then aggregated to generate the final inspection score, with lower scores reflecting more hygienic conditions. DOHMH classifies restaurants with scores of 13 and below as *A* restaurants, those with scores of 14-27 as *B* restaurants, and those with 28 or higher as *C* restaurants.[2] In addition to publicly posting grades, restaurants receiving an *A* are visited only annually for food safety inspections. Those receiving a *B* are inspected twice per year, and those receiving a *C* are inspected every four months. If an initial inspection leads to an inspection score in the *B* or *C* grade range, the restaurant is inspected again within one month. Therefore, a final grade is not assigned until after a re-inspection. Final inspections are those inspections that end with a provisional grade, which are comprised of initial *A* inspections and all re-inspections.

In addition to re-inspections, inspection scores and fines (and, therefore, grades) can be lowered (improved) through an adjudication process. Adjudication provides restaurants with an

[1] New York City Department of Health and Mental Hygiene. (2012).
[2] Restaurants can also be temporarily closed if they pose a large public safety risk.

opportunity to contest the violations, inspection scores, and grades at a tribunal administered by an independent agency, the Office of Administrative Trials and Hearings (OATH). Cases regarding violations received during the inspection process are adjudicated by hearing examiners (acting as judges), typically lawyers hired on by OATH. In FY2014, over 90% of the hearings conducted at the Health Tribunal concerned restaurant violations. Perhaps most importantly for purposes of the current analysis, restaurants have the right to post a placard that reads "Grade Pending" in lieu of posting *B* or *C* grades until they have their case heard at the OATH tribunal. In other work, we find that restaurants earning *B* and *C* grades at inspection are much more likely to have inspection scores reduced (improved) through the adjudication process in the post-period than they were in pre-period and this has substantial impacts on post-adjudicated grades, which are substantially better on average than the grades assigned at inspection (Silver et al.). These findings suggest that challenging grades in court is an important tool for restaurants motivated to post *A* grades in their window.

Both before and after the introduction of restaurant grading, the type and count of inspection violations determine the level of fines assessed. Fines range from $200-$2,000 per violation and are assessed at a restaurant's adjudication hearing at the discretion of a hearing officer – unless the grade is accepted and a lower fine is paid by the restaurant operator. It is important to note that after January 18, 2011, restaurants receiving an *A* grade at inspection were not fined for any inspection violations; therefore restaurants earning an *A* at inspection do not incur any fines for much of the post-period.

b. Empirical Literature Review

Grading, as a means of conveying information about the quality of services or goods, has been used in other policy contexts, including public education and public health. For example, many school districts grade public schools on their effectiveness (specifically improvements in test scores and other metrics), and they make these grades available to the public. Most studies focus on the variation in the grades themselves and how this differentiated information affects education-related outcomes. For example, there is some evidence that schools with lower grades have short-term improvement in aggregate student achievement (Rockoff and Turner 2010;

Winters and Cowen 2012). Figlio and Lucas (2004) find that the information provided by school grades has an effect on housing prices above and beyond that provided by test scores or other variables used to construct those same grades. However, the informative power of these grades appears to decline over time (although this might be a product of how they were implemented in their case site, Florida). Some of the earliest theoretical research on the information value of grading schemes has also been conducted in the context of agricultural markets. The findings from this body of work indicate that uniform grading schemes can increase the efficiency of those markets by succinctly classifying a heterogeneous commodity so that information is accessed in a less costly way and can lead to Pareto-optimal outcomes for the buyer/consumer. Net benefits for the producers are less certain, as some will experience reduced returns due to the information improvement provided by the grading scheme (Freebairn 1973, 1967).

As for public restaurant grading, the empirical research is scarce. Ho (2012) analyzes publicly available restaurant grading data for NYC, exploring the extent to which inspection scores in one period predict future scores. He observes that prior scores predict less than 2% of future grades, interpreting this as inconsistency in the inspection process resulting from complexity and imprecision in NYC's rules. Wong et al. (2015) provide new evidence of improved compliance since the beginning of the public grading programs, showing marked increases in the probability of a restaurant scoring in the A-range during unannounced initial inspections and offers survey evidence of the program's high approval ratings among New Yorkers. This is consistent with other surveys that have demonstrated that consumers use public inspection results to inform their dining decisions (Filion and Powell 2009).

As for impact studies, there are three. Two studies (Jin and Leslie, 2003; Simon et al., 2005) focus on the effects of the Los Angeles health inspection letter grade system, which required the posting of letter grades beginning in 1998. Jin and Leslie (2003) use OLS and difference-in-differences regression analyses to estimate the effect of the Los Angeles letter grades program on inspection scores, restaurant revenues, and foodborne illness hospitalizations. They find that posted grades improved restaurant inspection scores, that restaurant revenues responded to hygiene quality signals, and that foodborne-disease hospitalizations decreased in Los Angeles County following the implementation of the public letter grade program (the study by Simon et.

al. (2005) also provides evidence of reduced hospitalizations due to foodborne illnesses in Los Angeles County, compared to California overall). Jin and Leslie also suggest that the improvements in health outcomes cannot be explained by consumption choices alone, but are also likely a result of restaurant hygiene improvements. The most recent study, by Schwartz et. al. (2015), like many of the school grading studies, focuses on the differential impact of specific grades on restaurant food inspection compliance and economic activity. The authors find that a higher grade increases a restaurant's sales (and the associated sales taxes levied) and decreases the amount of fines assessed and the probability of the restaurant's closure. These results are also consistent with the expectation that public restaurant grading is providing new information for consumers' dining decisions (and, in turn, the restaurants' maintenance of sanitary conditions). Their results also suggest that there could be sorting of consumers away from the restaurants with lower grades to those with higher ones (although the authors do not test this directly).

While an important outcome, the incidence of foodborne illnesses is extremely difficult to link to the prevalence and use of posted inspection grades. Indeed, the correlation between inspection scores and foodborne disease outbreaks is inconsistent (Filion and Powell 2009) and empirically hard to identify. It relies entirely on restaurant patrons correctly identifying the foodborne illness and attributing its source, which is difficult to do for a number of reasons (duration of latency, expectations about food safety, proclivities towards gastrointestinal symptoms) (Jones and Angulo 2006; Mead et. al. 1999; Fein et. al. 1995;); moreover, it relies on their reporting the illness, which we know is done inconsistently (Jones and Angulo 2006 Mead et. al. 1999). For all of these reasons, we focus instead on food safety compliance and economic impacts (for both the restaurants and the municipal fisc).

III. Theoretical motivation

While restaurants have long been inspected and monitored by the government, the results of those inspections have not always been easily accessed. The obscurity of that information creates information asymmetries, whereby the restaurant operator knows of the sanitation conditions inside the establishment and the consumer knows only what is easily visible on site at the restaurant (or learns after-the-fact from some food-safety-related symptom). In this scenario, consumers make decisions about their eating habits based on incomplete information and the restaurants have fewer incentives to change their behaviors around sanitation issues that are not immediately discernible to the consumer. Theory suggests, then, that consumers should be affected by excessive incidents of foodborne illnesses. Public grading policies aim to address these information asymmetries (and the subsequent health risks) by making the sanitation inspection information more readily available to the consumer, in a way that minimizes his/her search costs. Specifically, the letter grading presents a format that is easily discernible, with a clear ranking or "mapping" of grades onto food-safety ratings (Thaler and Sunstein 2009). In addition, the posting of the letter grade in plain sight, at the point of purchase, makes the tool particularly salient and minimizes the effort required to gather and process the information (Thaler and Sunstein 2009).

We consider here the impacts of implementing a grading regime on both consumer and restaurant behaviors. These actions, in turn, will have fiscal implications for the municipality more generally. Whether or not the grade posting changes the consumers' dining decisions, and in turn the economic prospects for the restaurant operators and the municipality, is theoretically ambiguous. In the most optimistic scenario, the posting of grades will not only influence existing restaurant consumers to sort away from the establishments with lower grades towards those with the higher grades, but also induce consumers who did not patronize restaurants before (or at least not as frequently) to dine out (presumably more so at the establishments with higher grades). The restaurants, either in response or in expectation of this change in behavior on the part of the consumer, will adjust their sanitary practices to better comply and earn (and therefore post) higher grades. This means we would expect to see an improvement in food safety compliance (i.e. lower inspection scores and lower fines) and higher sales for the typical restaurant.

We consider this first scenario an optimistic, or upper bound, condition, as the grading policy's impact could be dampened in two important ways. First, it could be the case that no or few new consumers enter the restaurant market and that the implementation of the grading policy simply triggers a re-sorting of existing restaurant patrons. While inspection scores (and related fines) could still decline over time, any sales revenues could, on average, exhibit no or little change.

Spending would be reallocated from restaurants with lower grades to those with higher grades, resulting in little or no aggregate shift in revenues (and subsequent taxes). Second, we consider the expected outcomes when we drop the assumption of choice across restaurants. In this case, longstanding, dedicated patrons may prioritize information gathered from their first-hand experiences with the restaurant over the posted grade and continue their patronage in the same manner as before. Captive patrons, such as those without any other dining options nearby, may also not process the posted grade in the same way. Under these conditions, changes in both food safety compliance scores and revenues could be attenuated: the restaurants may be less motivated to invest in improving the posted grade and therefore any grade-induced sorting would be less evident in sales activity.

Since fiscal outcomes for the municipality closely track those for the restaurants, these implications are also ambiguous. Directly, the government will benefit from any increase in fines from the health inspections. Indirectly, any net increase in restaurant revenue will result in higher sales taxes. These revenues are up against any direct costs of administering the grading program, such as increased inspection manpower and program oversight. Anecdotally, we know that a grading system incurs significant manpower costs, mostly from its more frequent on-site inspections.

IV. Data and Measures

a. DOHMH Data

We obtained data on restaurant characteristics, restaurant zip codes, inspection data and score information, adjudication dates, grades assigned, and fines assessed from the NYC Department of Health and Mental Hygiene (DOHMH). Restaurant characteristics include number of seats, number of employees, an indicator for chain restaurant (at least 15 locations nationwide), and a series of variables indicating cuisine offered, service type, and venue type. Table 1 shows descriptive statistics for the restaurants in our sample. The mean restaurant has 3.25 final inspections over the study period, employs 6.2 workers and has about 29.6 seats. Just under 11% of restaurants in the sample are chains. DOHMH defines over 80 different cuisines, and records the type of service offered in a restaurant (i.e. wait service, counter service etc..) and the type of venue (i.e. diner, arena-stadium concession stand, bar/pub/brewery etc..).[3] We use data on restaurant characteristics recorded at the last inspection only, so these characteristics do not vary with time.

We also rely on scores assessed at each initial and final inspection to capture food safety compliance. Initial and final inspection scores will differ when restaurants do not get an *A* grade on their initial inspection and therefore have to undergo a re-inspection. The re-inspections typically take place four to six weeks after the initial inspection and scores tend to go down (i.e. improve). Whereas initial inspection scores and final inspection scores are the same for restaurants that earn an *A* at their initial inspection, re-inspections provide final inspection scores for restaurants that fail to earn an *A* at initial inspection. Thus, final inspection scores can reflect more learning or adjustment on the part of the restaurant and are the ones that result in a grade being posted (though restaurants reserve the right to post *Grade Pending* until their adjudication hearing). Importantly, both initial and re-inspections take place without advanced notice and inspectors are randomly assigned to their visits. Finally, we use data on fines to assess the program's revenue generation for the public sector (and conversely, the financial burden on restaurants) – focusing on the post-adjudication fines (because fines are only assessed after adjudication). All fine levels are adjusted using urban CPI to real 2013 dollars.

Our analytic sample for the impact on fines and closures includes the universe of final DOHMH food safety inspections from December 1, 2007 through February 28, 2013, spanning the two and a half years before public grading and the two and half years following the implementation of public grading(here forth referred to as "pre-period" and "post-period", respectively). This sample includes 159,588 initial inspections and 167,045 final inspections of 41,362 restaurants in all, including 29,864 restaurants that operate in the post-period and receive grades.

b. DOF Data

We obtain reported quarterly sales and sales tax liabilities (hereafter, sales taxes) for all NYC restaurants from the Department of Finance.[4] We then match DOF to the DOHMH data by

[3] A full list of cuisine, service, and venue types is available upon request.

Employer Identification Numbers (EINs). DOF provides data by quarter, while DOHMH data is date specific. We aggregate inspection data by restaurant (including grades and inspection scores) from DOHMH to quarters and match them to the quarterly DOF sales and tax data. We create a series of variables to indicate inspection scores and grades each quarter, including mean inspection score, inspection score at the beginning and the end of the quarter, share of days in each grade category and grade at the beginning and end of the quarter (for both provisional and final posted grades).

From DOF, we also access Real Property Assessment Database (RPAD) data to obtain the building class (for the restaurant's site) to control for locational or use characteristics that are associated with the amount of revenue generation and the likelihood of receiving a higher/lower grade (for example, the commercial classification likely identifies restaurants that serve mostly working and/or transient populations and tend to be patronized during more restricted daytime and weekday hours).

Because DOF cannot provide restaurant-level sales and tax data to outside researchers (to ensure confidentiality), grouped data were provided to us for the sales analyses. Specifically, DOF provided data for groups of 10 randomly assigned restaurants – that is, each observation provides data for a set of 10 restaurants randomly assigned to the same group, or "bin." A small number of groups have 11 rather than 10 restaurants in order to make sure all restaurants are included. To address attrition and entry, DOF first stratified the sample based upon quarters of operation and then assigned groups within these.[5]

The result is our matched and aggregated data set, which is organized by group-quarter and includes sales and tax information and summary inspection results. The data provides variables

---

[4] NYC restaurants are required to collect sales tax on food and beverage sales at a rate of 8.875% of gross sales - 4.875% for New York State and 4.0% for New York City. The State collects the entire sales tax from restaurants and remits the City's portion of sales tax revenue in the following month. Restaurants with $300,000 or less of sales in the previous quarter may remit sales taxes to New York State quarterly, while restaurants with more than $300,000 of sales in the previous quarter remit monthly to the State.

[5] Thus, the 5,145 restaurants operating in all 20 quarters of our study period were randomly assigned to 509 groups of 10 and five groups of 11; the 149 restaurants operating in all but the last quarter were grouped in five groups of 10 and 9 groups of 11; the 244 operating in all but the first were grouped in 20 groups of 10 and 4 groups of 11. They continue this process, sequentially, until all restaurants are assigned to groups, homogeneous in their quarters of operation and no group (and no observation) ever provides information on fewer than ten establishments.

summarizing the sales and tax activity in each group including quarterly means and standard deviations of sales, log(sales), sales taxes, and log(sales taxes). The data also includes quarterly means and standard deviations of inspection scores, number of seats and workers, daily mean provisional grade, daily mean posted grade as well as the share of group in each grade category at the beginning and the end of the quarter, in each zip code, operating in each building class and with each cuisine, venue, and service type. Our analytic sample for sales and taxes analyses includes 24,464 observations in 2,288 groups (including 254,216 restaurants or bars during the study period).[6]

V. Empirical Strategy

a. *Inspection scores and fines*

Our empirical strategy relies on a pre-post estimation model, which will allow us to compare various outcomes across time periods before and after the implementation of the restaurant grading policy. We know the precise start date of the policy, which provides an unambiguous identification of the pre- and post-periods. The pre-post identification strategy relies on the fact that restaurants are continuously inspected (and scored) throughout the study period, but only after the start of the grading policy are the inspection results made conspicuous via the posted grade. Therefore, any estimate of the policy effect captures the impact of new information provided through the posted grade. We test for this *posting* effect in two ways, both of which estimate the impact of the grading policy on outcomes that pertain to the restaurant's food safety compliance behavior, i.e. inspection scores, and the fiscal implications for the City, i.e. fines levied.[7] First, we rely on a standard pre-post model, which is specified in the following way:

$$(1) \quad y_{it} = \beta_0 + \boldsymbol{Grading\_Post}_{it}' \boldsymbol{\beta_1} + \boldsymbol{X_i}' \boldsymbol{\beta_2} + \beta_3 Pre\_Post_{it} + \gamma_i + \delta_t + \varepsilon_{it}$$

[6] The policy is implemented in the middle of the 2nd sales tax quarter in 2011. Our analytic sample includes data observed from the 4th quarter of 2008 to the 3rd quarter of 2013.

[7] We would also like to estimate changes in the likelihood of restaurant closures, but we are limited in how precisely we can identify the timing of closure. We identify restaurant closure as whether or not it is still operating at the time of inspection; since inspections occur irregularly, we are unable to precisely pinpoint the timing of closure, introducing considerable bias into our estimates. To address this issue, we would have to exclude all restaurants closing in the year during which the policy is implemented (approximately 20% of our sample) or during the last year of the panel (approximately 20% of our sample).

Here, y is a restaurant-specific outcome (inspection scores or fines). ***Grading_Post*** is a vector with two variables, which collectively estimate the impact of the grading policy: *Post* and *Post_trend*. *Post* takes on the value of 0 prior to the start of the grading policy (for $t<0$) and 1 thereafter (for $t \geq 0$); this can be interpreted as the immediate effect from the policy's implementation. *Post_trend* allows *Post* to vary over time, by interacting it with a linear time trend; this coefficient can be interpreted as the change per quarter following the policy's implementation. ***X*** is a vector of restaurant characteristics including cuisine, service, and venue type; *Pre_Post* is a linear time trend that extends the length of the study period (i.e. all 21 quarters, both before and after the grading policy's implementation); γ and δ are zip code and seasonal fixed effects, respectively; and ε is an error term with the usual properties. We also estimate model (1) controlling for restaurant fixed effects, μ_i, instead of γ_i and X_i.

While the first graded inspections occur on July 27, 2010, restaurants were not uniformly exposed to the new inspection regime (i.e. did not have to post a grade). We exploit the variation in grade posting during the roll-out period in an alternative specification, where we limit the sample to restaurant-quarter observations during the first year of the policy's legislative start, or the roll-out period. Here we are comparing the "early posters" to the "late posters" and this model takes on the following form:

$$(2)\ y_{it} = \beta_0 + \beta_1 \textit{Post\_Rollout}_{it} + \mathbf{X'}_{\boldsymbol{i}}\boldsymbol{\beta}_{\mathbf{2}} + \gamma_i + \delta_t + \varepsilon_{it}$$

Again, y is a restaurant-specific outcome (inspection scores and fines) and *Post_Rollout* takes on the value of 1 if the restaurant has posted a grade placard (A, B, C, or Grade Pending) by the beginning of the quarter t and 0 otherwise. The remaining variables are identical to those defined above.

b. Sales and Sales Taxes

We follow the same progression in regression models as is presented above for the sales and sales tax analyses. The equations take on a slightly different form, however, due to the fact that we are now using grouped data. The baseline pre-post model is as follows:

$$(3)\ y_{gq} = \boldsymbol{Grading\_Post_{gq}}'\boldsymbol{\tau_1} + \boldsymbol{X'_{gq}\tau_2} + \tau_3 Pre\_Post_{gq} + \gamma_g + \delta_q + \varepsilon_{gq}$$

Here, y_{gq} is the group's average daily restaurant sales or sales taxes in quarter q. As above, ***Grading_Post*** is a vector with two variables, which collectively estimate the impact of the grading policy: *Post* and *Post_trend*. *Post* takes on the value of 1 if quarter q is after the start of the grading policy and 0 otherwise; *Post_trend* allows *Post* to vary over time, by interacting it with a linear time trend. These estimates should be unbiased if restaurant *i*'s average grade (or more generally, the restaurant's response to the grading policy) in group g in quarter q only affects restaurant *i*'s sales and not the sales of other restaurants in group g. We find this assumption plausible due to restaurant random assignment to groups and the inclusion of the quarter fixed effects, δ_q. We note that while point estimates are unbiased estimates of the mean impact on restaurants, the standard errors are larger than if we observed individual restaurant sales. Finally, $\boldsymbol{X}$ is a vector of mean restaurant characteristics and building class; and γ_g and δ_q are group and quarter fixed effects, respectively.

And as above, we specify an alternative model, exploiting the roll-out sample:

$$(4)\ y_{gq} = \tau_1 Post\_Rollout_{gq} + \boldsymbol{X'_{gq}\tau_2} + \gamma_g + \delta_q + \varepsilon_{gq}$$

Where $y_{gq}$is the group's mean daily sales or sales taxes in quarter q and $Post\_Rollout$ is the average share of days in quarter q of posting a grade placard for restaurants in group g. Here grade placards can read *A, B, C,* or *Grade Pending,* because treatment begins with the first grades assigned at inspection. Insofar as *Grade Pending* is an unclear signal to consumers, sales and sales taxes results will be attenuated to zero. Again, these estimates should be unbiased (compared to those derived from models run on restaurant-level observations), for the reasons stated above. The remaining variables are identical to those defined above.

VI. Results

a. Inspection scores

We first discuss results from the pre-post analysis, estimating the grading policy's impact on the restaurant's food safety compliance, as measured by inspection scores. Recall that the scoring rubric assigns points for various kinds and degrees of violations and a higher score indicates more violations. Tables 2 and 3 display these results for the pre-post analysis, progressing from the model with the fewest controls to the most heavily controlled model. We also display results for both initial and final inspection scores, to see if restaurants' behaviors change when they have time to adjust and respond to the new grading regime. The initial inspection score results are the strictest test of improved food safety compliance, assessing if restaurants respond to the new policy in time for their first inspection under the new regime and maintain improved compliance each inspection cycle forward. The final inspection score results test the extent to which restaurants learn from poor initial inspection performance and maintain higher levels of food safety compliance after failing to earn an *A* on initial inspection. In addition, final scores are most closely representative of posted grades, though restaurants can challenge these results in court. In the first column of Table 2, we see that after the implementation of the grading policy initial inspection scores go down (i.e. health conditions improve) by about 1.3 points on average per inspection. This is about 6% percent of the sample mean in the pre-period. When we include additional controls, such as seasonal and ZIP fixed effects, restaurant characteristics and a pre-post trend line, the coefficient on *Post* turns positive and decreases slightly in magnitude. The coefficient on *Post* remains positive and significant when we include restaurant fixed effects (instead of ZIP fixed effects and time-invariant restaurant characteristics) and then when we include *Post_trend*. The final, most fully specified, model indicates that upon policy implementation initial inspection scores go up by about 1.2 points, but decline over time by about .33 per quarter, implying that mean initial inspection scores improve starting about one year after policy implementation. In general, we note that the other covariates display generally expected signs: initial inspection scores are lower (better) for chains and uncorrelated with number of seats and number of workers. There is also variation in scores depending on cuisine.

Table 3 shows the same models using final inspection scores. Just as above, the first column displays a significant and negative coefficient on *Post*, but much larger in magnitude: final inspection scores decline by about 7.6 points on average. As we add in controls to the model, the coefficient on *Post* remains negative and decreases slightly in magnitude, and in the final, preferred model, inspections scores decline by about 4 points on average upon implementation of the grading policy. Scores continue to decline over time, at about .29 per quarter. Since final inspection scores reflect food safety conditions after feedback or general learning from initial inspections, it is not surprising that the immediate effect (i.e. the coefficient on *Post*) is negative, or an improvement in food safety conditions (unlike the positive coefficient observed in regressions on the initial scores). The fact, however, that both initial and final scores decline over time after policy implementation does suggest improved food safety compliance. In general, we note that the other covariates display generally expected signs: final inspection scores are lower (better) for chains and uncorrelated with the number of workers and number of seats. There is also variation in scores depending on cuisine with direction of those relationships generally consistent with the relationships observed in the results for initial inspection scores.

As an alternative specification, we exploit the policy's roll out period to identify the impact of the policy change, i.e. the required posting of a grade, on inspection scores. One concern with this approach is that the restaurants exposed earlier to the policy were systematically different than those exposed later. To check this, we look at a range of statistics describing restaurant characteristics and sanitary conditions—these are displayed in Appendix B. In general, we find no meaningful difference between the early- and late-inspections, and fail to reject the null of group equivalence in a joint-significance F-test.[8] This mitigates some concerns of selection bias, based on observed characteristics (which we assume are at least somewhat correlated with unobserved characteristics). We will also be controlling for restaurant-level characteristics in the regression models, essentially implementing a within-restaurant comparison over time and further reducing unobserved heterogeneity across the restaurants that could introduce bias into the impact estimates.

[8] We also find little or no difference between the sample for the roll-out analysis and the larger sample for the pre-post analysis.

The results for the roll-out regression analysis are displayed in Table 4, and once again we begin with the most parsimonious model, controlling for restaurant characteristics and time trends. We show results for both initial and final inspection scores results. We see that both initial and final scores go down, although the declines are bigger for the latter measure, which is consistent with what we observed in the analysis on the full sample. Again, this is likely due to some combination of restaurant learning and improved compliance over somewhat shorter re-inspection windows. In the most fully specified models, when restaurant fixed effects are included, initial inspection scores for those restaurants exposed to the grading policy (compared to those not yet exposed) go down by just under 1 point. This is compared to a decline in final inspection scores of almost 4 points per inspection. We recognize that there could be a period of adjustment, even during the roll-out period. To test for this, we replicate the roll-out analysis, allowing the effect of the graded inspection to vary across time. These results are displayed in Appendix C. We see that the immediate effect of the graded inspection is positive for initial inspections and negative for final ones, and that over the course of the roll-out period this effect progressively becomes more negative (i.e. scores are improving). Thus, by the end of the first year of the grading policy, mean initial and final inspection scores are both lower than they were before public grading. Again, this is consistent with the findings from the pre-post analysis.

Altogether, the results for inspection scores indicate a period of adjustment on the part of the restaurants, which initially see a slight bump up in scores and then a decline over time. The initial increase in scores could mean two things. First, it suggests that restaurants were changing their food safety compliance behaviors in response to the policy (and the feedback from the inspections), but that it took time for it to manifest itself in the actual restaurants' conditions. The initial scores (and therefore food safety conditions) could have in many ways been more reflective of restaurants' conditions prior to the start of the grade-posting policy, and that later inspection scores are a product of their response to the change in policy. That is, the first set of initial inspection scores are not lower immediately following the policy, but progressively improve in the second and third inspection cycles. This pattern is supported by the descriptive statistics displayed in Table 5, which also show an improvement (i.e. decline) in inspection scores as the program progressed. During the period before public grading, restaurants earned inspection scores of 24.6 on average. In the first five quarters after public grading, mean final inspection scores improved to 18.2. This initial improvement was driven by improved compliance during re-inspections, as the initial scores are about level with the pre-implementation scores. In the next five quarters, average final inspection scores further improved to 15.6. This second improvement was driven in part by improved compliance on initial inspections (which on average went down to 22.2). We find similar changes in mean inspection scores for "continuously operating" restaurants that operate for two and half years before and two and half years after public grading. A second explanation for the initially increasing scores and subsequent decline relates to the inspectors' changing behavior—that they had incentives to improve scores under the new grading regime regardless of actual food safety compliance. While we cannot test this directly, there are two reasons why we think this mechanism is unlikely (and that any improvement in scores predominantly reflects an improvement in food safety compliance). First, inspectors are randomly assigned to their site visits, and therefore restaurants are dealing with different individuals for the initial and final inspections—it is unlikely that the randomly assigned final inspector would be colluding with the initial inspector to systematically reduce scores, other than based on an observed improvement in food safety compliance. Second, there are no new incentives in the grading policy to motivate inspectors to deflate inspection scores—their roles are unchanged from before the start of grading.[9]

b. Fines

Next we consider the impact of the grading policy on fines; this is a good indicator of financial benefits for the City (and conversely, financial burdens on the restaurants). One concern with the policy was that it was an excuse for the city to draw more revenues; we test the validity of this claim here. To start, we consider Figure 1, which shows mean restaurant fines by quarter. While fines per restaurant increase in the year immediately following program implementation, this extends a pre-existing trend (that temporarily discontinues in the second quarter of 2011, during program implementation). Quarterly fines reach a peak of $675 per operating restaurant in the first quarter of 2012 and then decline steadily, reaching pre-program levels by the third

[9] The new policy did include the hiring of more inspectors, but there is no evidence to suggest that the new inspectors were more lenient than the older ones; and again, they were still being randomly assigned to inspections. In addition, while the higher stakes of the posted grades could change the nature of the interaction between inspectors and restaurants, it is unlikely that this shift would be so systematic as to drive the effects we observe.

quarter of 2013 ($353 in fines for the average restaurant). The question is whether the post-grading trend is significantly different than what would have continued otherwise—we now turn to the results for the pre-post analysis on fines (displayed in Table 6).

Starting with the least specified model in the first column, we see that fines declined upon the policy's implementation. The coefficient on *Post* is negative and highly significant and indicates that on average fines went down by $271 per inspection. When we add in restaurant controls and ZIP and quarter fixed effects, the magnitude on the *Post* coefficient goes down substantially, but still remains negative: fines reduced by about $62 per inspection after the grading policy's implementation. When we instead rely on restaurant fixed effects, the coefficient on *Post* flips its sign to positive and retains this sign in our fully specified model in the final column of Table 6. Ultimately, we see that upon policy implementation, fines increase (by about $65 per inspection), but they decline over time, such that by the second quarter after implementation any increase in fines had been reversed. This immediate increase in fines is consistent with the short-term increase in initial inspection scores, which also goes down over the first year of the policy. Again, restaurants are likely adjusting to the new policy regime, which includes increased inspection frequency for restaurants with poor food safety compliance. The initial learning period may not reflect the full (and long-lasting) effects from the grade-posting. Altogether, these results suggest that the grading policy did not serve to increase fine-driven revenues for the city (which presumably would have been the case in the absence of the grading policy, based on the pre-post trend).

Again, we repeat the analysis using the roll-out sample, comparing those restaurants that were exposed to the grading regime earlier in the rollout period to those that were exposed later. The first column of Table 7 displays the results for the model without restaurant fixed effects; those restaurants exposed earlier to the grading regime pay higher fines on average than those later exposed—about $44 more per inspection. These results, however, could be due to a targeted roll-out strategy (for example, the restaurants subject to earlier inspections did exhibit higher fines; these statistics are displayed in Appendix B). Therefore, we add in restaurant fixed effects so that we can compare fines within restaurants, and the magnitude of the coefficient on *Post_Rollout* goes down and flips to a negative sign. Therefore, in the more fully-controlled model, fines go down by about $31 per inspection over the course of the initial year of the policy. This decline is consistent with that observed in the larger pre-post sample, which also produced fine declines over the course of the initial year (even though the immediate effect on fines was positive).

c. Sales revenues and taxes

Again, as a starting point, we look at unadjusted trends over the course of the study period. Figure 2 shows mean sales by quarter for the 10 quarters before public grading and the 9 quarters after. Mean sales remain similar in the post-grading period as compared to the pre-grading period in real 2013 dollars, though mean sales revenues rise slightly. Mean sales figures for this sample are quite high (over $150,000 a quarter) and mask a fairly high level of heterogeneity across restaurants. Figure 3 shows a similar story for sales taxes (as anticipated since sales taxes are mechanically derived from sales revenue). There is a small decline in sales taxes during the period 2 years to 1 year before program implementation. This trend reverses before public grading and mean taxes remain largely unchanged after public grading. Mean sales taxes are between $8,000-$9,000 a quarter in our study period.

To test whether or not this trend persists in the presence of restaurant and temporal controls, we consider the regression results in Tables 8 and 9, which display the results for the pre-post and roll-out models, respectively. The first column of Table 8 shows the results for the most parsimonious model and we see that sales for the average restaurant increased by about $10,700 after the implementation of the grading policy. In the most fully specified model, with seasonal and bin fixed effects as well as restaurant characteristics, the magnitude of the coefficient on *Post* goes down slightly, but remains positive and highly significant. Sales revenues, for the average restaurant, go up by about $8,000 immediately following the policy's implementation and continue to increase by about $730 per quarter. However, we note that over the same time period, we also see a sustained increase in total sales taxes (which includes revenues from other activities, such as retail trade more broadly and businesses services) and hotel taxes (and presumably revenues) (NYC OMB 2014), which makes it challenging to conclude that any observed increase in restaurant sales revenue is independently due to the implementation of the grading policy. In addition, compared to the time trends for inspection scores and fines, which display clear reversals in slopes (suggesting that changes are grading-induced), the trends for sales revenues are less discontinuous across pre- and post-grading periods.

The results for the roll-out sample (displayed in the next table) confirm this ambiguity. In the most fully specified model, the coefficient on *Post* is negative and smaller in magnitude (almost half the size) than that in the pre-post results, albeit statistically insignificant. These results suggest that within the first year of the program, there was no significant revenue effect on those exposed earlier to the grading regime compared to those exposed later.[10] The difference in these results and those from the pre-post analysis could be explained by some adjustment on the part of consumers in using the grades to dictate their dining choices. In addition, the program may not have provided clear information to consumers in the first couple of months of the policy. Only restaurants subject to earlier exposure and earning *A*'s posted grades in the first couple of months post-grading. Furthermore, the time between initial inspection and re-inspection is 2 to 4 weeks and between re-inspection and adjudication is between 4 to 6 weeks; treated restaurants could post nothing during the first window and *Grade Pending* during the second window if they did not earn an *A* initially. It is not clear whether or not this distinction was meaningful enough to influence dining choices, and therefore revenues, during the initial months of the policy. Over time, however, it is more likely that grading-induced sorting was reflected in the restaurants' business activity (and their subsequent revenues), some of which could be captured by the pre-post estimates.

Not surprisingly, the results for sales taxes echo those for restaurant sales revenues (indeed the former is a function of the latter). Tables 10 and 11 display these results. The pre-post analysis indicates that the grading policy has an immediate positive impact on sales taxes, and one that continues to increase over time. Specifically, controlling for temporal and restaurant characteristics, taxes increase by about $225 per restaurant and increase by about $107 per quarter following the policy's implementation. As with sales revenues, the coefficient on *Post_Rollout* is insignificant (and negative), suggesting no immediate differential impact on sales tax revenues (between those inspected earlier and later under the grading regime).

d. Falsification tests

[10] We also tested for differential revenue effects across the roll-out period, but none of these results were statistically significant. These results are available from the authors upon request.

We implement one final test to see if the impact we observe is actually due to the grading policy implementation (i.e. the posting of the grade), and not just a stage in the process of adjusting to the policy that was perhaps anticipated in some way before the policy's actual start. To do this, we replicate the fully specified models (the ones with restaurant fixed effects and pre-post and post linear trends), assigning the policy start date one year prior to the actual policy start date. We find that the program was already being discussed in the popular press one year prior to the program's actual start, and so we use this date for our falsification tests. The results from this analysis are displayed in Table 12, which display a model including *Post_false, Post_trend_false* and also *Post* and *Post_trend*.[11]

In the case of initial and final scores, we see that the false start date is associated with an immediate decline in scores (which is statistically insignificant for final scores), and thereafter a similar increase per quarter. The true policy start date (while still controlling for the false start date) is associated with statistically significant declines for both initial and final scores (the latter one about ten times larger), and then a continued decline over the post-period. The discontinuity in intercept and reversal in slope both suggest that any decline in inspection scores is in fact associated with the grade posting and not a continuation of a prior trend or expectation.

For fines, we conduct the same test and see that the false start date is associated with an increase in fines and a positive slope thereafter. The actual policy start date, however, indicates a drop in fines and a subsequent decline (that rather quickly reverses any prior increase in fines). Again, this reversal in effect suggests that any initial bump up in fines (as observed in the pre-post analysis) could be driven by trends prior to the actual start date, and that the grading policy itself seems to be associated with a sustained drop in fines.

Finally, we run falsification tests for the sales revenue/tax analysis. When we include in the model indicators for the false policy start date, we find that both the initial and continued effects are significant and positive. Upon the actual policy start date, revenues go up (by almost 30 percent less than at the time of the false start) and there is no significant trend after that initial

[11] We run more parsimonious models, including only *Post_false* and *Post_trend_false*, but for purposes of brevity we display only the more comprehensive models since they better control for all the possible points of inflection during the pre- and post-grading regimes. We further specify models with quadratic and cubic time trends, finding similar results which are available upon request of the authors.

effect. Results for sales taxes display the same patterns except that there is no significant effect at the time of actual policy implementation (controlling for any announcement effect one year prior) and tax revenues decline thereafter. The more modest change in intercept between false and actual policy starts makes it difficult to conclude that the policy imposed any shocks in revenue for the restaurants; however the absence of any grade-induced revenue trend reinforces the null effect observed in the roll-out models. Together, we can infer that at best the grading policy had a small immediate effect on restaurant revenues, and at worst the average restaurant experienced a slight decline in revenues in the quarters thereafter. The fuzziness in the discontinuity in the falsification tests (as compared to the starker shifts for scores and fines) confirms any hesitation about the independent impact of the grading policy on sales revenues (and taxes). It is possible that any noise in the results is likely due to the broader economic context, which was still in the beginnings of a post-recession recovery at about the same time as the grading policy's implementation.

VII. Conclusion

Cities have long inspected restaurants for their sanitary conditions, but the availability of that information and its incorporation into the regular dining decisions of patrons is a relatively new phenomenon. The motivation for a grading policy is to make the sanitary conditions of a restaurant more transparent and uniformly understood as a means of reducing the incidence of foodborne illnesses. Therefore, it is very much a health policy. However, if theory is correct, and consumers use this information to change their behaviors, restaurants (and the municipal fisc) could bear economic repercussions as well.

We systematically test these predictions using a rich panel dataset on restaurants' food safety compliance and sales activity, both before and after the implementation of the grading policy in New York City. Our data compilation efforts are not only an example of coordinating across multiple administrative datasets (and agencies), but also of extracting valuable information from otherwise confidential data (i.e. business' sales revenues). Altogether, the results from this project shed light on the direct food safety compliance goals of the grading policy, as well as the broader economic implications for businesses and the local fisc.

Our results suggest that New York City's restaurant grading policy does improve sanitary conditions (as measured by inspection scores), and, after an adjustment period, reduces revenues collected through fines and unconvincingly affects sales revenues (and taxes). Specifically, final inspection scores decline (i.e. sanitation improves) by about 4 points (about 17% of the pre-period mean) upon policy implementation and then continue to decline at about ¼ point per quarter. Fines increase immediately after that start of the grading policy (by about $65 per inspection, or 6% percent of the pre-period mean fine), but decline thereafter such that any gain is reversed by the second quarter post-implementation. It is difficult to tease out any independent impact on sales revenues (and taxes) from the grading policy, as it is unclear the extent to which the grading policy increases revenues or if increases in revenues reflect broader economic recovery trends (as indicated by growth of revenues for more inclusive retail trade and hotel tax forecasts). Furthermore, models that estimate effects during the policy's roll-out year show no significant revenue change for restaurants that are first exposed to the grading regime, compared to those that are later exposed.

Therefore, the health goals, as they relate to the restaurant's food safety environment, seem to be addressed through the improvement of inspection scores. Moreover, this is in addition to any tangible fiscal savings from reduced incidences (and treatments of) foodborne illnesses (which we were unable to measure directly here). However, we should consider this up against the increased administrative and labor costs required of the program (an amount that, according to our calculations, is between $245 and $320 per inspection, averaging approximately $2.3 million in total annually to the City[12]). We can already see, that these costs are not being mitigated by significant program-induced revenues, such as fines and/or sales taxes, in the long run. This is in contrast to the fear that the policy was simply a mechanism for raising revenues for the City.

Apart from the City's overall welfare, we should also be concerned about the distributional effects of such a policy. While the current analysis obscures any variation across restaurants over time, a related paper (using the same case and dataset) finds that there are meaningful differences in economic performance across restaurants with different grades: restaurants that

[12] The average cost per inspection is calculated by dividing all spending in DOHMH's Food Safety budget by the reported number of actual restaurant inspections. Budget figures come from OMB Budget Function Analysis, and actual inspections come from DOHMH reporting to OMB.

post *A*s are less likely to close, owe fewer fines and bring in more revenues compared to *B* restaurants (Schwartz et. al. 2015). Therefore, the relative benefits and burdens of the policy differ across restaurants. Certain restaurants may be able to more easily absorb the costs of managing higher stakes inspections and will likely benefit more from improved compliance.

Altogether, these results indicate that, after some period of adjustment, grading policies can be a powerful tool to improve compliance (in this case, improving sanitary conditions on the part of the restaurant), presumably through posting these conditions in conspicuous locations. We interpret this as an indication that restaurants expect and/or observe that consumers will sort towards the establishments with higher grades, and change their food compliance practices accordingly. The inconclusive revenue results, however, suggest that restaurants overall are not necessarily benefitting from more business—but they are not being deprived of business either. Therefore, the economic implications appear to be less pronounced than any health-related food safety outcomes, on which the policy clearly improves.

VIII. References

Fein, S. B., Lin, C. T. J., & Levy, A. S. 1995. Foodborne illness: perceptions, experience, and preventive behaviors in the United States. *Journal of Food Protection*, 58(12): 1405-1411.

Figlio, David N., and Maurice E. Lucas. 2004. "What's in a grade? School report cards and the housing market." *American Economic Review*, 94(3): 591-604.

Filion, Katie, and Douglas A. Powell. 2009. "The use of restaurant inspection disclosure systems as a means of communicating food safety information." *Journal of Foodservice*, 20(6): 287-297.

Freebairn, John W. "Grading as a market innovation." Review of Marketing and Agricultural Economics 35.3 (1967): 147-162.

Freebairn, John W. "The value of information provided by a uniform grading system." Australian Journal of Agricultural Economics 17.2 (1973): 127-139.

Ho, D.E. (2012). Fudging the Nudge: Information Disclosure and Restaurant Grading. *Yale Law Journal*, Vol. 112(3): 574-688.

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

Jones TF, Angulo FJ. 2006. Eating in restaurants: a risk factor for foodborne disease? *Clin Infect Dis.*, 43(10):1324-1328.

Mead, Paul S., Laurence Slutsker, Vance Dietz, Linda F. McCaig, Joseph S. Bresee, Craig Shapiro, Patricia M. Griffin, and Robert V. Tauxe. 1999. Food-related illness and death in the United States. *Emerging infectious diseases*, 5(5): 607-625.

New York City Department of Health and Mental Hygiene. 2012. How We Score and Grade. Retrieved from http://www.nyc.gov/html/doh/downloads/pdf/rii/how-we-score-grade.pdf.

New York City Office of Management and Budget. 2014. *Tax Revenue Forecasting Documentation*. Retrieved from http://www.nyc.gov/html/omb/downloads/pdf/methodology_2014_08.pdf

Rockoff, J.E., and Turner, L.J. (2010). Short-Run Impacts of Accountability on School Quality. *American Economic Journal: Economic Policy*, Vol. 2 (4):119-147.

Schwartz, Amy Ellen, Michah W. Rothbart, Thad Calabrese, Zachary Papper, Tod Mijanovich, Rachel Meltzer, Diana Silver and Meryle Weinstein. 2015. “The Impact of Restaurant Letter Grades on Taxes and Sales: Micro Evidence from New York City.” Working paper.

Silver, D., Rothbart, M., Bae, J., Schwartz, A.E., Mijanovich, T. (in preparation). *Public Restaurant Grades and Inspection Adjudication in New York City*. Manuscript in preparation.

Simon, P.A., Leslie, P., Run, G., Jin, G.Z., Reporter, R., Aguirre, A., Fielding, J.E. (2005). Impact of Restaurant Hygiene Grade Cards on Foodborne-Disease Hospitalization in Los Angeles County. *Journal of Environmental Health*, Vol. 67: 32-36.

Thaler, Richard H. and Cass R. Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New York: Penguin Books.

Winters, M.A., and Cowen, J.M. (2012). Grading New York: Accountability and Student Proficiency in America's Largest School District. Educational Evaluation and Policy Analysis. http://epa.sagepub.com/content/early/2012/05/14/0162373712440039

Wong, M.R., McKelvey, W., Ito, K., Schiff, C., Jacobson, B., and Kass, D. (2015). Impact of Letter Grade Program on Restaurant Sanitary Conditions and Diner Behavior in New York City. *American Journal of Public Health* 105(3): 81-87.

Table 1. Restaurant Descriptive Statistics, Full Sample

	Pre-Public Grading	Post-Public Grading
Number		
Inspections	3.3	6.2
Final Inspections	3.3	3.2
Workers	6.5	6.7
Seats	29.6	29.5
Cuisine		
American	0.22	0.24
Chinese	0.09	0.11
Pizza	0.04	0.06
Latin	0.04	0.04
Café/Coffee/Tea	0.03	0.04
Others	0.38	0.51
Missing	0.20	0.00
	1.00	1.00
Service		
Takeout-Limited Eat in	0.35	0.39
Wait Service	0.15	0.18
Wait and Counter Service	0.11	0.17
Takeout Only	0.08	0.08
Counter Service	0.07	0.12
Others	0.06	0.07
Missing	0.20	0.00
	1.00	1.00
Chain	0.10	0.10
Annual Closure Rate	0.16	0.12
N	30,405	34,917

Notes: Inspections include initial and re-inspections. Final inspections include all inspections in the pre-period, initial *A* inspections in the post-period, and re-inspections for those initially receiving *B* or *C* in the post-period. Workers, seats, cuisine, service, and chain reflect restaurant characteristics at the most recent restaurant inspection and are time-invariant variables. Annual closure rate is the fraction of open restaurants closing each year.

Table 2. Regression Results, Impact on Initial Inspection Scores, Pre-Post Sample

VARIABLES	(1)	(2)	(3)	(4)
Post	-1.33***	1.21***	2.41***	1.23***
	(0.09)	(0.17)	(0.20)	(0.20)
Post*Linear Trend	__	__	__	-0.33***
				(0.04)
Linear Trend	__	-0.22***	-0.28***	-0.10***
		(0.01)	(0.02)	(0.03)
Seasonal FE	N	Y	Y	Y
Rest. Char.	N	Y	N	N
Restaurant FE	N	N	Y	Y
Constant	24.15***	16.10***	22.68***	23.40***
	(0.07)	(4.16)	(0.13)	(0.16)
Inspections	159,588	159,588	116,228	116,228
Restaurants	41,362	41,362	20,641	20,641
R-squared	0.00	0.06	0.30	0.30

Notes: Robust standard errors clustered by restaurant in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Initial Inspection Score reflects restaurants inspection score on the first inspection each cycle. Post takes a value of 1 if the quarter is after the implementation of public grading. Linear Trend reflects the number of quarters before/after public grading (negative for number before and positive for number after). Post*Linear trend reflects the number of quarters into the public grading regime and takes a value of 0 in the pre-period. Seasonal FE are fixed effects for quarters with quarter 1 running from December through February, quarter 2 from March through May, and so forth. Restaurant characteristics include indicators variables for zip code, chain, cuisine, venue, and service type, seats, and number of workers.

Table 3. Regression Results, Impact on Final Inspection Scores, Pre-Post Sample

VARIABLES	(1)	(2)	(3)	(4)
Post	-7.62***	-5.08***	-4.25***	-4.05***
	(0.08)	(0.16)	(0.18)	(0.18)
Post*Linear Trend	__	__	__	-0.29***
				(0.03)
Linear Trend	__	-0.20***	-0.18***	-0.05*
		(0.01)	(0.02)	(0.03)
Seasonal FE	N	Y	Y	Y
Rest. Char.	N	Y	N	N
Restaurant FE	N	N	Y	Y
Constant	24.55***	13.27***	22.57***	23.12***
	(0.07)	(2.33)	(0.12)	(0.14)
Inspections	167,045	167,045	125,036	125,036
Restaurants	40,554	40,554	32,142	32,142
R-squared	0.06	0.11	0.31	0.31

Notes: Robust standard errors clustered by restaurant in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Final Inspection Score reflects restaurants inspection score on the final inspection each cycle. Post takes a value of 1 if the quarter is after the implementation of public grading. Linear Trend reflects the number of quarters before/after public grading (negative for number before and positive for number after). Post*Linear Trend reflects the number of quarters into the public grading regime and takes a value of 0 in the pre-period. Seasonal FE are fixed effects for quarters with quarter 1 running from December through February, quarter 2 from March through May, and so forth. Restaurant characteristics include indicators variables for zip code, chain, cuisine, venue, and service type, seats, and number of workers.

Table 4. Regression Results, Impact on Inspection Scores, Rollout Sample

	Initial Inspection Score		Final Inspection Score	
VARIABLES	(1)	(2)	(3)	(4)
Graded Inspection	-3.553***	-0.949***	-3.382***	-3.778***
	(0.184)	(0.159)	(0.181)	(0.160)
Quarter Of Grading Policy				
2	0.696***	0.379***	1.217***	1.573***
	(0.054)	(0.053)	(0.071)	(0.070)
3	0.543***	-0.071	0.574***	1.233***
	(0.091)	(0.066)	(0.123)	(0.093)
4	0.635***	0.334***	0.756***	2.051***
	(0.112)	(0.073)	(0.150)	(0.113)
Restaurant FE	N	Y	N	Y
Constant	25.981***	25.536***	20.219***	19.791***
	(0.109)	(0.059)	(0.099)	(0.066)
Observations	122,886	122,886	93,788	93,788
R-squared	0.007	0.848	0.011	0.763

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Initial Inspection Scores are inspection scores on a restaurant's initial inspection. Final Inspection Scores are restaurant final inspection scores for each inspection cycle. Quarter 1 post grading is the omitted period. Graded Inspection measures if a restaurant's most recent inspection occurs after the restaurant grade program begins in July 2010. Graded Inspection is the share of days the restaurant is exposed to the grading regime in the quarter. Sample includes all restaurants open during the first four quarters of restaurant grading.

Table 5. Inspection Scores and Count of Inspections by Treatment Period, Pre-Post Sample

		Inspections: All Operating Restaurants	Inspections: Continuously Operating Restaurants
Pre		25.07 (76,231)	23.50 (29,804)
Quarters Post			
1-5	Initial	25.31 (41,933)	24.26 (17,723)
	Final Inspection Score	21.87 (27,874)	20.88 (11,743)
6-10	Initial	22.46 (46,180)	21.62 (18,409)
	Final Inspection Score	19.52 (29,135)	18.78 (11,544)

Includes pre-adjudicated inspection scores. Mean score shown on top; number of inspections shown parenthetically. Final Inspection Score is the mean restaurant inspection score for final inspections each cycle (all A-graded initial inspections and re-inspections of restaurants that do not get an A grade on initial inspection). Continuously Operating Restaurants are open for every quarter of the sample period.

Table 6. Regression Results, Impact on Inspection Fines, Pre-Post Sample

VARIABLES	(1)	(2)	(3)	(4)
Post	-270.72***	-61.59***	88.67***	65.26***
	(5.77)	(10.79)	(10.54)	(10.52)
Post*Linear Trend	—	—	—	-55.98***
				(1.84)
Linear Trend	—	-14.35***	-22.16***	11.10***
		(0.87)	(0.89)	(1.45)
Seasonal FE	N	Y	Y	Y
Rest. Char.	N	Y	N	N
Restaurant FE	N	N	Y	Y
Constant	1,141.52***	245.74*	947.74***	1,081.79***
	(5.03)	(137.72)	(7.04)	(8.44)
Inspections	233,642	233,642	172,098	172,098
Restaurants	41,362	41,362	32,142	32,142
R-squared	0.01	0.05	0.28	0.29

Notes: Robust standard errors clustered by restaurant in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Fines reflect restaurants' fines each inspection. Post takes a value of 1 if the quarter is after the implementation of public grading. Linear Trend reflects the number of quarters before/after public grading (negative for number before and positive for number after). Post*Linear Trend reflects the number of quarters into the public grading regime and takes a value of 0 in the pre-period. Seasonal FE are fixed effects for quarters with quarter 1 running from December through February, quarter 2 from March through May, and so forth. Restaurant characteristics include indicators variables for zip code, chain, cuisine, venue, and service type, seats, and number of workers.

Table 7. Regression Results, Impact on Fines by Quarter, Rollout Sample

VARIABLES	(1)	(2)
Graded Inspection	50.918***	-132.357***
	(10.547)	(13.919)
Quarter Of Grading Policy		
2	95.591***	163.405***
	(9.496)	(10.108)
3	102.872***	235.971***
	(12.863)	(15.448)
4	99.762***	279.118***
	(11.105)	(15.132)
Restaurant FE	N	Y
Constant	358.596***	333.149***
	(5.546)	(5.741)
Observations	94,752	94,752
R-squared	0.003	0.551

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Fines Per Quarter measures average fine assessed on a restaurant during each inspection cycle by quarter. Graded Inspection measures if a restaurant's most recent inspection occurs after the restaurant grade program begins in July 2010. Graded Inspection is the share of days the restaurant is exposed to the grading regime in the quarter. Sample includes all restaurants open during the first four quarters of restaurant grading.

Table 8. Regression Results, Impact on Sales by Quarter, Food and Beverage Pre-Post Sample

VARIABLES	(1)	(2)	(3)	(4)
Post	10,697.98***	18,980.28***	8,892.17***	7,936.05***
	(2,238.69)	(1,847.45)	(915.70)	(908.06)
Post*Linear Trend	__	__	__	729.23***
				(208.03)
Linear Trend	__	1,356.40***	452.47***	180.88
		(225.99)	(141.66)	(183.78)
Seasonal FE	N	Y	Y	Y
Rest. Char.	N	Y	Y	Y
Group FE	N	N	Y	Y
Constant	174,368.92***	253,566.97***	204,764.41***	205,187.51***
	(4,153.39)	(41,492.83)	(58,240.93)	(58,156.70)
Observations	24,464	24,464	18,897	18,897
Restaurant-Quarters	254,216	254,216	195,279	195,279
R-squared	0.00	0.42	0.95	0.95

Notes: Robust standard errors clustered by group in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Revenues measures mean bin sales by quarter, using aggregated group data. Post takes a value of 1 if the quarter is after the implementation of public grading. Linear Trend reflects the number of quarters before/after public grading (negative for number before and positive for number after). Post*Linear Trend reflects the number of quarters into the public grading regime and takes a value of 0 in the pre-period. Seasonal FE are fixed effects for quarters with quarter 1 running from December through February, quarter 2 from March through May, and so forth. Restaurant characteristics include indicator variables for borough, chain, number of workers, and building class. Group FE are fixed effects for restaurant grouping.

Table 9. Regression Results, Impact on Sales by Quarter, Food and Beverage Rollout Sample

VARIABLES	(1)	(2)
Graded Inspection	31,552.417	-4,485.497
	(39,710.175)	(6,686.311)
Quarter Of Grading Policy		
2	-7,763.466	1,866.532
	(10,710.803)	(2,298.146)
3	-26,199.939	-4,923.041
	(23,558.118)	(4,723.712)
4	-16,070.555	11,821.031***
	(31,101.395)	(5,603.230)
Group FE	N	Y
Constant	212,102.946***	212,970.365***
	(5,506.497)	(1,023.340)
Observations	3800	3800
Restaurant-Quarters	39,188	39,188
R-squared	0.002	0.981

Notes: Robust standard errors clustered by group in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Impact estimates of graded inspections on restaurant revenues by quarter, using aggregated group data. Graded Inspection measures identifies if most recent inspection occurs after the restaurant grade program begins in July 2010. Graded Inspection is the average share of days in a quarter restaurants in the group are exposed to graded inspections. Sample includes all restaurants open during the first four quarters of restaurant grading as well as the quarter prior to grading and the quarter following the first year (to reduce bias resulting from imperfect measures of restaurants opening and going out of business on model estimates).

Table 10. Regression Results, Impact on Sales Taxes by Quarter, Food and Beverage Pre-Post Sample

VARIABLES	(1)	(2)	(3)	(4)
Post	987.16***	599.00***	85.93**	225.69***
	(96.70)	(81.29)	(39.01)	(39.52)
Post*Linear Trend	__	__	__	-106.60***
				(9.42)
Linear Trend	__	132.61***	114.58***	154.28***
		(10.08)	(6.26)	(8.27)
Seasonal FE	N	Y	Y	Y
Rest. Char.	N	Y	Y	Y
Group FE	N	N	Y	Y
Constant	7,364.81***	11,264.64***	9,487.84***	9,426.00***
	(174.07)	(1,834.40)	(1,667.80)	(2,752.04)
Observations	24,464	24,464	18,897	18,897
Restaurant-Quarters	254,216	254,216	195,279	195,279
R-squared	0.01	0.41	0.95	0.95

Notes: Robust standard errors clustered by group in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Sales taxes measures mean bin sales taxes by quarter, using aggregated group data. Post takes a value of 1 if the quarter is after the implementation of public grading. Linear Trend reflects the number of quarters before/after public grading (negative for number before and positive for number after). Post*Linear Trend reflects the number of quarters into the public grading regime and takes a value of 0 in the pre-period. Seasonal FE are fixed effects for quarters with quarter 1 running from December through February, quarter 2 from March through May, and so forth. Restaurant characteristics include indicator variables for borough, chain, number of workers, and building class. Group FE are fixed effects for restaurant grouping.

Table 11. Regression Results, Impact on Sales Taxes by Quarter, Food and Beverage Rollout Sample

VARIABLES	(1)	(2)
Graded Inspection	1,368.139	-125.606
	(1,811.143)	(282.394)
Quarter Of Grading Policy		
2	-323.622	75.535
	(487.026)	(102.862)
3	-1,136.577	-254.665
	(1,073.538)	(210.262)
4	-686.096	469.990*
	(1,417.316)	(243.447)
Group FE	N	Y
Constant	9,573.287***	9,609.241***
	(247.443)	(43.117)
Observations	3800	3800
Restaurant-Quarters	39,188	39,188
R-squared	0.002	0.983

Notes: Robust standard errors clustered by group in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Impact estimates of graded inspections on restaurant sales taxes by quarter, using aggregated group data. Graded Inspection measures identifies if most recent inspection occurs after the restaurant grade program begins in July 2010. Graded Inspection is the average share of days in a quarter restaurants in the group are exposed to graded inspections. Sample includes all restaurants open during the first four quarters of restaurant grading as well as the quarter prior to grading and the quarter following the first year (to reduce bias resulting from imperfect measures of restaurants opening and going out of business on model estimates).

Table 12. Falsification Test, Inspection Scores, Fines, Revenues, and Sales Taxes, Pre-Post With 1 Year Lead

VARIABLES	Initial Score	Final Score	Fines	Sales	Sales Taxes
Policy Start					
Post	-0.64**	-6.76***	-138.92***	5,105.54***	-80.11
	(0.31)	(0.27)	(15.86)	(1,398.83)	(59.94)
Post*Linear Trend	-1.38***	-1.27***	-98.96***	-107.88	-210.41***
	(0.13)	(0.12)	(7.01)	(447.34)	(20.53)
One Year Before					
Post	-0.76**	-0.43	88.83***	7,056.19***	732.68***
	(0.38)	(0.33)	(20.04)	(1,115.81)	(49.68)
Post*Linear Trend	1.30***	1.27***	53.91***	3,614.20***	349.71***
	(0.14)	(0.12)	(7.37)	(548.74)	(25.11)
Linear Trend	-0.34***	-0.34***	-0.32	-2,635.19***	-97.23***
	(0.06)	(0.05)	(2.85)	(333.29)	(13.28)
Seasonal FE	Y	Y	Y	Y	Y
Restaurant/Group FE	Y	Y	Y	Y	Y
Constant	23.15***	22.59***	989.98***	181,983.07***	7,240.30***
	(0.18)	(0.15)	(9.39)	(55,302.25)	(2,558.02)
Inspections	109,197	120,261	159,617	--	--
Restaurants	20,480	20,482	20,485	--	--
Observations	--	--	--	18,234	18,234
Rest-Quarters	--	--	--	188,384	188,384
R-squared	0.31	0.32	0.32	0.95	0.95

Robust standard errors clustered by restaurant in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Initial Score and Final Score reflect restaurants inspection score on the first and final inspection each cycle, respectively. Fines reflect restaurants' fines each inspection. Sales and Sales Taxes reflect mean quarterly sales and sales taxes in a group, respectively. Policy Start Post takes a value of 1 if an inspection occurs after grading policy implementation and takes a value of 0 during the real pre-period. Policy Start Post*Linear Trend is an interaction term between the real policy start date and a temporal trend line. One Year Before Post takes a value of 1 if an inspection occurs no more than one year before grading policy implementation or any time after and takes a value of 0 during the time more than a year before grading. One Year Before Post*Linear Trend is an interaction term between the one year lead period and a temporal trend line. Linear Trend reflects linear changes over time by quarter. Seasonal FE are fixed effects for quarters with quarter 1 running from December through February, quarter 2 from March through May, and so forth. Inspection and fines models include restaurant fixed effects. Sales and sales tax models include group fixed effects.

Figure 1. Average Fines by Quarter, Operating Restaurants

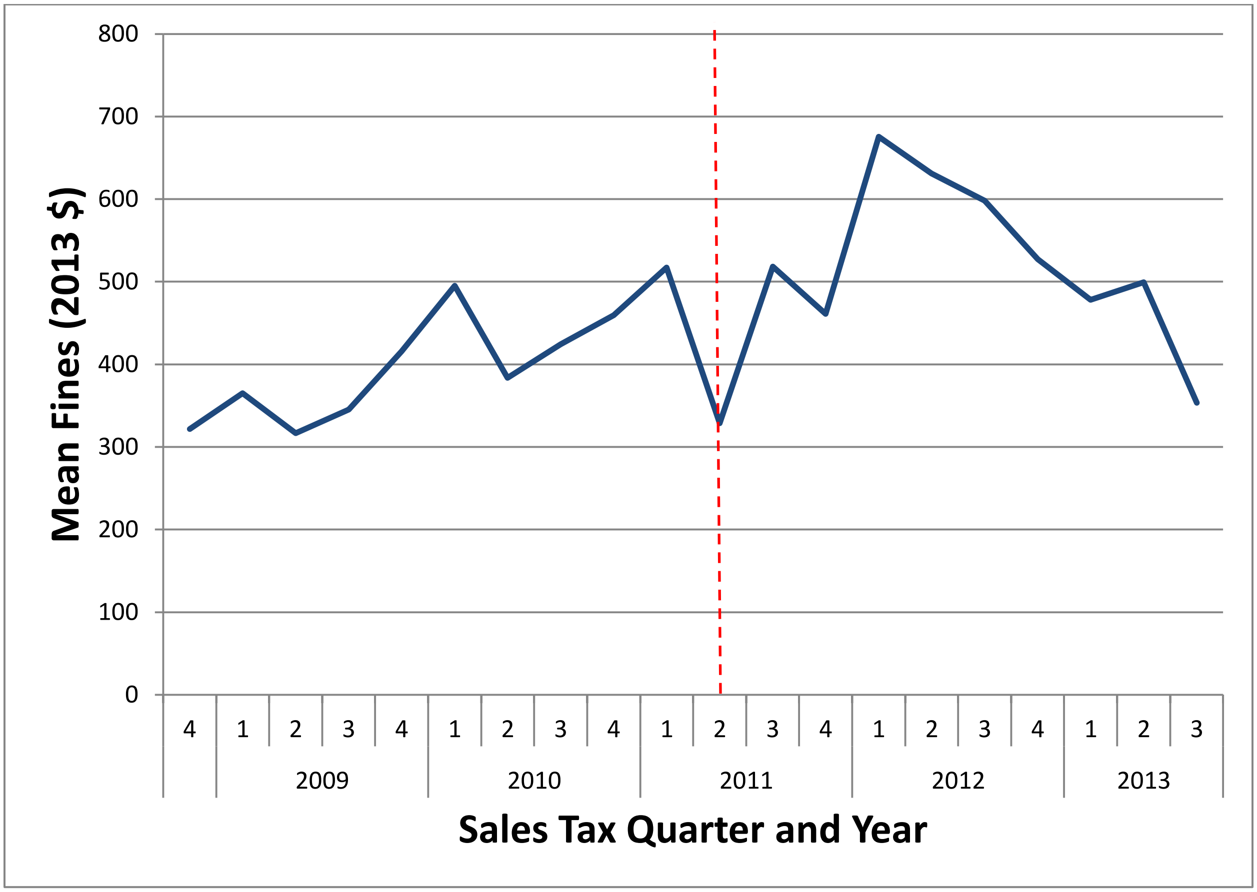


Figure 2. Average Sales, Operating Food And Beverage Entities

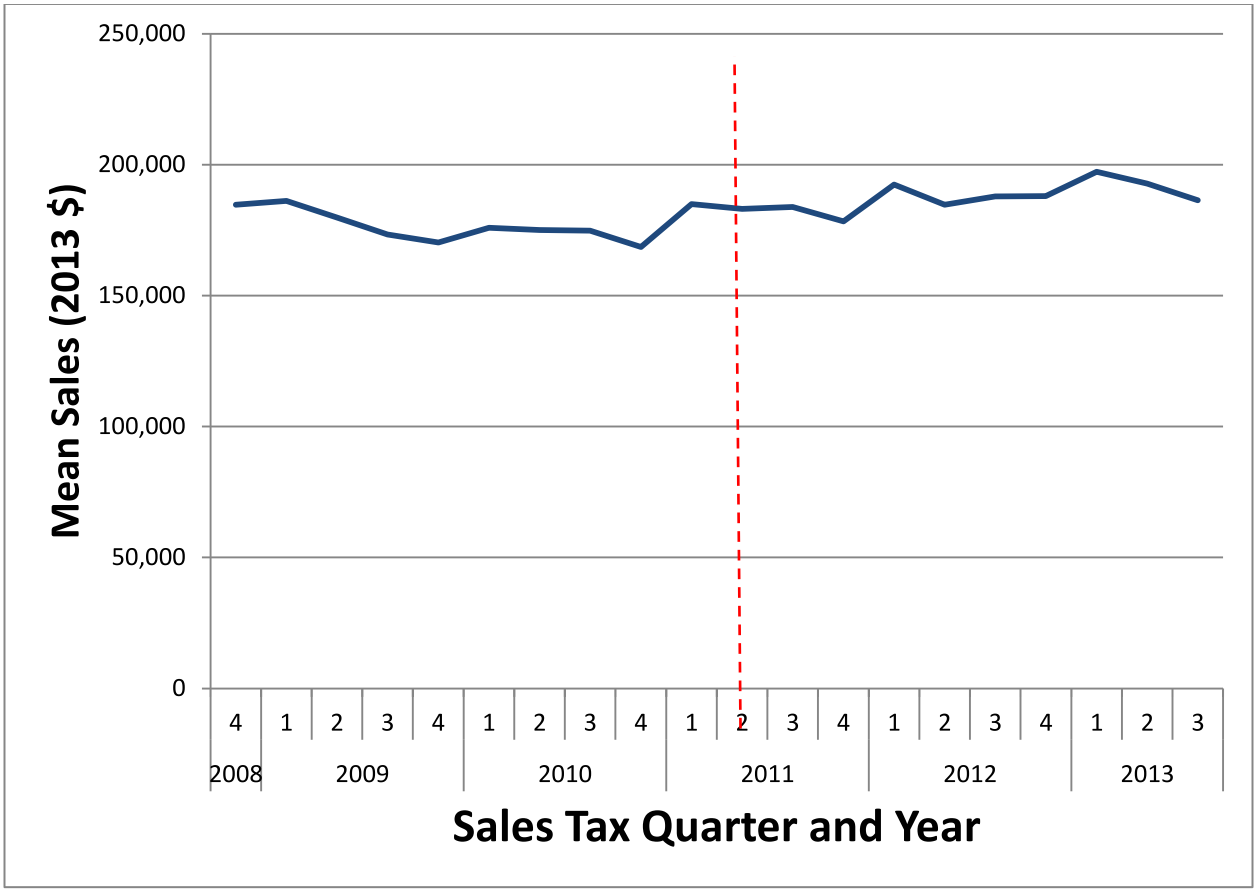
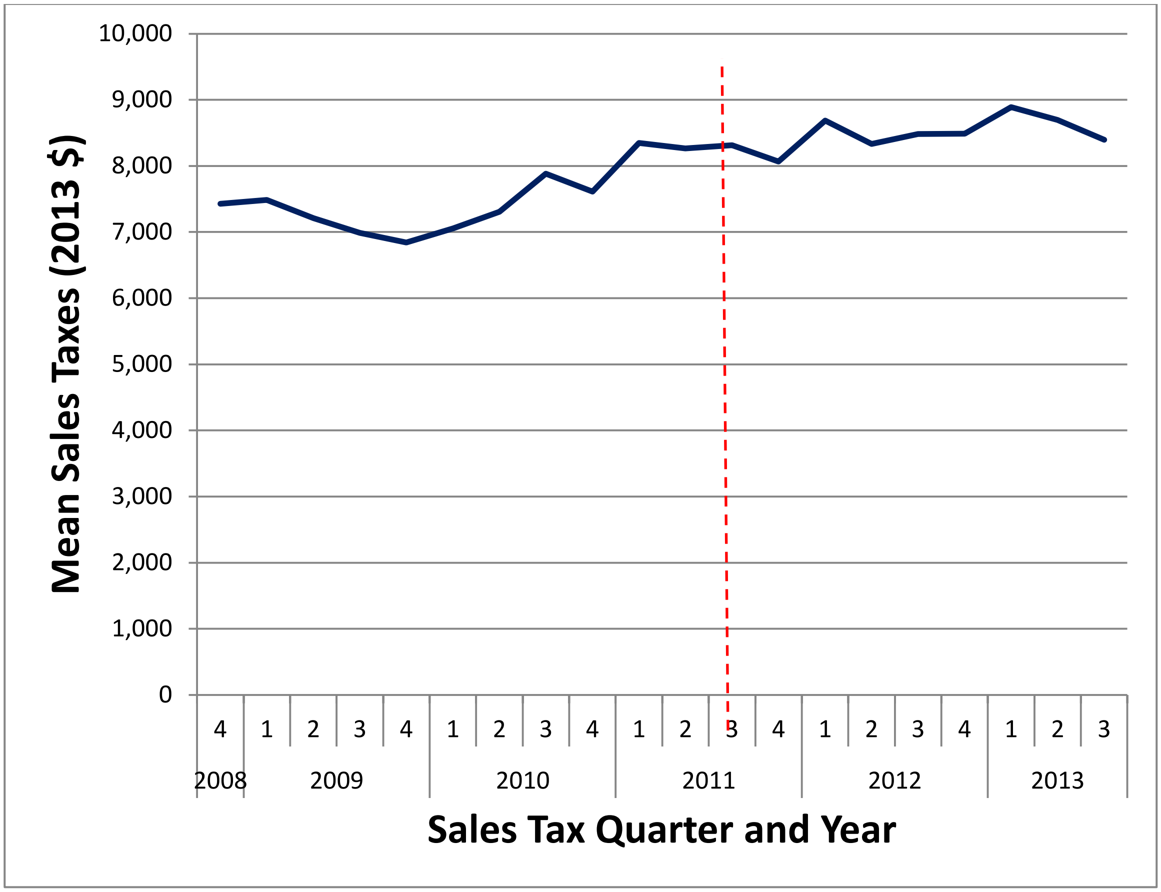


Figure 3. Average Sales Taxes, Operating Food And Beverage Entities



Appendix A. Sample Posted B Grade

Appendix B. Mean Treatment and Control Group Characteristics, Rollout Sample By Quarter

	Jun – Aug, 2010		Sept – Nov, 2010		Dec – Feb, 2011		Mar – May, 2011	
	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control
Characteristics:								
Borough								
Manhattan	39.6%	41.4%	40.3%	41.8%	41.5%	40.8%	41.6%	40.6%
Bronx	9.4	9.8	10.2	9.4	10.0	9.0	9.6	9.3
Brooklyn	24.8	22.8	23.2	22.6	22.9	22.6	22.9	22.0
Queens	25.4	21.8	23.2	21.8	22.2	22.7	22.2	23.5
Staten Island	0.9	4.2	3.1	4.4	3.4	4.9	3.7	4.5
Chain	15.6%	12.3%	13.6%	12.1%	13.8%	9.9%	13.5%	7.8%
Workers	5.9	7.3	6.8	7.5	7.0	7.5	7.2	6.6
Building:								
Assessed Value	7,322,578	9,503,942	9,108,455	9,557,440	9,151,351	9,293,507	9,061,502	10,200,000
Building Type								
Office/Commercial	9.2%	7.2%	8.0%	6.9%	8.0%	5.6%	7.6%	4.7%
Retail/Commercial	34.7	34.7	35.5	34.1	34.4	34.8	34.5	34.0
Mixed Retail	41.0	43.3	42.7	43.5	43.8	42.8	43.7	41.8
Other Commercial	4.5	4.5	4.0	4.8	4.1	5.4	4.3	6.4
Residential	9.0	8.2	7.9	8.4	8.3	8.2	8.2	8.7
Government/Public	1.6	2.2	1.9	2.4	1.5	3.4	1.7	4.4
Joint Significance	F(12, 937) = 0.80 Prob > F = 0.6523		F(12, 937) = 1.08 Prob > F = 0.3779		F(12, 937) = 0.71 Prob > F = 0.7279		F(11, 937) = 1.39 Prob > F = 0.1738	
Initial Insp. Score	23.5	24.3	26.0	22.9	25.0	21.8	24.1	22.6
Final Insp. Score	18.0	20.1	20.9	19.2	20.5	18.3	19.9	18.8
Fines per Quarter	$341	$250	$329	$202	$285	$175	$262	$170
N	1,560	17,326	7,597	11,118	13,098	5,755	16,514	2,717
N with Building Code	1,082	12,122	5,364	7,817	9,333	4,000	11,742	1,858

Notes: Restaurants in the treatment group have their first graded inspection by the end of the fiscal quarter. Control group restaurants do not have a graded inspection until after the quarter ends. Each observation is a restaurant-quarter.

Appendix C. Regression Results, Impact on Inspection Scores and Fines by Quarter, Rollout Sample

VARIABLES	Initial Inspection	Final Inspection	Fines
Graded Inspection			
In Quarter 1:	1.897***	-2.137***	153.759***
	(0.677)	(0.379)	(21.950)
In Quarter 2:	0.810***	-2.377***	26.014
	(0.180)	(0.177)	(17.981)
In Quarter 3:	-0.693***	-3.616***	-147.830***
	(0.170)	(0.173)	(20.183)
In Quarter 4:	-2.071***	-5.561***	-275.079***
	(0.182)	(0.216)	(18.357)
Quarter Of Grading Policy			
2	0.074	1.243***	130.444***
	(0.052)	(0.072)	(10.915)
3	-0.133**	1.179***	251.674***
	(0.063)	(0.094)	(19.331)
4	0.974***	3.409***	387.343***
	(0.076)	(0.151)	(17.798)
Restaurant FE	Y	Y	Y
Constant	25.483***	19.736***	326.148***
	(0.059)	(0.067)	(6.021)
Observations	122,886	93,451	94,752
R-squared	0.849	0.764	0.553

Notes: Robust standard errors clustered by restaurant in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Impact estimates of graded inspections on initial inspection scores, final inspection scores, and fines by quarter. Graded Inspection captures if the most recent inspection occurs after the restaurant grade program begins in July 2010. Quarters Post * Graded is an interaction of the number of quarters after the grading policy is implemented and share of a quarter a restaurant has been treated. In Quarter 1, 2, 3, 4 are a vector of interactions between the quarter of observation and share of quarter with a graded inspection. Sample includes all restaurants continuously operating from the quarter before grading to five quarters following grading.

Appendix D. Regression Results, Impact on Revenues and Taxes by Quarter, Food and Beverage Rollout Sample

VARIABLES	Revenues	Sales Taxes
Graded Inspection		
In Quarter 1:	27,399.824	1,178.587
	(31,831.663)	(1,408.232)
In Quarter 2:	-400.594	93.322
	(5,882.097)	(237.801)
In Quarter 3:	-7,223.270	-269.696
	(10,282.449)	(428.125)
In Quarter 4:	-9,342.567	-341.486
	(11,618.780)	(516.547)
Quarter Of Grading Policy		
2	1,444.116	43.155
	(2,115.147)	(93.446)
3	-2,473.286	-134.734
	(7,043.024)	(305.405)
4	16,464.542*	673.659
	(9,861.280)	(440.550)
Group FE	Y	Y
Constant	212,202.897***	9,577.850***
	(1,213.733)	(51.758)
Observations	3800	3,800
Restaurant-Quarters	39,188	39,188
R-squared	0.981	0.983

Notes: Robust standard errors clustered by group in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Impact estimates of graded inspections on restaurant revenues and sales taxes by quarter, using aggregated group data. Graded Inspection captures if the most recent inspection occurs after the restaurant grade program begins in July 2010. Quarters Post * Graded is an interaction of the number of quarters after the grading policy is implemented and share of a group's restaurants with their most recent inspection occurring after the restaurant grade program begins. In Quarter 1, 2, 3, 4 are a vector of interactions between the quarter of observation and share of group that have had a graded inspection. Sample includes all restaurants continuously operating from the quarter before grading to five quarters following grading.