DRAFT - NOT FOR CIRCULATION OR QUOTATION

THE IMPACT OF RESTAURANT LETTER GRADES ON TAXES AND SALES: MICRO EVIDENCE FROM NEW YORK CITY

Amy Ellen Schwartz, Michah W. Rothbart, Thad Calabrese, Zachary Papper, Todor Mijanovich, Rachel Meltzer and Diana Silver

This paper presents difference-in-differences estimates of the impact of public restaurant grades on economic activity and public resources controlling for underlying food safety compliance, using evidence from New York City. We find that A grades reduce probability of closure and increase restaurant revenues, while increasing sales taxes remitted and decreasing fines relative to B grades. Conversely, C grades increase probability of closure and decrease restaurant revenues, while decreasing sales taxes remitted relative to B grades. These findings suggest that policymakers may be able to incorporate public information efforts into their regulatory approach with positive financial implications for the private and public sectors.

Keywords: Public Grades, Restaurant Revenues, Public Resources

JEL classification codes: D00, H22, H27

Amy Ellen Schwartz: Institute for Education and Social Policy, Steinhardt and Wagner Schools, New York University, New York, NY, USA and Maxwell School, Syracuse University, Syracuse, NY, USA <amy.schwartz@nyu.edu>

Michah W. Rothbart: Institute for Education and Social Policy and Wagner School PhD Student, New York University, New York, NY, USA <michah.rothbart@nyu.edu>

Thad Calabrese: Wagner School, New York University, New York, NY, USA <thad.calabrese@nyu.edu>

Zachary Papper: New York City Department of Finance, New York, NY, USA <PapperZ@finance.nyc.gov>

Todor Mijanovich: Steinhardt School, New York University, New York, NY, USA <tod.mijanovich@nyu.edu>

Rachel Meltzer: Milano School, The New School, New York, NY, USA <meltzerr@newschool.edu>

Diana Silver: Steinhardt School, New York University, New York, NY, USA <diana.silver@nyu.edu>

I. INTRODUCTION

In July 2010, the New York City Department of Health and Mental Hygiene (DOHMH) began requiring restaurants to post summary results of food safety inspections in the form of a letter grade (*A*, *B*, or *C*) in a conspicuous location near the restaurant's entrance. New York City (NYC) Mayor Michael R. Bloomberg touted the success of the letter grades, claiming they had beneficial health and economic effects. In 2012, he noted increases in total restaurant sales in NYC (by 9.3 percent or $800 million) in the first nine months of the grading program as compared to the year before, declines in reported cases of *Salmonella* and hospitalizations due to food borne illness, and improvements in compliance with food safety regulations.[1] However, this enthusiasm was not shared by the restaurant industry, which charged that the new public grading program hampered business. Andrew Rigie, executive vice president of NYC's chapter of the New York State Restaurant Association, contended "[i]f you define success as taxing small-business owners and making their lives miserable, then letter grades have been a complete success" (Saul, 2012).

Public restaurant letter grades have a clear intuitive appeal: grades provide easy to access information about restaurant food safety that allows consumers to make informed decisions about where to eat; to "vote with their feet," directing their business dollars to restaurants with high grades and perhaps lower risks of food-borne illness. Thus, an *A* grade could benefit a restaurant while a *C* could damage it.[2] The magnitude of the effect is unknown, depending, in part, on the distribution of letter grades and behavioral responses by consumers and restaurant owners. Restaurants may also increase compliance with health regulations in order to earn a

[1] City of New York, Office of the Press Secretary, 2012.
[2] A similar logic motivates letter grading of public schools and hospitals, and the perceived success of these efforts fuels the enthusiasm to spread grading to other areas such as subway stations and transportation services, street vendors, among others.

higher grade – a response clearly intended by program champions in NYC and elsewhere (Jin and Leslie, 2003; Simon et al., 2005; Wong et al., 2015). The extent to which this provision of information influences consumers to change their behavior, induces restaurants to improve adherence to health regulations and ultimately, affects sales, fines, taxes or restaurant closures are empirical questions which we address in this paper.

We draw from a wider literature on public grading and the value of information in framing this work. A key assumption of the restaurant letter grades law is that consumers will use this information, which is now available at the point of sale, when making consumption decisions. This may lead to decreases in sales for restaurants where there are other restaurant options that have earned better grades. Such assumptions have been studied in other contexts, including public education and public health. For example, many school districts grade public schools on their effectiveness (measured by improvements in test scores and other information), and they make these grades available to the public. There is some evidence that schools with low grades have short-term improvement in aggregate student achievement (Rockoff and Turner 2010; Winters and Cowen 2012). As another example, many cities require fast food restaurants to post the caloric content of menu items so that consumers may make more informed choices at the point of purchase. The evidence regarding the public's use of this information has been mixed and Elbel et al. (2009) found no impact of calorie labeling on the number of calories ordered at the point of purchase.

Here, we use detailed longitudinal data on taxes, fines and health inspections for all NYC restaurants over a 5-year period to gain insight into the impact of the grades on the economic activity of restaurants, and, by extension, on the City's tax and fine revenues. We explore one key feature of the program: the 'return' to getting a good grade to restaurants and the City. We

estimate the impact of grades on restaurant closure, fines, sales, and sales taxes. Is there evidence that grades provide incentive to improve? Are restaurant sales and closures affected by the grade posted in the window? And what is the impact on the level and mix of City revenues (i.e. total revenues and revenue sources)? Using a difference-in-differences model with fixed effects we estimate the impact of posting an *A* (versus *B* and versus *C*) on economic activity (revenues and closures) and payments to the City (fines and sales taxes), while controlling for underlying food safety compliance scores, a range of restaurant characteristics, and spatial and temporal fixed effects.[3] These results provide evidence on the extent to which grades provide restaurants financial incentives to improve their food safety compliance and the fiscal consequences for the public sector.

Importantly, we observe underlying inspection scores in each period, which are health inspector assessment of a restaurant's food safety compliance and are used to assign grades. Controlling for inspection score, either with a linear control variable or by focusing on restaurants near the grade threshold, allows for unbiased estimation of the impact of grades if assignment of grades near the grade threshold is "as good as random," which is likely due to the design and implementation of the program.[4]

Results suggest that, indeed, grades matter. Receiving an *A* grade – rather than a *B* – increases a restaurant's sales (approximately $78-$145 dollars per day) and sales taxes collected (between $4 and $6 per day), decreases the amount of fines assessed (more than $2,000 annually), and decreases the probability by 3-5 percentage points that a restaurant closes in the period after earning the grade. A *C* grade has the opposite effect – restaurants receiving a *C* are

[3] In this paper, "sales taxes" are sales taxes collected by restaurants that are owed to New York City.
[4] In particular, inspections are unannounced, randomly assigned within an inspection window, inspectors rotate restaurants, and inspectors are randomly assigned each inspection. We further assess whether the data hold up to the difference-in-differences assumptions (including continuity of observed variables and with a pre-period falsification test), and conclude, on net, that the assumptions hold.

more likely to close and their sales (and taxes) decrease compared to *B* restaurants. Together, our results indicate that grading mechanisms, as a means of making more available to the public otherwise obscured information, is effective at changing both consumers' and restaurant operators' behaviors. They also indicate that grading policies designed to improve well-being (i.e. health), may be able to bolster an effective, perhaps more palatable, source of revenue (i.e. sales taxes) in lieu of more punitive fine-based revenue streams.

The rest of the paper is organized as follows. We begin with a brief history of the restaurant grade policy in NYC. In section three, we review previous literature. Section four presents the data and measures. In the fifth section we discuss our empirical strategy, and in section six, the results. We then conclude with a summary of findings and implications.

II. BACKGROUND ON RESTAURANT GRADING IN NEW YORK CITY

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 that remain unaddressed during inspection. Prior to grading, all restaurants were inspected annually, and violations found were posted on DOHMH's website in a searchable database. Inspections occur on a regular basis, but inspectors are randomly assigned, and the precise timing of inspections is random 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. DOHMH also added each restaurant's grade to its website.

While restaurants were inspected and assigned scores prior to the introduction of letter grades, the reform made the health inspection information more accessible both by (1) moving

consumer information from an offsite database to the point of purchase, potentially increasing its salience, and (2) creating more easily interpretable discrete grades (*A*, *B*, and *C*). These reforms were intended to improve compliance with food safety regulations by increasing the stakes of enforcement and increasing the salience of information provided by inspection scores.

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.[5]

Additional points can be added to each violation to reflect the extent of the violation (on a scale of 1 to 5).[6] Points from violations are then aggregated to generate the final inspection score, with lower scores reflecting more hygienic conditions than higher scores.

DOHMH assigns *A* grades for scores of 13 and below, *B* grades for scores of 14—27, and *C* grades for scores of 28 or higher.[7] 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 yielding a provisional grade, which are comprised of initial *A* inspections and all re-inspections. In addition to re-inspections,

[5] New York City Department of Health and Mental Hygiene. (2012).
[6] See NYC Department of Health and Mental Hygiene (2010) for detailed explanation of the relationship between violation severity and violation points assessed.
[7] Restaurants can also be temporarily closed if they pose a large public safety risk, but we do not estimate the impacts of these temporary closures in this paper.

inspection scores (and, therefore, grades) can be lowered (improved) through an adjudication process, which in turn can reduce fines assessed as well.[8] During the period between a final inspection and the adjudication date, the restaurant can post a sign reading "*Grade Pending*" in lieu of their *B* or *C* grade. In addition to publicly posting grades, restaurants assigned an *A* at an initial inspection are visited only annually for food safety inspections. Those receiving a *B* level score are inspected twice per year, and those receiving a *C* level score are inspected every four months. A simplified diagram of the inspection process can be found in Figure 1.

<Insert Figure 1, about here>

As for fines, 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. Beginning January 19, 2011, those restaurants assigned an *A* grade at inspection are exempt from fines. As a result, restaurants with an *A* inspection do not incur any fines for much of the post-period. By design, this creates a sharp increase in fine levels for *B* graded inspections near the *A*-*B* grade threshold as compared to otherwise similar *A*s. Those assigned an *A* grade through adjudication, however, are still fined based on number and severity of violations.

III. LITERATURE REVIEW ON PUBLIC GRADING

[8] After two inspections (an initial and a re-inspection), restaurants have the right to due process and may challenge inspection violations at a third party tribunal. Silver et al., in preparation, estimate the impact of public grades and public grading on the adjudication process for NYC food safety inspections.

Despite the attention in the popular press and political arena, there is relatively little research examining the efficacy of public restaurant grades, even though they are a common tool in the United States and international settings. Perhaps most relevant, Ho (2012) found that, in NYC, prior scores predicted less than two percent of future grades, and interpreted this as program inconsistency. Notice, however, that the new restaurant public grading law was explicitly intended to encourage restaurants to take actions to improve (and change) their grades in future inspections. Therefore, the "inconsistency" observed by Ho may, instead, be interpreted as a sign of the success of the program. More recently, 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.

Two recent studies (Jin and Leslie, 2003; Simon et al., 2005) estimated the effects of the Los Angeles health inspection letter grade system, which began 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 found that posted grades improved restaurant inspection scores, that restaurant revenues increase in response to positive signals of hygiene quality, and that foodborne-disease hospitalizations decreased in Los Angeles County following the implementation of the public letter grade program. They also suggested that the improvements in health outcomes cannot be explained by consumption choices alone, but are also likely a result of restaurant hygiene improvements. Simon et al. (2005) use California hospitalization data to estimate the effect of grading on foodborne illness hospitalizations. Using OLS regressions, they

compare LA to the rest of California, finding a decrease in foodborne-illness hospitalizations that is sustained for at least three years.

While providing the best evidence to date that public restaurant grades influence economic activity and consumer choice, the difference-in-differences models used in Jin and Leslie (2003) omit underlying restaurant hygiene practice as measured by inspection scores. Consumers may observe differences between two restaurants of dissimilar food safety practice (such as rodents versus no rodents) and choose the one with better practices even in the absence of public grades. A simple difference-in-differences may not be appropriate if consumers use other signals to assess restaurant hygiene.

One way to address this form of endogeneity is by controlling for underlying food safety compliance scores that are used to assign grades. For example, one can restrict the sample to restaurants near and on either side of the grade assignment threshold and estimate the impact of the grade, which yields unbiased estimates of *A* and *C* grades near the grade thresholds if restaurants cannot manipulate grades except through improved food safety compliance (and, therefore, improved inspection scores). A regression discontinuity model would allow for unbiased estimation, but due to the distribution of inspection scores in NYC, we have pause using this approach. Instead, we assess the extent to which difference-in-differences estimates are robust to the inclusion of controls for inspection scores or restricting the sample to restaurants near grade cut points. No studies to date have examined the impact of grades on economic activity or restaurant viability while also controlling for restaurant food safety practice.

In addition, current studies do not consider public finance effects of grades, which are increasingly relevant during periods of fiscal stress. Previous work has not examined the impact of public grades on the level and mix of revenues paid to local municipal governments (i.e. total

public revenues and public revenue sources). In a time of increasing competition for public resources, understanding the potential effects of these public health initiatives on government revenues is critical and yet unexplored.

IV. DATA AND MEASURES

This study utilizes richly detailed, longitudinal, inspection and restaurant data from the DOHMH matched on Employer Identification Numbers (EINs) with longitudinal sales tax data from the New York City Department of Finance (DOF). In this section we first describe the data from DOHMH, and then the data from the DOF. Finally, we explain how the data is matched and aggregated in order to preserve privacy of restaurant sales information.

A. DOHMH Data

Our analytic sample for the impact on fines and closures includes the universe of final DOHMH food safety inspections from July 27, 2010 through February 28, 2013, spanning the two and a half years following the implementation of public grading.[9] We use data on restaurant characteristics and zip codes, inspection and adjudication dates, inspection scores, grades and fines. We include all inspections for which a restaurant receives a provisional grade. This sample includes 82,977 inspections of 29,742 restaurants in all; this will constitute the first dataset for our analysis.

We use inspection scores assessed at final inspections to capture food safety compliance. *Final inspection scores* are used to grade restaurants. On average, a restaurant has 3.2 final inspections from July 27, 2010 through February 28, 2013 (as shown in Table 1). Table 2

[9] In addition, we show some descriptive statistics from the "pre-period" two and half years before public grading to provide policy context.

provides summary statistics of inspection scores, which improve (i.e. decline) over time. During the period before public grading, average *final inspection scores* were 24.6. In the first five quarters after public grading, average *final inspection scores* improved to 18.2, driven by improved compliance during re-inspections. In the next five quarters, average *final inspection scores* further improved to 15.6, driven in part by improved compliance on *initial inspection scores*.[10] We find similar changes in mean inspection scores for the set of "continuously operating" restaurants, which are those that operate for two and half years before and two and half years after public grading.

<Insert Tables 1 and 2, about here>

Grades summarize *final inspection scores* into discrete measures of food safety compliance. As stated previously, *A* grades reflect scores of 13 and below, *B* grades reflect scores of 14—27, and *C* grades reflect scores of 28 or higher. *ITT Grades* are provisional grades based on *final inspection scores*. We refer to these as the "intention-to-treat" (*ITT Grade*), as grades may be contested and changed through adjudication. *ITT Grades* include *A, B,* and *C.*

Posted Grades reflect the grades consumers see at the point of sale (the "treatment"). *Posted Grades* include *A, B, C,* and *Grade Pending*, and reflect *adjudicated scores*, which can provide improvements on *final inspection scores* through a third party tribunal. The same thresholds are used for *Posted Grades* as *ITT Grades*. *Posted Grades* are always at least as good as *ITT Grades*, because *final inspection scores* can only improve or remain the same during an adjudication hearing.

While *Posted Grades* is arguably a better measure of what consumers actually observe,

[10] These changes are descriptive. Meltzer et al. (2015) estimate the impact of the public grading policy on inspection scores.

ITT Grades is our preferred measure of food safety compliance because *Posted Grades* can be manipulated through the adjudication process. For example, restaurants might hire tribunal representatives, increasing the probability of earning an *A*. Further, the incentive to do so may depend on the potential returns to a better grade. Thus, *ITT Grades* are a better and more conservative measure of food safety compliance, while *Posted Grades* are a better measure of the information readily available at the point of sale, but might be more easily manipulated.

The distributions of *ITT Grades* and *Posted Grades* by quarter are shown in Figures 2 and 3, respectively. The trends shown in Table 2 are reflected in Figures 2 and 3 as well; grades improve over time. For example, as shown in Figure 3, only about 65 percent of restaurants post *A* grades in the fifth quarter of public grading, but over 80 percent post *A* grades by the tenth quarter of public grading.

<Insert Figures 2 and 3, about here>

A restaurant is closed or "out-of-business" (*OOB*) if it is reported as not operating for three straight inspection attempts (on different days and at different times of day). The *OOB* date is then assigned as the first failed inspection attempt (that is, the first day for which the restaurant is observed not operating). For each final inspection, an indicator variable (*OOB*) takes a value of 1 if a restaurant goes out-of-business within 365 days of an inspection and 0 otherwise.[11] Notice that our estimate of the timing of the closure typically lags the true closure date – except in the unlikely event that the restaurant closed on the exact day of the first inspection attempt. As shown in Table 1, about 12 percent of restaurants go out of business (closes) each year.

[11] We test the sensitivity of our estimates to a wider time window (closures within 390 days) and find that the results are consistent with the findings reported in this paper. These results are available upon request of the authors.

Fines capture restaurant payments to the City resulting from inspection violations. *Fines* are final (post-adjudication) fines assessed for each restaurant inspection. *Fines* are adjusted using urban wage earners CPI to real 2013 dollars (United States Bureau of Labor Statistics). Figure 4 shows mean restaurant fines by quarter. Fines average about $500 per restaurant per quarter. While, on average, fines 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 quarter of 2013 ($353 in fines for the average restaurant).[12]

<Insert Figure 4, about here>

Other restaurant characteristics used as control variables include number of seats, number of employees, an indicator for chain restaurant (at least 15 locations nationwide), and a series of indicators for cuisine, service type, and venue type. Table 1 shows descriptive statistics of the restaurants in our sample. On average, a restaurant has 3.2 final inspections, employs 5.9 *workers* and has 29.5 *seats*. *Chains* comprise 11 percent of restaurants. DOHMH defines over 80 different *cuisine* types; the most common *cuisines* are American, Chinese, and Pizza. There are 13 *service* types; *service* type includes wait service only, wait service and counter service, takeout only, etc. *Venue* type summarizes a restaurant's setting and there are indicators for 26 different *venue* types including diner, arena-stadium concession stand, bar/pub/brewery (food

[12] Meltzer et al. (2015) find the initial increase in mean fines is driven by increased frequency of inspections, while fines per inspection have fallen dramatically over time. Fines per inspection two years after program implementation are lower than in any pre-period quarter.

served), night club, restaurant (with bar), and restaurant (no bar).[13] We use data on restaurant characteristics recorded at the last inspection only, so these do not vary over time.

B. DOF Data

Our sample for the impact on sales and sales taxes includes graded food and beverage establishments with any sales from July 27, 2010 through November 31, 2012. Some DOHMH graded establishments are not primarily restaurants – and a large share of sales can derive from non-restaurant activity.[14] Including entities that primarily earn revenue through alternative streams, such as hotels, increases statistical noise. Therefore, we use a subsample of primarily food and beverage providers as identified by their North American Industry Classification System (NAICS) code for the analyses in this paper.[15][16]

DOF data include reported quarterly sales and sales tax liabilities (hereafter, *Sales* and *Sales Taxes*) from graded establishments.[17] Figure 5 shows mean *Sales* by quarter. On average, *Sales* figures are quite high ($175,000 to close to $200,000 a quarter) and mask a fairly high level of heterogeneity across restaurants. Mean *Sales* increase slightly after grading, suggesting that food and beverage entities have gotten a bump in *Sales* following the grading program implementation, but this could also be caused by other macro-level trends in addition to public grades.

[13] A full list of cuisine, service, and venue types is available upon request.
[14] For example, hotels usually operate restaurants that are graded by DOHMH. Sales tax returns, however, aggregate all sales for the hotel and do not separate them by individual business activities.
[15] We use all NAICS codes beginning with 722, as well as 445299, 445291, and 445120 to flag primary food and beverage establishments. We use the same set of code identifiers used and recommended by the NYC DOF.
[16] We also estimate the impact on sales using a sample of all graded entities as a robustness check, finding consistent results, which are available upon request of the authors.
[17] NYC restaurants are required to collect sales tax on food and beverage sales at a rate of 8.875 percent of gross sales - 4.875 percent for New York State and 4.0 percent 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.

Sales Taxes mirror *Sales* revenue. *Sales Taxes* during this period range from about $8,000 to nearly $9,000 a quarter, with a slight (statistically significant) rise over time. Like mean *Sales*, mean *Sales Taxes* masks large levels of heterogeneity across restaurants.

<Insert Figure 5, about here>

Building Class is a vector of indicator variables constructed from The Real Property Assessment Database (RPAD) data from DOF. *Building Class* is used as a set of control variables for locational or use characteristics, which may be associated with revenue generation and/or higher/lower grades. Six *Building Class* types include office commercial, retail commercial, mixed use retail, other commercial, residential, and government/public.[18] We match restaurant BBL to *Building Class* for each year in our sample period.[19]

C. Matching Tax and Inspection Data

As noted previously, we match DOHMH and DOF data on Employer Identification Numbers (EINs). To do so, we match quarterly DOF data to date specific DOHMH data. We aggregate inspection data by restaurant (including grades and inspection scores) to quarters and match 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 *ITT Grades* and *Posted Grades*). The matched sample includes 15,899 restaurants or bars observed over the nine quarters post-policy-implementation.

[18] We use the RPAD variable AV-BLDGCL to construct our building class variable. See Appendix A for more information.

[19] The sample restaurants mostly operate in mixed use retail buildings and retail commercial buildings (44 percent and 34 percent, respectively), with less than 10 percent of the sample operating in each of the office commercial, other commercial, residential, and government/public buildings.

D. Aggregating / Privacy

Confidentiality rules prohibit DOF from providing establishment-level sales and tax data to outside researchers. We received data aggregated in 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."[20] To address attrition and entry, DOF first stratified the sample based upon quarters of operation and then assigned groups within these.[21]

The result is our second data set, which is organized by group-quarter and includes sales and tax information and summary inspection results. The data provides variables 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 *ITT 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 9,182 observations in 1,538 groups (which are comprised of 15,899 restaurants or bars during the study period).[22] As a robustness check, we use the full set of matched establishments (including those not primarily food and beverage purveyors), randomly assign all establishments to a new set of groups, and find consistent results.[23]

[20] A small number of groups have 11 rather than 10 restaurants in order to include all restaurants.
[21] 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 four 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.
[22] The policy is implemented in the middle of the 2nd sales tax quarter in 2011. Our analytic sample includes data observed from the 3rd quarter of 2011 to the 3rd quarter of 2013.
[23] These estimates are available upon request of the authors.

V. EMPIRICAL STRATEGY

We first use difference-in-differences models comparing fines, closures, sales, and sales taxes across restaurants with *A*, *B*, and *C* grades. We then add controls for underlying inspection scores, controlling for *final inspection scores* and, alternatively, restricting the sample to restaurants near the grade cut points. Finally, we add restaurant fixed effects to control for unobserved differences in restaurant quality that might be correlated with food safety compliance. Our results provide unbiased estimates of the impact of grades if assignment of grades near the grade threshold is "as good as random." The likelihood of restaurant manipulation is minimized by program design; inspections are unannounced, randomly assigned within an inspection window, inspectors rotate restaurants, and inspectors are randomly assigned each inspection. The program is purposefully organized so that *final inspection scores* accurately reflect food safety environments rather than gaming. ___put in here?????___

Our estimates are unbiased if *final inspection scores* reflect the food safety environment or random noise (such as random inspector assigned) rather than some managerial skill that leads to better performance near grade thresholds, which are also associated with economic activity and City revenues.

One aspect of the program design might be subject to managerial skill and gaming: the adjudication process. Restaurants may choose to challenge grades at a third-party tribunal and winning such cases may reflect managerial skill or resources in addition to food safety practice. We, therefore, estimate the impact of grades in two ways. First, we estimate impacts using an intent-to-treat (ITT) framework, using provisional grades assigned through the inspection

process (*ITT Grades*).[24] The coefficients from the ITT analyses provide estimates of the impact of inspections yielding *A* or *C* grades, ignoring the adjudication process. If consumers only respond to publically posted grades, then the ITT estimates are conservative estimates of the impact of posted grades (attenuated towards zero), because some *ITT Grades* improve through the adjudication process.

Second, we control for inspection scores, which reflect the inspector's assessment of food safety compliance, and estimate the effect of *Posted Grades*, which plausibly may reflect ability to challenge grades in court. F*inal inspection scores* do not perfectly predict *Posted Grades*, but do affect the probability of restaurants posting *A*, *B*, *C*, or *Grade Pending*. Compared to *ITT Grades*, *Posted Grades* is arguably a better measure of information readily available to consumers at the point of sale, but also might be somewhat more easily manipulated. Estimates from the *Posted Grades* analyses are unbiased if restaurants cannot manipulate grades through the adjudication process and changes in grades reflect corrections from inspector error, which are randomly distributed across restaurants. Coefficients from these analyses provide estimates of the impact of grades that are required to be conspicuously posted in restaurant windows, though estimates from regressions using *Posted Grades* – as a result of adjudication hearings – might overstate the impact of grades if improving management quality increases both the probability of winning an adjudication hearing and revenues earned.

___Fill in throughout__???

Grades are assigned with certainty based on underlying inspection scores, so this study at first appears like an ideal place for a regression discontinuity design. Underlying inspection

[24] For fines and closure ITT estimates (establishment level), we use *ITT Grades* earned during final inspection. For sales and tax ITT estimates (aggregate data), we use *ITT Grades* at the beginning of the quarter because *ITT Grades* are provisional and best reflect which grades are posted during the upcoming quarter due to the restaurant food safety inspection cycle. For example, a restaurant can post "Grade Pending" after earning an *ITT C*, while the *C* is often not posted until late in the quarter (if at all) due to adjudication timing.

scores are not, however, smoothly distributed through the grade cut points, one of the key tests for the appropriateness of a regression discontinuity (see, for example, Figure 6 Panel A). Moreover, breaks in the score distribution near grade cut points raises concerns that some managers of restaurants are able to manipulate their scores, earning better grades when near an assignment threshold. Importantly, restaurant manipulation would have to occur at the point of inspection, because inspection timing is randomly assigned within a window, and would have to be successful with a large enough share of the randomly assigned inspectors to justify the effort.

For restaurant manipulation to bias estimates, manipulation ability would have to be correlated with a restaurant's ability to earn more revenue (avoid closure, accrue fines, etc.). Further, in estimates with restaurant fixed effects, ability to manipulate would have to increase over time as a restaurant's ability to earn revenue (avoid closure, etc.) increases. That is, even if an RDD strategy cannot be employed, difference-in-differences estimates are unbiased if grade assignment is exogenous. For that reason, in this paper we focus on final inspection scores in part because it is more likely restaurants can manipulate scores during the adjudication process, when they have opportunities to prepare arguments, hire experts to represent them, and timing is not random. That is, for example, scores are unsmooth following adjudication because restaurants are more likely to improve scores through the adjudication process if they fail to achieve an A grade (Silver, Rothbart, Bae, Schwartz and Mijanovich, in preparation).

In fact, final inspection scores are likely to be unsmooth without any restaurant manipulation due to program design. Only restaurants that fail to achieve an A grade on initial inspection are reinspected, thus leading to an uneven score distribution near the A grade threshold. This sort of mechanical effect on the score distribution would not bias estimates because restaurants near the grade threshold still pull from the same score distribution, just non-

A restaurants pull twice. Moreover, scores are a sum of violation points, which accrue in units other than one, which leads to an integer problem. Taken in sum, programmatic details explain a large portion of the lack of smoothness in the score distribution, but not all of it. Panel B of Figure 6 shows that scores are more smoothly distributed for initial inspections than after adjudication. Finally, Panel C of Figure 6 shows that the integer problem exists two years before grading, but scores are smoother than the score distributions after public grading.

Instead of leaning on the RDD, we focus on identifying exogenous differences-in-differences estimates. That is, we control for restaurant characteristics and fixed effects, along with underlying inspection scores, to give credibility to the assumption that grade assignment is conditionally random near the thresholds, even though probability of a better grade (*A*) is higher than a lower grade (*B*). This is possible if some inspectors only assign A grades for marginal restaurants, but because inspectors are randomly assigned to restaurants each inspection, this would not lead to underlying bias in the impact estimates. Unlike the previous conditions, this is not directly testable because inspector IDs were not made available to the research team due to privacy concerns. It can, however, be assessed indirectly by examining if restaurants near a grade threshold that end up with a better grade are better equipped to earn that grade on average than those who end up with a worse grade. That is, testing grade on next initial inspection for the subset of restaurants included in our Wald estimates' sample. Here, we find that restaurants that barely get an *A* are no more likely to get an inspection score that leads to an *A* in the following inspection cycle than those that just barely miss an *A* and instead get an inspection score that would lead to a *B*. Column 1 (and 2 and 3) of _____ shows the probability that a restaurant one point (and two and three points) within the *A*-range are no more or less likely to be so again than those restaurants one point (and two and three points) away from the *A*-range.

1. With restaurant/group fixed effects

2. With controls for the slope effect of inspection scores

3. Wald estimates - restricting the sample to inspections near the inspection cut points.

A. Fines and Closures – Establishment Level

The starting point for our empirical work is a regression linking restaurant outcomes to grades and other variables:

$$(1)\ y_{it} = \beta_0 + \beta_1 A_{it} + \beta_2 C_{it} + \beta_3 \text{Score}_{it} + \mathbf{X'_i\beta_4} + \gamma_i + \delta_t + \varepsilon_{it}$$

where y is a restaurant-specific outcome (i.e. *OOB*, *fines*); *A* (*C*) is an indicator variable that takes a value of one if restaurant *i* is awarded an *A* (*C*) grade for an inspection in period t; *Score* is the *final inspection score* for restaurant *i* at inspection *t*; *X* is a vector of restaurant characteristics including cuisine, service, and venue type; γ and δ are zip code and quarter-by-year fixed effects, respectively; and ε is an error term with the usual properties. The omitted category in these and all subsequent analyses is a *B* grade. In alternative specifications, we add restaurant fixed effects, μ_i, dropping γ_i and X_i. Model (1) yields estimates of the impact of *A* (*C*) grades identified by the restaurant fixed effects model and controls for underlying food safety compliance score. β_1 and β_2 will be unbiased estimates of the impact of *A* and *C* grades, respectively, if restaurants cannot manipulate their ITT grade near the thresholds and there is a linear relationship between *Score* and y_{it}. As stated previously, the first assumption is plausible because DOHMH randomly assigns inspectors to restaurants each cycle, the timing of inspections are random within two month windows, and we use *ITT Grades* to address any

possible manipulation through the adjudication process. We find the second assumption plausible because the correlation between *Score* and y_{it} is nearly zero conditional on grade. Moreover, we relax the linearity assumption, estimating the impact on restaurants with inspection scores near the *A* (*C*) cutoff, using a Wald estimator, and restricting the sample to inspection scores just above or below the grade assignment thresholds. While we acknowledge the lack of smoothness in the score distribution, we ultimately derive the local linear regression discontinuity estimates using the optimal bandwidth, which minimizes MSE, as a robustness check (Imbens and Kalyanaraman, 2009).

We further test the robustness of our closure results using an alternative measure of restaurant closure. Instead of DOHMH direct measures of closure (*OOB*), we estimate the impact of *ITT Grades* on the probability that a restaurant no longer receives revenue in subsequent quarters (Z*ero Revenue*). This robustness check is estimated using the aggregate (DOHMH-DOF merged) data.

A last falsification test for the impact of grades on closures uses *final inspection score* and *OOB* dates from the period before public grading. We use *Synthetic ITT Grades* based on *final inspection scores* and estimate the model above. That is, what if grades were assigned using the same grading formula in the period before public grades and this information were kept private?

B. Sales and Sales Taxes – Aggregate Data

We estimate the impact of *A* (*C*) grades on sales and taxes using model (2) below:

$$(2)\ y_{gq} = \tau_1 A_{gq} + \tau_2 C_{gq} + \tau_3 GP^B_{gq} + \tau_4 GP^C_{gq} + \tau_5 Score_{gq} + \boldsymbol{X'_{gq}}\tau_6 + \gamma_g + \delta_q + \varepsilon_{gq}$$

where y_{gq} are the group's average daily restaurant sales or sales taxes in quarter q;[25] A_{gq} (C_{gq}) is the mean share of days in quarter q restaurants in group g hold an A (C) grade; GP_{gq}^{B} (GP_{gq}^{C}) is the mean share of days in quarter q the restaurants in group g have a B (C) grade, but are allowed to post "Grade Pending"; $Score_{gq}$ is a group's average inspection score over the course of the quarter weighted by restaurant-days; X_{gq} is a vector of mean restaurant characteristics and building class; and γ_g and δ_q are group and quarter-by-year fixed effects, respectively. Model (2) yields estimates of the impact of A (C) grades on sales. τ_1 and τ_2 are estimates of the impact of A or C grades on a restaurant's daily sales and are unbiased if restaurant i's average grade in group g in quarter q only affects restaurant i's sales and not the sales of other restaurants in group g.[26] We find this assumption plausible due to random assignment of restaurants to groups and the inclusion of the quarter-by-year fixed effects, δ_q. While point estimates are estimates of the mean impact on restaurants, the standard errors are larger than if we observed individual restaurant sales.

Estimates of the impact using *ITT Grades* assigned to restaurants based on *final inspection scores* (regardless of whether the grade is "pending" or if the restaurant wins an adjudication hearing) are estimated using the model below:

$$(3)\ y_{gq} = \tau_1 ITTA_{gq} + \tau_2 ITTC_{gq} + \tau_3 ITTScore_{gq} + \tau_4 X_{gq} + \delta_q + \varepsilon_{iq}$$

[25] As a robustness check, we estimate the impact of grades on percent changes in average daily sales and sales taxes using average of log(sales) and log (sales taxes) as the outcome variables, respectively. Log model estimates are qualitatively similar to the results shown in this paper and are available upon request of the authors.

[26] Note that if a restaurant in group g earns an A for an entire quarter the value of A_{gq} will be 0.1 higher than if that restaurant earns a B for the entire quarter. The impact on sales would be divided evenly among the whole group, but should be inflated by a magnitude of 10 if own restaurant grades only affect own restaurant sales. Thus, impact estimates are first divided by 10 to reflect the impact of one restaurant's grade change on mean grades, but then multiplied by 10 to reflect that impact on group sales are a result of only one restaurant's grade change.

Where y_{iq} is the restaurant's mean daily sales or sales taxes in quarter q and *ITT A*, *ITT C*, and *ITTScore* denote restaurant intention-to-treat grades and score at the beginning of quarter q.[27] τ_1 and τ_2 are interpreted as the ITT impact of an *A* or *C* on daily sales. As noted above, we expect these coefficients to be of smaller magnitude than for the *Posted Grades* because only the posted grades are salient (inspection scores are online, but *Grade Pending*, *B* or *C* may be in the window). τ_1 and τ_2 are interpreted as the effect of earning an *A* or *C* at inspection on a restaurant's daily sales, controlling for a restaurant's most recent *final inspection score* as of the beginning of the quarter. Estimates are unbiased if restaurant *i*'s average grade in group *g* in quarter *q* only affects restaurant *i*'s sales and not the sales of other restaurants in group *g* and if restaurants cannot manipulate *ITT Grades*.

VI. RESULTS

A. The Impact on Closures

Table 3 shows difference-in-differences estimates of the impact of *A* and *C* grades on restaurant closure (*OOB)*. As shown in column 1, *A* graded restaurants are 4.2 percentage points less likely to close than those earning a *B*. The results from models that include a linear control for scores are shown in column 2 and the effect of *A* grades is estimated to be a 2.6 percentage point decline in the likelihood of closure as compared to a *B*, still a significant and meaningful reduction. Column 3, including controls for restaurant characteristics and zip code fixed effects, yields statistically and substantively similar impact estimates. Results are also robust to including

[27] As noted previously, we use *ITT Grades* at the beginning of the quarter because these best reflect which grades are posted throughout the following quarter due to the restaurant food safety inspection cycle. Beginning-of-quarter *ITT Grades* are predictive of *Posted Grades* later in the quarter and best capture the impact of grades on sales and taxes. We present results from two other model specifications measuring inspection grades by daily average and at the end of the quarter in Appendix E.

restaurant fixed effects (column 4). Thus, *A* restaurants are less likely (by about 2.6 percentage points) to close than *B* restaurants.

<Insert Table 3, about here>

Does a *C* grade hurt? As shown in Table 3 (column 1), *C* graded restaurants are 4.9 percentage points more likely to close than *B*s. Our preferred model controls for inspection score and restaurant fixed effects; in column 4, a *C* is 2.1 percentage points more likely to close than a *B*, a significant increase relative to the closure rate overall of 12 percent.

Next, we relax the linearity assumption, estimating the impact on restaurants with inspection scores near the grade thresholds. Table 4 shows Wald estimates of the impact of *A* and *C* grades on closure. Our preferred model (shown in columns 1 and 4) limits the sample to scores one point above and below the cut point. The estimate in column 1 shows that restaurants earning *A* grades are 4.9 percentage points less likely to close within a year than those earning *B* grades. The estimate in column 4 shows that restaurants earning *C* grades are 4.2 percentage points more likely to result in closure than those earning *B* grades (although only marginally significant). The point estimates are not sensitive to increasing the bandwidth to two points or to inclusion of restaurant characteristics and zip code fixed effects.[28] As a robustness check, estimates using a local linear regression discontinuity estimator yields similar results (Appendix B).[29]

[28] As an additional sensitivity analysis, we estimate the impact using a slightly wider window of 390 days following final inspection. Using our preferred model specifications (columns 1 and 4), we find that *A* inspections are 3.2 percentage points more likely to lead to closure and *C* inspections are 1.6 percentage points more likely to lead to closure, respectively. Both point estimates are not significantly different from the coefficients reported in this paper.
[29] Local linear estimates use an optimal bandwidth, which minimizes MSE, as in Imbens and Kalyanaraman (2009). The optimal bandwidth for determining the effect of an A inspection is 1.293. The optimal bandwidth for determining the effect of a C inspection is 1.939. These are not meaningful in the context of inspection scores, so we use one and two points in the Wald regressions.

<Insert Table 4, about here>

To test the robustness of our findings to an alternative measure of closure, we estimate the impact using *Zero Revenue* as an indicator of closure. Recall that, by construction, if any restaurants in a group cease to have sales revenue in a quarter, then all restaurants cease earning revenue in the same quarter.[30] Using group *ITT Grades* we estimate the impact of grades on closure (*Zero Revenue*), finding estimates that are qualitatively similar to our inspection-level estimates. Using revenues to reflect restaurant open/closed status and our preferred models, we find an *ITT A* is 3.6 percentage points less likely to lead to closure and an *ITT C* is 6.6 percentage points more likely to lead to closure than an *ITT B*.[31]

Finally, a falsification test estimates the impact of *Synthetic ITT Grades* in the pre-grading period on closure (*OOB)* in the period before grading. Results in Appendix C and D show the impact of *Synthetic ITT Grades* is smaller in the pre-period, and is insignificant at the 95 percent level for all but one model (column 4 of Appendix C, showing estimated impacts of *Synthetic C Grades*), suggesting that the impact of grades are a result of the grades themselves rather than other confounders (such as observed differences in food safety practices).

B. The Effect on Fines

As shown in Table 5 column 1, *A* grades yield fines $518.46 lower than otherwise similar restaurants that earned a *B*. Results are insensitive to widening the bandwidth (column 2). As shown in column 3, *C*s do not yield fines appreciably different than similar *B* restaurants. While the results change in terms of statistical significance in column 4, the difference in fines using a

[30] As outlined previously, restaurants are randomly assigned to groups of 10, stratified on quarters of operation.
[31] Regression tables are available upon request of the authors.

two point bandwidth is equivalent to the *de facto* observed price of two additional inspection points (that is, the linear relationship between points and fines is $40 per point on average), suggesting that this may reflect changes in inspection scores rather than effects from inspection grades. Estimating using local linear regression yields similar results (Appendix B).[32] In sum, *A* grades reduce fines as compared to *B*s, but no similar effect is observed between *B* and *C* grades.

<Insert Table 5, about here>

C. The Effects on Sales and Sales Taxes

Table 6 shows estimates of the impact of grades on sales revenue (in 2013 dollars).[33] The simple difference-in-differences estimates in column 1 show that *A* grades increase and *C* grades decrease sales (the impact of *C* grades is statistically insignificant, but qualitatively meaningful, in this model). If grades only affect own restaurant sales within the randomly assigned group, then an *A* is associated with a $145 increase in mean daily sales as compared to a *B*. For a *C*, average daily sales are estimated to decrease by $104, but this result is not statistically significant in this model specification. The estimated effect of days in which a restaurant is allowed to post *Grade Pending* is insignificant as compared to the period in which a restaurant must post a *B*. This could be due to the fact that consumers perceive *Grade Pending* as equivalent to a *B* grade, or do not understand it consistently enough to convey any systematic consumption patterns. Further, during the period that restaurants may post *Grading Pending*, there is no significant difference in sales between restaurants with *B*s and those with *C*s. We cannot say for certain that

[32] Local linear estimates use an optimal bandwidth, which minimizes MSE, as in Imbens and Kalyanaraman (2009). The optimal bandwidth for determining the effect of an A inspection is 3.774. The optimal bandwidth for determining the effect of an C inspection is 10.35.
[33] Results from models that estimate the impact of *Posted Grades* at the beginning or the end of the quarter can be found in Appendix E.

the sales impact is the same for *B* and *C* grade during the grade pending period, but this does suggest that the effect of a *C* grade as compared to a *B* grade is obscured.

<Insert Table 6, about here>

Column 2 of Table 6 shows difference-in-differences estimates of the impact of *A* and *C* grades using *ITT Grades* rather than *Posted Grades*, showing estimates of the impact of a restaurant's most recent inspection-assigned grade on sales. We find that *A*s increase restaurant sales by about $78 a day and that *C*s decrease restaurant sales by about $122 per day. These estimates are similar to the estimated impacts of *Posted Grades*, but are more precisely estimated due to increased power (due to increased reference group size). While marginal restaurants may be able to improve grades through the adjudication process, *ITT Grades* have an impact on quarterly sales on average.

Columns 3 and 4 of Table 6 show estimates of the impact of grades on sales, controlling for the linear effect of inspection score and group fixed effects. Inspection scores are uncorrelated with restaurant sales once controlling for grades. In other words, increased sales from improved food safety practice are driven by improved grades. The independent effect of *Posted Grades* on sales is quite large (column 3). Comparing restaurants with very similar food safety scores, we estimate that those with *A* grades earn $123 more in sales a day than *B*s. Conversely, those with *C*s earn about $113 less than *B*s, but this is insignificant. Our estimates once controlling for *final inspection score* are similar to those presented above, which suggests that consumers mostly use the posted grade information when making consumption decisions and that the local average treatment effect of *A* and *C* grades on sales is similar throughout the inspection score distribution. Again, the impact of *Grade Pending* on sales are insignificant as

compared to the period in which a restaurant must post a *B* grade for restaurants receiving either provisional *B* or *C* grades, suggesting the impact of a *C* is muted during the grade pending period.

Column 4 of Table 6 shows ITT estimates of the impact of *A* and *C* grades assigned during inspection. We find that *A*s increase restaurant sales by about $83 a day and that *C*s decrease restaurant sales by about $144 per day.[34] Like the previously presented results, these coefficients are similar to the *Posted Grade* estimates, but are more precisely estimated. Moreover, these estimates are not exposed to the potential bias introduced via the adjudication process. Once again, the results suggest that grades are the primary source of food safety information for consumers, rather than underlying inspection scores.

Table 7 shows the sales tax implications of restaurant grades. Simple difference-in-differences estimates shown in column 1 reveal that restaurants with *A*s remit about $6 a day more in taxes to the City than those with *B*s. Similarly, restaurants with *C*s remit about $5 less a day in taxes than those with *B*s, though this is statistically insignificant. These are small effects each day, but imply that the City should expect approximately $2,000 per restaurant more in taxes annually for restaurants with *A* grades than otherwise similar restaurants with *B*s. There is no statistical or substantively meaningful difference in the sales taxes remitted between *B* and *C* graded restaurants during the grade pending period.

<Insert Table 7, about here>

[34] We present results from models using *ITT Grades* at the beginning of the quarter because these best reflect a restaurant's grades over the course of the quarter. Results from the other two model specifications measuring *ITT Grade* by daily average and at the end of the quarter are shown in Appendix F.

Column 2 of Table 7 shows ITT estimates of the impact of grades on sales taxes, which are of similar magnitude and direction to the *Posted Grades* estimates. Beginning a quarter with an *ITT A* leads to about $4 more in sales taxes remitted each day on average than beginning the quarter with an *ITT B*. Conversely, beginning a quarter with an *ITT C* leads to remitting about $6 less in sales taxes each day in the following quarter than a *B*, on average.[35]

VII. DISCUSSION AND CONCLUSIONS

This paper offers compelling evidence that public restaurant grades have important financial implications for both the private and public sector. We contribute to the growing literature evaluating the consequences of public grades, offering estimates of the impact on economic activity – controlling for underlying food safety compliance – and estimates of the fiscal impact on the public sector. Using data on restaurant closures, fines, sales, and sales taxes in NYC, we find that restaurants with better grades are less likely to close, receive fewer fines, have greater sales, and remit more taxes than they would have with worse grades. These impacts are large and are robust to controls for a variety of restaurant characteristics, restaurant fixed effects, and underlying inspection scores. Improving food hygiene compliance helps restaurants who move from a *B* to an *A* grade (or a *C* to a *B* grade) by increasing revenues and decreasing probability of closure. Further, moving from a *B* to an *A* increases sales taxes remitted to the City, but reduces fines, implying that public grades have important implications for NYC government revenues.

We find that the effects on economic activity are both statistically significant and substantively meaningful. The impact of earning an *A* rather than a *B* is a 3-5 percentage point decrease in probability of restaurant closure within a year. Conversely, the impact of earning a *C*

[35] The alternative set of RDD estimates of impact on sales taxes are shown in Appendix G.

rather than a *B* is a 2-5 percentage point increase in probability of restaurant closure. This is compared to the mean closure rate of 12 percent. Similarly, sales at restaurants earning an *A* rise by $80-120 a day as compared to a *B* grade (about $30,000 to 40,000 a year). In contrast, *C* graded restaurant sales decline by about $110-140 a day as compared to a *B* (about $40,000 to $50,000 a year). In addition, restaurant grades (both *Posted Grades* and *ITT Grades*) improve substantially during the observed period, suggesting that restaurants respond to the economic incentives. Combining our estimates of the impact of grades and the observed change in the citywide grade distribution, we can predict that mean restaurant sales would rise in the period following grading by as much as $7,000 to $10,000 annually if restaurant sales are not zero-sum.[36] Based on this upper bound estimate, the change in grade distribution may explain as much as 30 percent of the observed increase in sales since the start of restaurant grading.

Together, our results indicate that grading mechanisms, as a means of making obscured information more salient, are effective at changing both consumers' and restaurant operators' behaviors. First, patrons appear to migrate towards the establishments with better grades, as evinced by the grade-induced increase in sales revenues. Second, restaurants appear to improve their food safety compliance (as indicated by the decrease in non-compliance fines and the increase in the number of *A* restaurants), presumably in pursuit of maintaining and/or increasing their patronage. There are, however, trade-offs: restaurants that do not (or cannot) improve their food safety compliance and achieve higher grades suffer in terms of revenues and the likelihood

[36]A 10 percentage point increase in the *A* (*C*) share, for example, is associated with a $3,000-$4,400 increase ($4,000-$5,000 decrease) in sales. Even after the initial rollout period, the share of *ITT A* grades rises by about 20 percentage points annually (while the share of *ITT C* grades declines by about two percentage points annually).

of survival.[37] The upside, presumably, is the welfare (and fiscal) gains from forgone health consequences due to poor food safety compliance in those establishments.

In addition, we find statistically significant and substantive effects on public revenues. While earning an *A* rather than a *B* leads to a large decrease in fines assessed on the restaurant, there is little evidence of such an effect from earning a *B* over a *C* (by program design). The impact of *A* grades on fine revenues (i.e. a reduction), however, is offset by the impact of *A* grades on sales taxes. *A* grades increase sales taxes as compared to *B* grades by a magnitude commensurate with the changes in sales. Thus, *A* grades increase sales taxes by $4-$6 a day (about $1,500 to $2,000 a year) as compared to a *B* grades. Similarly, *C* grades decrease sales taxes by $5-$6 a day (about $1,800 to $2,000 a year). Again combining our estimates of the impact of grades and the observed change in the citywide grade distribution, we can predict that mean restaurant sales taxes would rise in the period following grading by as much as $350 to $500 annually if restaurant sales are not zero-sum.[38] Conversely, observed mean fines per operating restaurant fall between $360 and $650 annually after initial program rollout (as shown in Figure 4), largely resulting from the change in citywide grade distribution (more *A* grades) and improved food safety compliance.

These results show that the composition of NYC's revenue sources change as a result of the policy. In particular, restaurants earning *A* grades are more likely to stay open, pay fewer fines, and remit more sales taxes (because of increased sales) than restaurants earning *B*s. As a result, the City is likely to become increasingly reliant on these establishments' for long-term revenues and decreasingly reliant on restaurants with worse food safety practices. This also

[37] This possible "zero sum" outcome is documented in Meltzer et. al. (2015), in which the authors find less robust sales revenue gains for the City overall.

[38] A 10 percentage point increase in the *A* grade (*C* grade) share, for example, is associated with a $150-$220 increase ($180-$220 decrease) in sales. Share of *ITT A* (*ITT C*) grades rises (falls) by about 20 percentage points (2 percentage points) annually after initial rollout.

affects the long-term viability of food safety fine revenue, which has been on steady decline since its peak four quarters into the grading policy. As more restaurants earn *A* grades, fewer pay fines, and the City will become more reliant on the sales tax remittances relative to fine revenues. This is of particular importance in NYC, because fines from food safety inspections go directly to NYC while sales taxes first are remitted to New York State, then NYC, and only then allocated to the different branches of the City's municipal government. These present important fiscal concerns for NYC to consider as it continues to use public restaurant grades as a tool to regulate food safety practice. More broadly, these findings suggest that municipal governments weighing the efficacy of restaurant grades should consider the impact on both total public revenues and the sources of public revenues, rather than just the potential health gains. In addition to improving food safety compliance, our results suggest that grading policies may also raise municipal revenues via sales taxes and may substitute a more palatable source of public revenue (i.e. sales taxes) for a more punitive one (i.e. fines).

In sum, we find that integrating consumers into the regulatory framework for food safety compliance has provided new and substantial incentives for restaurants to improve their practices. Moreover, restaurants have responded quickly to these incentives with improved food safety compliance, as would be predicted by the size of the impact on sales revenues relative to the size of fines. In arenas where standards of practice, like food safety, are commonly established, policymakers may be able to incorporate public information efforts into their regulatory approach with positive financial implications for both consumers and service providers as well as for the public sector.

ACKNOWLEDGMENTS AND DISCLAIMERS

The results and opinions shown are entirely of the authors, but this research would not have been possible without the contributions of the NYC Department of Health and Mental Hygiene, the NYC Department of Finance, and the Robert Wood Johnson Foundation. Special thanks are due to Meryle Weinstein, Dan Kass, Wendy McKelvey, J. Bryan Jacobson, Melissa Wong, Corinne Schiff, Karen Schlain, and Steven Robinson.

DISCLOSURES

The authors have no financial arrangements that might give rise to conflicts of interest with respect to the research reported in this paper.

REFE**RENCES:**

City of New York, Office of the Press Secretary, 2012. *Mayor Bloomberg, Deputy Mayor Gibbs and Health Commissioner Farley Announce Decrease in Foodborne Illness and Increase in Restaurant Revenue Since Letter Grading Began.* [Press release]. Retrieved from http://www.nyc.gov/portal/site/nycgov/menuitem.c0935b9a57bb4ef3daf2f1c701c789a0/index.jsp?pageID=mayor_press_release&catID=1194&doc_name=http%3A%2F%2Fwww.nyc.gov%2Fhtml%2Fom%2Fhtml%2F2012a%2Fpr076-12.html&cc=unused1978&rc=1194&ndi=1

Elbel, Brian, Rogan Kersh, Victoria L. Brescoll, and L. Beth Dixon, 2009. "Calorie Labeling And Food Choices: A First Look At The Effects On Low-Income People In New York City." *Health Affairs* 28 (6), w1110—w1121.

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

Imbens, Guido, and Karthik Kalyanaraman, 2009. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." NBER Working Paper No. 14726. National Bureau of Economic Research, Cambridge, MA.

Jin, Ginger Zhe, and Phillip Leslie, 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *The Quarterly Journal of Economics* 118 (2), 409—451.

New York City Department of Health and Mental Hygiene, 2010. *Self-Inspection Worksheet for Food Service Establishments*. Retrieved from http://www.nyc.gov/html/doh/downloads/pdf/rii/self-inspection-worksheet.pdf

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.

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

Saul, Michael H., 2012. "Grading Eatery Grading." *The Wall Street Journal*, March 7. http://online.wsj.com/articles/SB10001424052970203458604577265852174980164

Meltzer, Rachel, Michah Rothbart., Amy E. Schwartz, Thad Calabrese, Diana Silver, Todor Mijanovich, and Meryle Weinstein, 2015. "Is Public Grading of Restaurants on Food Safety Worth the Costs? An Evaluation of New York City's Restaurant Grades Policy." Manuscript submitted for publication.

Silver, Diana, Michah Rothbart, Jean Bae, Amy E. Schwartz, and Todor Mijanovich, 2015. "Do the Stakes in Food Safety Inspections Change Restaurant Outcomes in Court." Unpublished manuscript.

Simon, Paul A., Phillip Leslie, Grace Run, Ginger Zhe Jin, Roshan Reporter, Arturo Aguirre, Jonathan E. Fielding, 2005. "Impact of Restaurant Hygiene Grade Cards on Foodborne-Disease Hospitalization in Los Angeles County." *Journal of Environmental Health* 67 (7), 32—36.

Winters, Marcus A. and Joshua M. Cowen, 2012. "Grading New York: Accountability and Student Proficiency in America's Largest School District." *Educational Evaluation and Policy Analysis* 34 (3), 313—327.

Wong, Melissa R., Wendy McKelvey, Kazuhiko Ito, Corinne Schiff, J. Bryan Jacobson, and Daniel Kass, 2015. "Impact of Letter Grade Program on Restaurant Sanitary Conditions and Diner Behavior in New York City." *American Journal of Public Health* 105 (3), e81—e87.

Table 1

Restaurant Descriptive Statistics

	Post-Public Grading
Number	
Inspections	6.2
Final Inspections	3.2
Workers	6.7
Seats	29.5
Cuisine	
American	0.24
Chinese	0.11
Pizza	0.06
Latin	0.04
Café/Coffee/Tea	0.04
Other*	0.51
Missing	0.00
	1.00
Service	
Takeout-Limited Eat in	0.39
Wait Service	0.18
Wait and Counter Service	0.17
Takeout Only	0.08
Counter Service	0.12
Other*	0.07
Missing	0.00
	1.00
Chain	0.10
Annual Closure Rate	0.12
N	34,917

Notes: Inspections include initial and re-inspections. Final inspections include initial *A* inspections and re-inspections for those initially receiving *B* or *C*. Workers, seats, cuisine, service, and chain are observed once per restaurant based on inspector reports. Annual closure rate is the fraction of open restaurants closing each year. *Other includes 76 additional cuisine types and 8 additional service types.

Table 2

Inspection Scores Statistics, by treatment period

		All Restaurants	Continuously Operating
Before Public Grading		25.1 (76,231)	23.5 (29,804)
Quarters Post Public Grading			
1-5	Initial	25.3 (41,933)	24.3 (17,723)
	Final Inspection Score	21.9 (27,874)	20.9 (11,743)
6-10	Initial	22.5 (46,180)	21.6 (18,409)
	Final Inspection Score	19.5 (29,135)	18.8 (11,544)

LOWER SCORES INDICATE MORE HYGIENIC RESTAURANT CONDITION

Data are pre-adjudicated inspection scores. Mean score shown on top; number of inspections shown in parentheses. Final inspection score includes all A-graded inspections and re-inspections of restaurants that do not get an *A* grade on initial inspection. An inspection score of 13 or lower leads to an *A*. A final inspection score of 14—27 leads to a *B* and restaurants can post *Grade Pending* until adjudication. A final inspection score of more than 27 leads to a *C* and restaurants can post *Grade Pending* until adjudication. Continuously operating restaurants are open for two and half years before and after the introduction of public grading.

Table 3

Effect of Restaurant Inspection Scores on Closures, Dif-in-Dif Estimates

	(1)	(2)	(3)	(4)
A	-0.042***	-0.026***	-0.027***	-0.026***
	(0.002)	(0.004)	(0.003)	(0.004)
C	0.049***	0.021***	0.016***	0.021***
	(0.004)	(0.006)	(0.006)	(0.006)
Inspection Score		0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)
Constant	0.104***	0.076***	0.115*	0.076***
	(0.008)	(0.009)	(0.065)	(0.009)
Quarter-Year FE	Y	Y	Y	Y
Rest. Char.	N	N	Y	N
Zip FE	N	N	Y	N
Restaurant FE	N	N	N	Y
# of Inspections	82,977	82,977	82,977	82,977
Restaurants	34,917	34,917	34,917	34,917

LOWER SCORES INDICATE MORE HYGIENIC RESTAURANT CONDITION

Robust standard errors, adjusted for within-restaurant clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Closure=1 if a restaurant is permanently closed within the next four fiscal periods. Columns (2), (3), and (4) include a control for the inspection's score. Column (3) includes restaurant controls for chain restaurants, number of workers, number of seats, and a set of indicator variables for restaurant cuisine, service type, and venue type as well as zip code fixed effects. Restaurant controls are time invariant and are measured in the most recent restaurant inspection. Column (4) includes a restaurant fixed effect and excludes time invariant restaurant and location controls. The reference group is inspections assigned a *B* grade.

Table 4

Effect of Restaurant Inspection Scores on Closures, Wald Estimate

	A - B			C - B		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A	-0.049*** (0.017)	-0.040*** (0.017)	-0.048*** (0.010)	--	--	--
C	--	--	--	0.042* (0.023)	0.039* (0.024)	0.043*** (0.015)
Constant	0.116*** (0.017)	0.281* (0.165)	0.115*** (0.010)	0.114*** (0.006)	-0.223 (0.209)	0.112*** (0.005)
Q-Y FE	Y	Y	Y	Y	Y	Y
Rest. Char.	N	Y	N	N	Y	N
Zip FE	N	Y	N	N	Y	N
# Inspections	7,387	7,387	17,113	2,921	2,921	5,398
Restaurants	6,812	6,812	13,609	2,710	2,710	4,744
Bandwidth	1 points	1 points	2 points	1 points	1 points	2 points

Robust standard errors, adjusted for within-restaurant clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Columns (1), (2), (4), and (5) restrict the sample to inspections one point above and one point below the grade cutoff. Columns (3) and (6) restrict the sample to inspections two points above and two points below the grade cutoff. The optimal bandwidth for a local linear RD estimate of an *A* inspection effect, which minimizes MSE as in Imbens and Kalyanaraman (2009), is 1.293 points. The optimal bandwidth for a local linear RD estimate of a *C* inspection effect is 1.939 points. The estimated effects in local linear models are exactly equal to columns (1) and (4), respectively. Columns (2) and (5) include a include restaurant controls for chain restaurants, number of workers, number of seats, and a set of indicator variables for restaurant cuisine, service type, and venue type as well as zip code fixed effects. Restaurant controls are time invariant and are measured in the most recent restaurant inspection. The reference group is inspections assigned a *B* grade.

Table 5

Effect of Restaurant Inspection Scores on Inspection-Level Fines ($), Wald Estimate

	A - B		C - B	
	(1)	**(2)**	**(3)**	**(4)**
A	-518.46*** (34.53)	-553.00*** (18.19)	—	—
C	—	—	-17.40 (53.29)	100.37** (39.01)
Constant	1129.91*** (46.00)	1157.09*** (29.75)	1016.17*** (124.21)	1127.49*** (93.89)
Q-Y FE	Y	Y	Y	Y
Rest. Char.	N	N	N	N
Zip FE	N	N	N	N
# Inspections	7,387	17,113	2,921	5,398
Restaurants	6,812	13,609	2,710	4,744
Bandwidth	1 points	2 points	1 points	2 points

Robust standard errors, adjusted for within-restaurant clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Columns (1) and (3) restrict the sample to inspections one point above and one point below the grade cutoff. Columns (2) and (4) restrict the sample to inspections two points above and two points below the grade cutoff. The optimal bandwidth for a local linear RD estimate, which minimizes MSE as in Imbens and Kalyanaraman (2009), is 3.77 for columns (1) and (2) and 10.35 for columns (3) and (4). The estimated effects in local linear models are qualitatively similar (-609.21 and significant for A; 22.32 and insignificant for *C*). The reference group is inspections assigned a *B* grade.

Table 6

Effect of Restaurant Grades on Sales ($), Dif-in-Dif Estimates

	(1)	(2)	(3)	(4)
A	144.71***	77.61**	123.33**	82.86**
	(51.58)	(31.91)	(55.06)	(35.77)
C	-103.77	-122.26***	-113.04	-143.65***
	(124.06)	(43.83)	(126.41)	(53.67)
Grade Pending:				
B	1.54	--	6.40	--
	(67.38)		(67.59)	
C	-20.00	--	13.12	--
	(84.37)		(88.38)	
Inspection Score	--	--	-0.68	0.80
			(1.47)	(1.44)
Ungraded	45.26	58.65	61.40	73.34
	(63.43)	(44.87)	(63.60)	(45.31)
Building Class FE	Y	Y	Y	Y
Quarter-Year FE	Y	Y	Y	Y
Group FE	Y	Y	Y	Y
Constant	2,347.31***	2,393.44***	2,396.65***	2,380.98***
	(449.56)	(448.36)	(449.78)	(449.27)
Observations	9,182	9,182	9,182	9,182
Groups	1,538	1,538	1,538	1,538
R-squared	0.98	0.98	0.98	0.98

Robust standard errors, adjusted for within-group clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Table shows estimated impact of restaurant grades on daily sales. Columns 1 and 3 show estimates of the impact of average posted grade on sales. Columns 2 and 4 shows estimates of the impact of ITT grade earned at inspection by the beginning of the quarter on sales. *A* and *C* are share of a group with an *A* or *C* grade, respectively. Because estimates are reported on the means of all variables, these are estimates of impacts on a single restaurant. *Grade Pending* are the share of group with the option to post either "Grade Pending" or the grade indicated in their window. Columns 3 and 4 include controls for inspection score and all models control for building class, group fixed effects (which controls for time invariant group characteristics such as share serving each cuisine type), and quarter-by-year fixed effects. The reference group is restaurants posting *B* grades.

Table 7

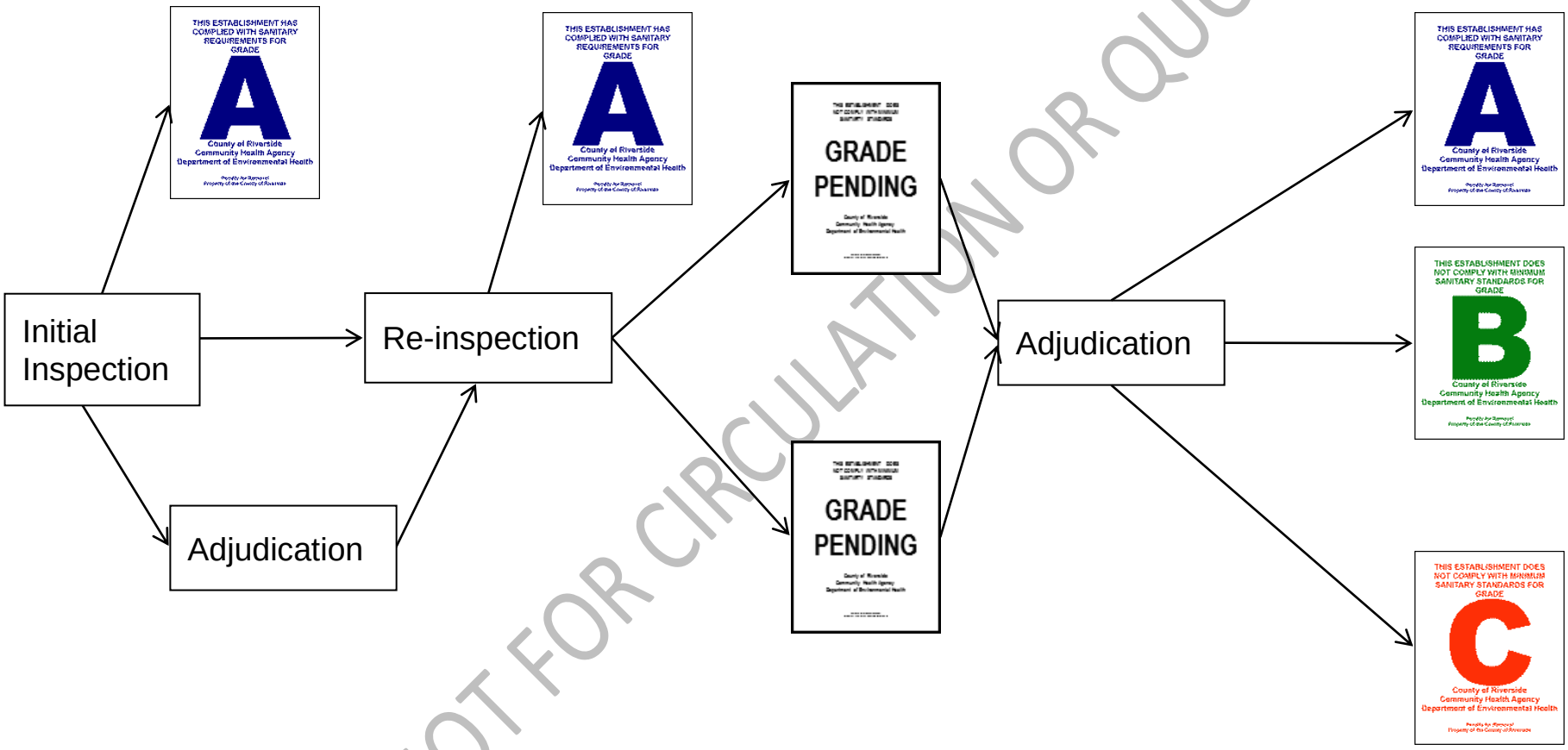
Effect of Restaurant Grades on Sales Taxes ($), Dif-in-Dif Estimates

	(1)	(2)
A	5.93**	4.08***
	(2.44)	(1.58)
C	-5.07	-5.97**
	(5.59)	(2.37)
Grade Pending:		
B	0.83	--
	(2.99)	
C	0.41	--
	(3.91)	
Inspection Score	-0.01	0.05
	(0.07)	(0.06)
Ungraded	2.13	2.61
	(2.81)	(2.00)
Building Class FE	Y	Y
Quarter-Year FE	Y	Y
Group FE	Y	Y
Constant	107.84***	107.30***
	(19.89)	(19.88)
Observations	9,182	9,182
Groups	1,538	1,538
R-squared	0.98	0.98

Robust standard errors, adjusted for within-group clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Table shows estimated impact of restaurant grades on daily sales taxes. Column 1 shows estimates of the impact of average posted grade on sales taxes. Column 2 shows ITT estimates of the impact of grade earned at inspection by the beginning of the quarter on sales taxes. *A* and *C* are share of a group with an *A* or *C* grade, respectively. Because estimates are reported on the means of all variables, these are estimates of impacts on a single restaurant. *Grade Pending* are share of group with the option to post either "Grade Pending" or the grade indicated in their window. All models control for inspection score, building class, group fixed effects (which controls for time invariant group characteristics such as share serving each cuisine type), and quarter-by-year fixed effects. The reference group is restaurants posting *B* grades.

Figure 1

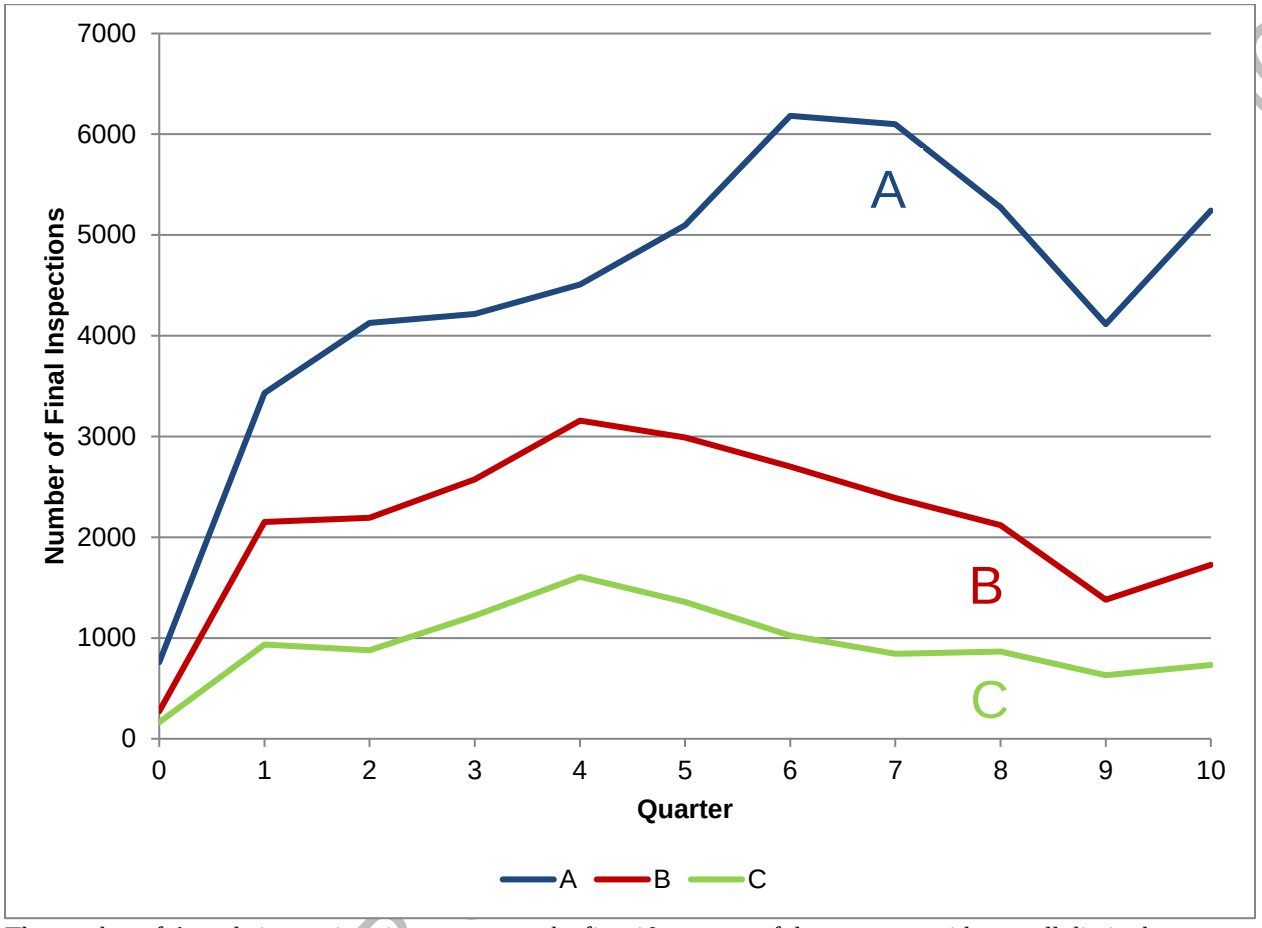
A Simplified Model Of The Inspection Cycle Post-Grading



Fines assessed for violations at each inspection that does not lead to an *A* (assessed for *A* inspections for first six months of grading)

Figure 2

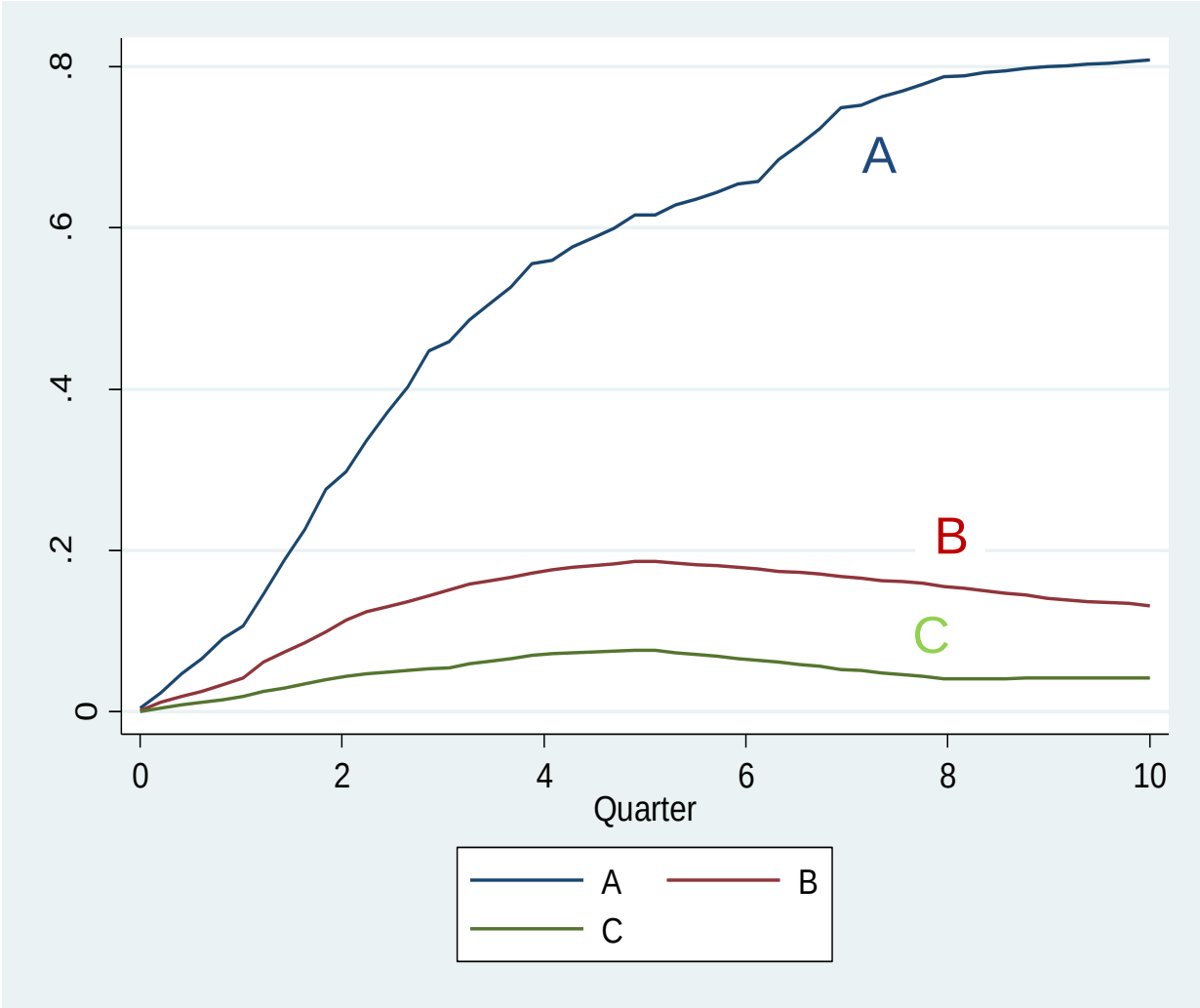
Inspection Grades Awarded by Quarters Post-Grading, Intention-to-Treat



The number of *A* grade inspections increases over the first 10 quarters of the program, with a small dip in the quarter of and following Hurricane Sandy. The number of *B* and *C* inspections increases for the first four quarters and then begins to decline.

Figure 3

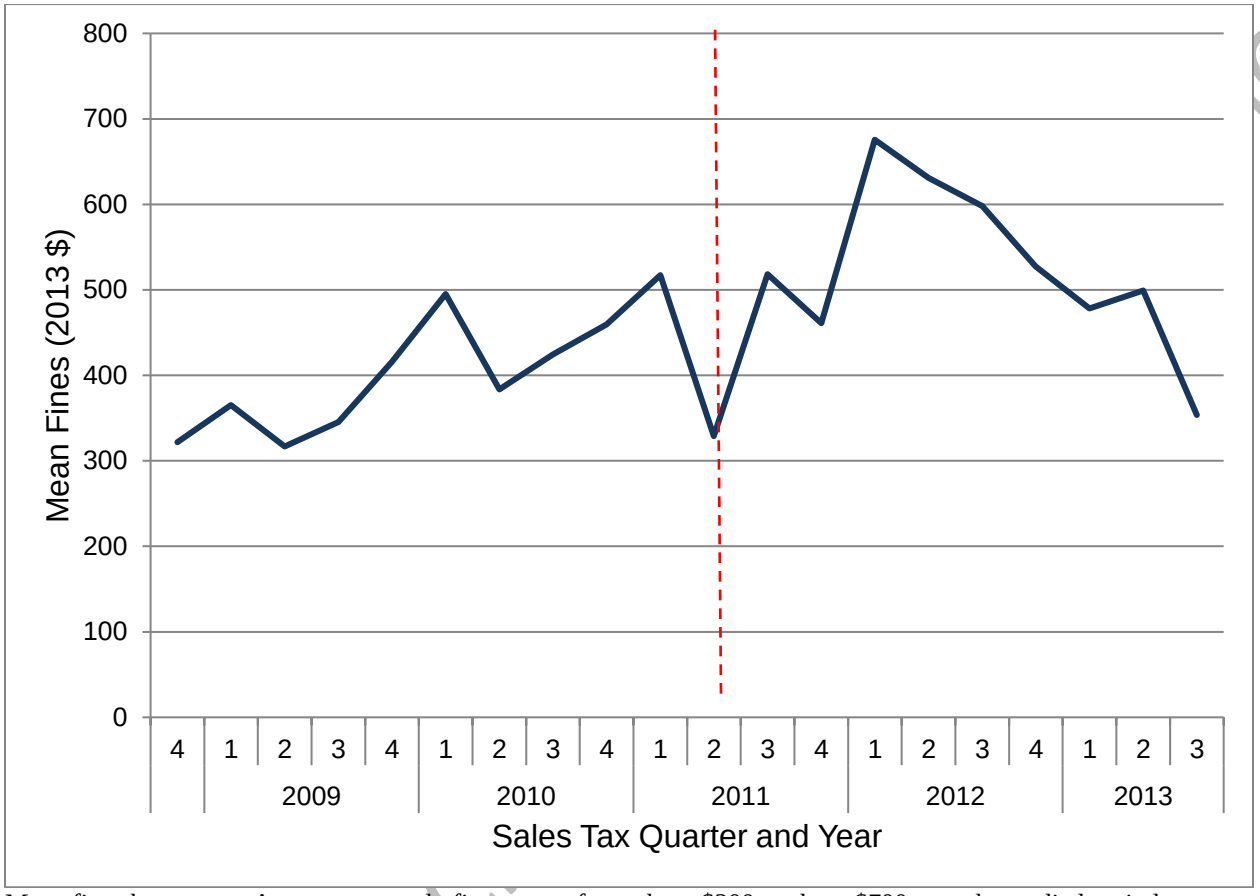
Grades in the Window by Quarters Post-Grading, Rollout of Treatment



The share of restaurants with *A* grades posted in the window increases over the first 10 quarters of the program and reaches 80 percent by the end of the sample period. The share of restaurants with *B* and *C* grades posted in the window increases for the first five quarters and then begins to decline.

Figure 4

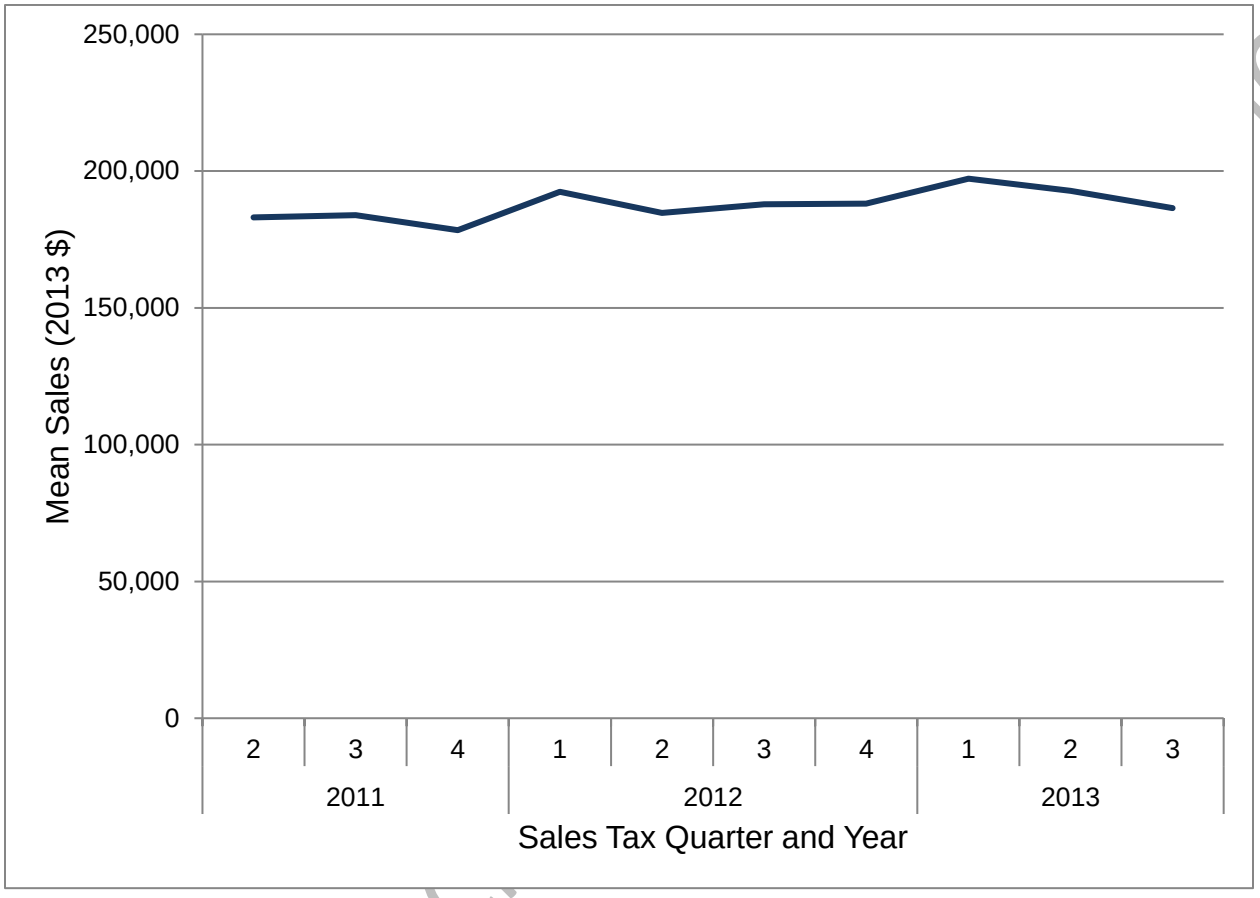
Average Fines by Quarter, Operating Restaurants, Pre and Post-Grading



Mean fines by quarter. Average quarterly fines range from about $300 to about $700 over the studied period. Average fines increase substantially in the first year after the restaurant grades law and then steadily decline from there. Average fines levied citywide are at pre-program levels on a quarterly basis starting in the middle of the 2013 sales tax year.

Figure 5

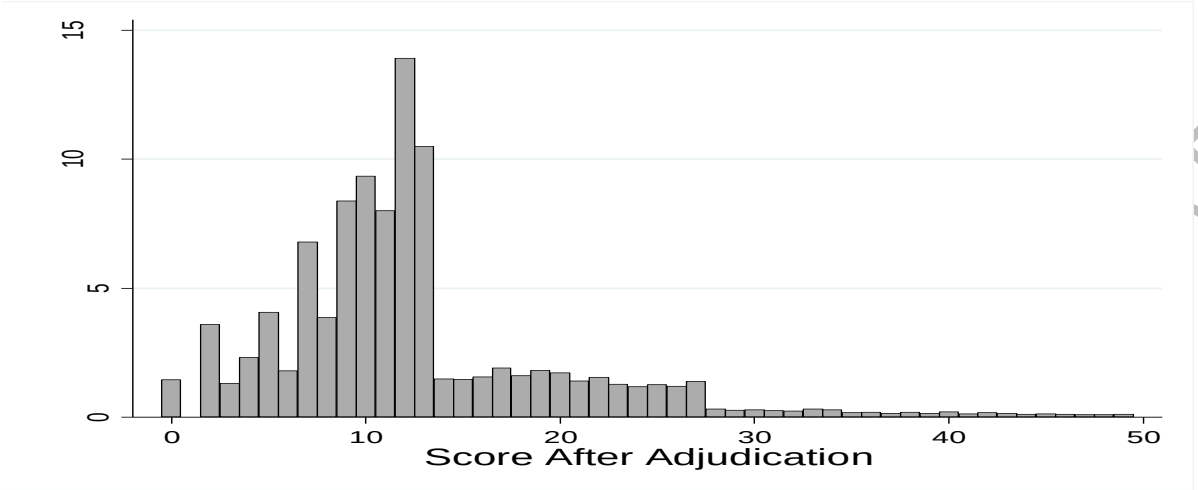
Average Sales, Operating Food And Beverage Entities, Post-Grading



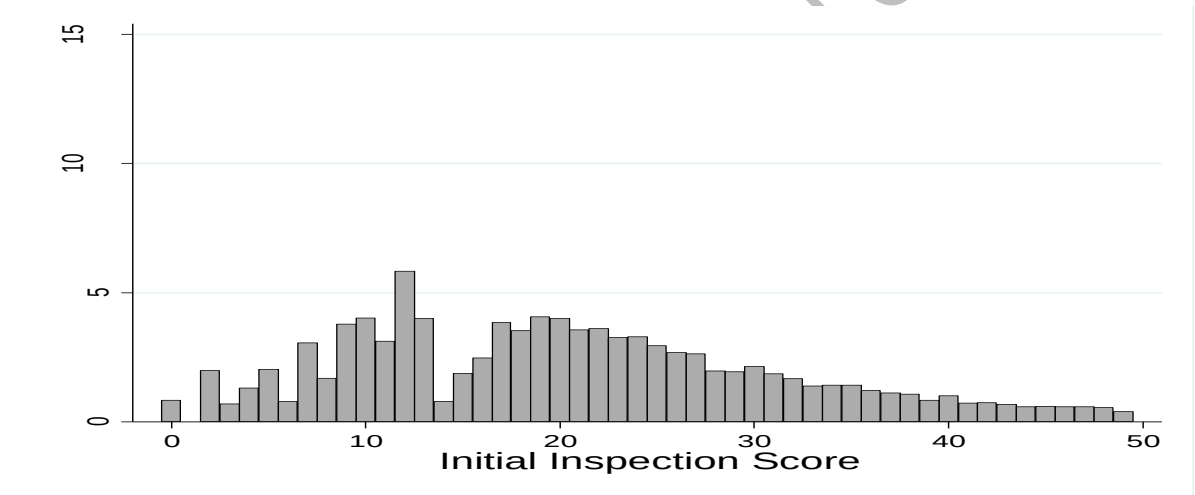
Average restaurant sales increase following the implementation of the public grades program in the second quarter of FY 2011 (July, 2010). Mean sales during this period ranges from about $175,000 a quarter to close to $200,000 a quarter. Mean sales masks large levels of heterogeneity across restaurants, including heterogeneous impacts of grades that this paper explores.

Figure 6. Inspection Score Distributions, After and Before Grading

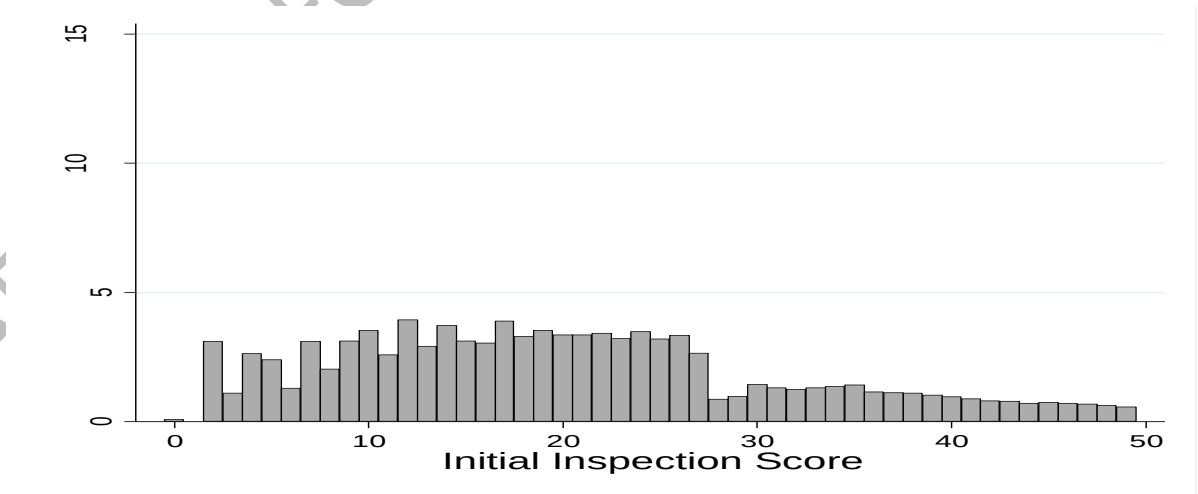
Panel A. Post-Adjudication Score Distribution, Second Year of Grading



Panel B. Initial Inspection Score Distribution, Second Year of Grading



Panel C. Initial Inspection Score Distribution, Two Years Before Grading



APPENDIX A

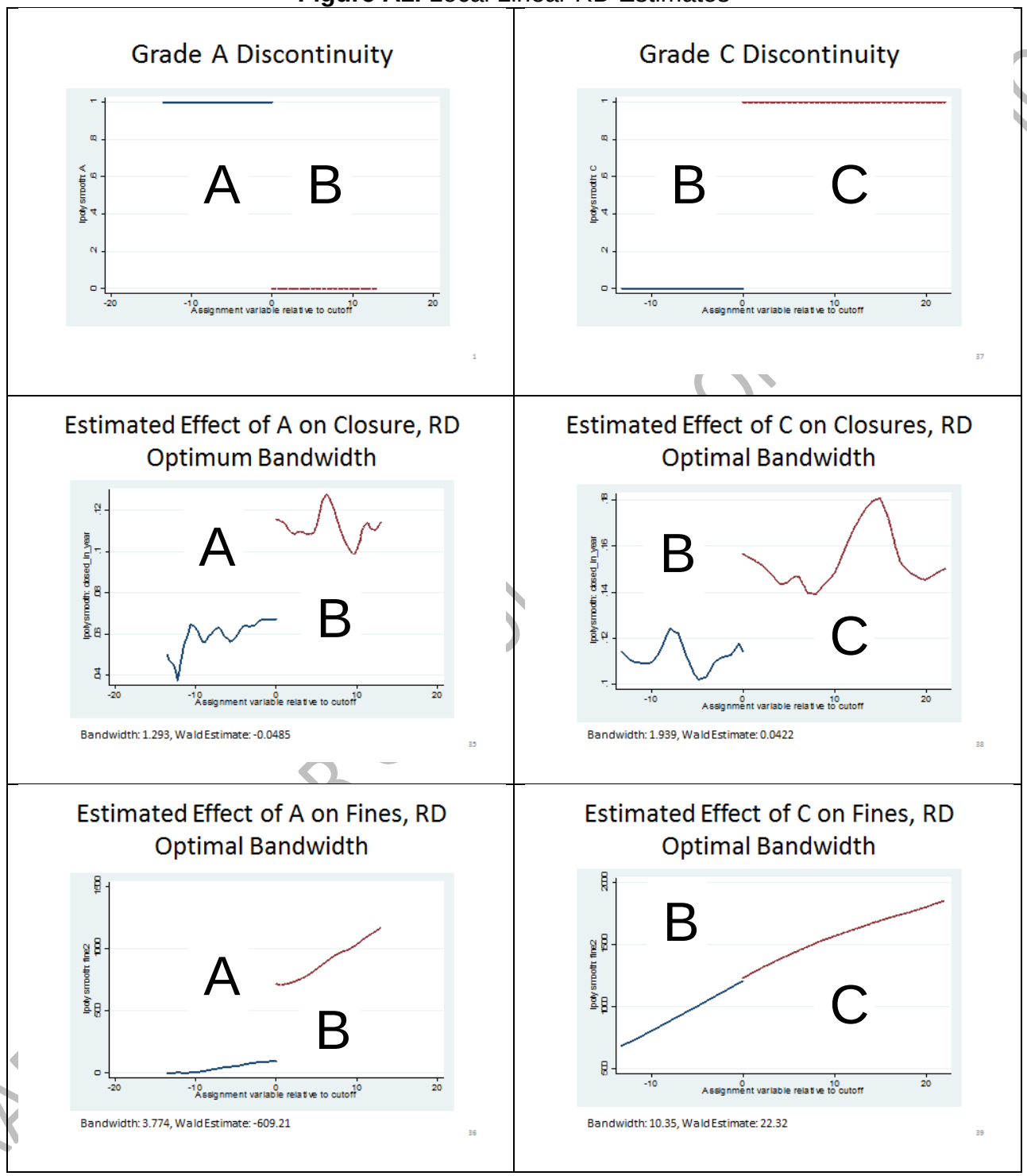
Table A1. Measures of Building Class

Building Class	**Included AV-BLDGCL**
Office Commercial	O
Retail Commercial	K, L8, RK
Mixed Use Retail	C7, D6, D7, S
Other Commercial	E, F, G, H, I, J, M, L (except L8), N, RA—RW (except RK)
Residential	A, B, C (except C7), D (except D6 and D7), R0—R4
Government/Public	P, Q, T, U, V, W, Y, Z

Notes: AV-BLDGCL codes come from the 2010—2013 RPAD data.

APPENDIX B

Figure A1. Local Linear RD Estimates



APPENDIX C

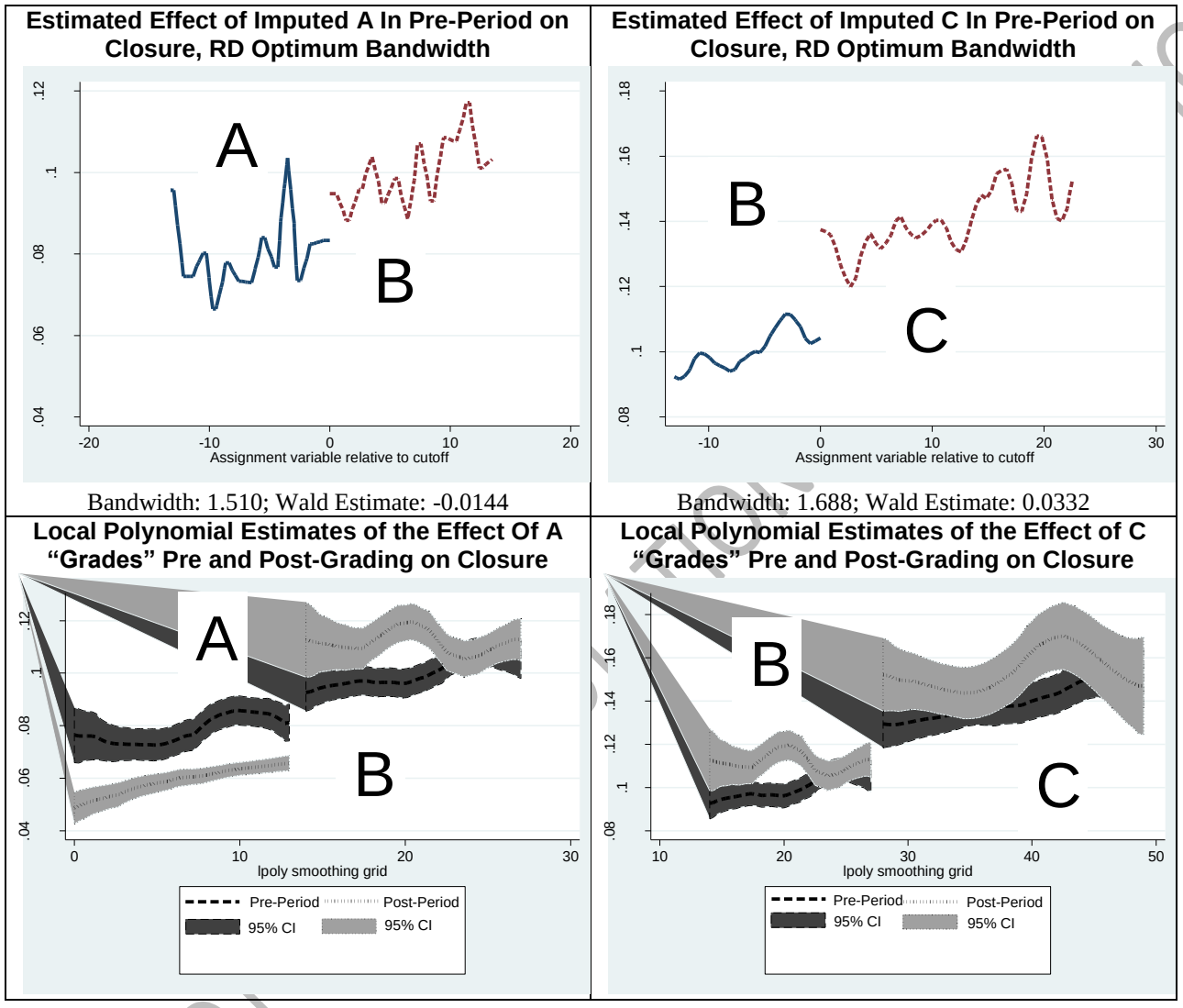
Table A2. Falsification Test, Restaurant Closure, Pre-Period

	A - B			C - B		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A	-0.012 (0.009)	-0.014* (0.008)	-0.009 (0.006)	--	--	--
C	--	--	--	0.031** (0.015)	0.001 (0.013)	0.030 (0.020)
Constant	0.064*** (0.014)	-0.293*** (0.041)	0.059*** (0.009)	0.066*** (0.0173)	0.753*** (0.029)	0.057*** (0.011)
Q-Y FE	Y	Y	Y	Y	Y	Y
Rest. Char.	N	Y	N	N	Y	N
Zip FE	N	Y	N	N	Y	N
# Inspections	4,436	4,436	9,431	2,745	2,745	5,960
Restaurants	4,156	4,156	8,268	2,626	2,626	5,400
Bandwidth	1 points	1 points	2 points	1 points	1 points	2 points

Robust standard errors, adjusted for within-restaurant clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Columns (1), (2), (4), and (5) restrict the sample to inspections one point above and one point below the grade cutoff. Columns (3) and (6) restrict the sample to inspections two points above and two points below the grade cutoff. The optimal bandwidth for a local linear RD estimate of an *A* inspection effect, which minimizes MSE as in Imbens and Kalyanaraman (2009), is 1.510 points. The optimal bandwidth for a local linear RD estimate of a *C* inspection effect is 1.688 points. Columns (2) and (5) include a include restaurant controls for chain restaurants, number of workers, number of seats, and a set of indicator variables for restaurant cuisine, service type, and venue type as well as zip code fixed effects. Restaurant controls are time invariant and are measured in the most recent restaurant inspection. The reference group is inspections assigned a *B* grade.

APPENDIX D

Figure A2. Falsification Test, Restaurant Closure, Local-Linear RD Pre-Period Graphs



APPENDIX E

Table A3. Regression results, Difference-in-Differences Model, Sales Post-Grading

	Posted			ITT		
Measurement Time:	Beginning	End	Daily Average	Beginning	End	Daily Average
A	115.56**	111.12**	144.71***	77.61**	24.50	48.18
	(49.42)	(47.16)	(51.58)	(31.91)	(31.63)	(35.48)
C	-147.05	-105.88	-103.77	-122.26***	-47.08	-117.77**
	(118.27)	(112.59)	(124.06)	(43.83)	(42.95)	(53.06)
Grade Pending:						
B	-70.57	79.08	1.54	--	--	--
	(56.97)	(54.95)	(67.38)			
C	-76.42	44.20	-20.00	--	--	--
	(70.20)	(68.38)	(84.37)			
Ungraded	52.49	0.62	45.26	58.65	-74.96	18.55
	(56.91)	(57.58)	(63.43)	(44.87)	(51.23)	(51.38)
Building Class FE	Y	Y	Y	Y	Y	Y
Quarter-Year FE	Y	Y	Y	Y	Y	Y
Group FE	Y	Y	Y	Y	Y	Y
Constant	2,373.86***	2,383.63***	2,347.31***	2,393.44***	2,452.30***	2,422.82***
	(449.30)	(449.23)	(449.56)	(448.36)	(448.40)	(448.91)
Observations	9,182	9,182	9,182	9,182	9,182	9,182
Groups	1,538	1,538	1,538	1,538	1,538	1,538
R-squared	0.98	0.98	0.98	0.98	0.98	0.98

Robust standard errors, adjusted for within-group clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Table shows estimated impact of restaurant grades on sales. "Posted" shows estimates of the impact of grade on sales where grade is measured at the beginning, end, or averaged over the quarter where indicated. "ITT" shows estimates of the impact of grade earned at inspection on sales, where grade is measured at the beginning, end, or averaged over the quarter where indicated. *A* and *C* are share of a group with an *A* or *C* grade, respectively. Because estimates are reported on the means of all variables, these are estimates of impacts on a single restaurant. *Grade Pending* are share of group with the option to post either "Grade Pending" or the grade indicated in their window. All models control for building class, group fixed effects (which controls for time invariant group characteristics such as share serving each cuisine type), and quarter-by-year fixed effects. The reference group is restaurants posting *B* grades.

APPENDIX F

Table A4. Regression results, Alternative Measures of Quarterly Grades, Sales

	Posted			ITT		
Measurement Time:	Beginning	End	Daily Average	Beginning	End	Daily Average
A	106.01**	80.16	123.33**	82.86**	7.61	33.21
	(51.88)	(49.73)	(55.06)	(35.77)	(35.97)	(39.61)
C	-137.16	-83.74	-113.04	-143.65***	-21.09	-104.68
	(119.86)	(113.55)	(126.41)	(53.67)	(54.22)	(64.33)
Grade Pending:						
B	-69.52	73.96	6.40	--	--	--
	(57.30)	(54.89)	(67.59)			
C	-64.05	62.07	13.12	--	--	--
	(72.14)	(70.13)	(88.38)			
Inspection Score	-0.70	-0.75	-0.68	0.80	-1.06	-0.66
	(1.17)	(1.25)	(1.47)	(1.44)	(1.59)	(1.69)
Ungraded	70.01	26.49	61.40	73.34	-57.36	24.76
	(57.29)	(57.26)	(63.60)	(45.31)	(50.79)	(51.20)
Building Class FE	Y	Y	Y	Y	Y	Y
Quarter-Year FE	Y	Y	Y	Y	Y	Y
Group FE	Y	Y	Y	Y	Y	Y
Constant	2,397.90***	2,469.73***	2,396.65***	2,380.98***	2,540.30***	2,471.45***
	(450.09)	(447.04)	(449.78)	(449.27)	(446.62)	(449.11)
Observations	9,182	9,182	9,182	9,182	9,182	9,182
Groups	1,538	1,538	1,538	1,538	1,538	1,538
R-squared	0.98	0.98	0.98	0.98	0.98	0.98

Robust standard errors, adjusted for within-group clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Table shows estimated impact of restaurant grades on sales. "Posted" shows estimates of the impact of grade on sales where grade is measured at the beginning, end, or averaged over the quarter where indicated. "ITT" shows estimates of the impact of grade earned at inspection on sales, where grade is measured at the beginning, end, or averaged over the quarter where indicated. *A* and *C* are share of a group with an *A* or *C* grade, respectively. Because estimates are reported on the means of all variables, these are estimates of impacts on a single restaurant. *Grade Pending* are share of group with the option to post either "Grade Pending" or the grade indicated in their window. All models control for inspection score, building class, group fixed effects (which controls for time invariant group characteristics such as share serving each cuisine type), and quarter-by-year fixed effects. The reference group is restaurants posting *B* grades.

APPENDIX G

Table A5. Regression results, Alternative Measures of Quarterly Grades, Sales Taxes

	Posted			ITT		
Measurement Time:	Beginning	End	Daily Average	Beginning	End	Daily Average
A	4.47*	4.00*	5.93**	4.08***	0.48	1.97
	(2.30)	(2.20)	(2.44)	(1.58)	(1.59)	(1.75)
C	-5.13	-3.39	-5.07	-5.97**	-0.83	-4.23
	(5.30)	(5.02)	(5.59)	(2.37)	(2.40)	(2.85)
Grade Pending:						
B	-3.37	3.42	0.83	--	--	--
	(2.54)	(2.43)	(2.99)			
C	-2.99	3.04	0.41	--	--	--
	(3.19)	(3.10)	(3.91)			
Inspection Score	-0.02	-0.02	-0.01	0.05	-0.04	-0.01
	(0.05)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)
Ungraded	1.86	1.13	2.13	2.61	-3.08	0.52
	(2.53)	(2.53)	(2.81)	(2.00)	(2.25)	(2.26)
Building Class FE	Y	Y	Y	Y	Y	Y
Quarter-Year FE	Y	Y	Y	Y	Y	Y
Group FE	Y	Y	Y	Y	Y	Y
Constant	108.45***	111.14***	107.84***	107.30***	114.66***	111.27***
	(19.91)	(19.77)	(19.89)	(19.88)	(19.76)	(19.86)
Observations	9,182	9,182	9,182	9,182	9,182	9,182
Groups	1,538	1,538	1,538	1,538	1,538	1,538
R-squared	0.98	0.98	0.98	0.98	0.98	0.98

Robust standard errors, adjusted for within-group clusters, in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Table shows estimated impact of restaurant grades on sales taxes. "Posted" shows estimates of the impact of grade on sales taxes where grade is measured at the beginning, end, or averaged over the quarter where indicated. "ITT" shows estimates of the impact of grade earned at inspection on sales taxes, where grade is measured at the beginning, end, or averaged over the quarter where indicated. *A* and *C* are share of a group with an *A* or *C* grade, respectively. Because estimates are reported on the means of all variables, these are estimates of impacts on a single restaurant. *Grade Pending* are share of group with the option to post either "Grade Pending" or the grade indicated in their window. All models control for inspection score, building class, group fixed effects (which controls for time invariant group characteristics such as share serving each cuisine type), and quarter-by-year fixed effects. The reference group is restaurants posting *B* grades.