A TALE OF REPETITION:
LESSONS FROM FLORIDA RESTAURANT INSPECTIONS

Ginger Zhe Jin
Jungmin Lee

A Tale of Repetition: Lessons from Florida Restaurant Inspections
Ginger Zhe Jin and Jungmin Lee
NBER Working Paper No. 20596
October 2014, Revised February 2018
JEL No. D02,D81,D82,K32,L51

ABSTRACT

This paper examines the role of repetition in government regulation using Florida restaurant inspection data from 2003 to 2010. In the raw data, inspectors new to inspected restaurants tend to report 27% more violations than repeat inspectors. After ruling out regulatory capture and endogenous inspector rotation as potential explanations, we find that the new-repeat gap is best explained by the following two effects: first, restaurants target compliance in response to heterogenous stringency and tastes of different inspectors; second, inspectors pay greater attention in their first visit than in subsequent visits. After controlling for heterogenous inspector criteria, we find that a new inspector reports 13-18% more violations than the second visit of the previous inspector, likely due to a higher level of attention. Counterfactual simulations highlight the importance of inspector training and rotation in regulatory outcomes.

Ginger Zhe Jin
University of Maryland
Department of Economics
3115F Tydings Hall
College Park, MD 20742-7211
and NBER
jin@econ.umd.edu

Jungmin Lee
Sogang University
Seoul, Korea
junglee@sogang.ac.kr

1 Introduction

From nuclear power to food safety, government employees inspect economic entities on a regular basis. Such an inspection introduces a classical double-moral-hazard problem: on the one hand, government-employed inspectors may not detect or report every violation as the principal desires; on the other hand, regulated firms may not comply with every rule set by the principal. Theorists have made many suggestions to alleviate the moral hazard problem, including outcome-based contracts,[1] targeted auditing, reduction of information rents, high penalties for corrupt inspectors, or intentional selection of biased employees,[2] but they are often difficult to implement due to such constraints as rigid compensation schemes and limited resources.

This paper focuses on inspector repetition/rotation, a feature that exists in almost all inspection programs but is rarely studied in the literature. In the program we study, the Florida Division of Hotels and Restaurants (DHR hereafter) conducts regular unannounced inspections at restaurants. In our sample, an individual inspector is new to a particular inspected restaurant on average 27% of the time. Comparing new and repeat inspectors, we find a stark difference in inspection outcomes. Figure 1A shows the average number of reported violations in a typical new-repeat history of a restaurant, based on the DHR inspection records from July 2003 to March 2010. It starts with the first new inspector for a restaurant, followed by up to four repeated visits of the same inspector and then the next round for the next new inspector. The new-repeat gap is striking[3]: a new inspector tends to report more violations than her predecessor, but the number of reported violations declines sharply when she returns. Not only does the cycling pattern hold for different types of violations, it also persists after we control for restaurant age, year-month fixed effects and restaurant fixed effects (Figure 1B).

What drives the cycling pattern? We propose two explanations. The first one is based on *heterogeneous inspector criteria*. Inspectors differ in their stringency and tastes. Since compliance is not costless to restaurants, they should determine the effort of compliance

---

[1]For example, the principal may set inspectors' compensation conditional on the reported violations or design a dynamic contract to prevent collusion between inspectors and the regulated (Tirole, 1986; Martimort, 1999).

[2]See Laffont and Tirole (1993), Mookherjee and Png (1989, 1995) for specific theories and Prendergast (1999) and Dixit (2002) for comprehensive reviews. Prendergast (2007) focuses on biased bureaucrats in particular.

[3]More details about Figure 1 will be discussed in Section 2.2.

given the expectation about inspector type. Suppose that a restaurant believes that the same inspector returns for the next inspection. Then it is optimal for the restaurant to comply more in the areas where the previous inspector focused on. Such targeted compliance may be reinforced after the inspector returns multiple times, because repeat interactions reveal the inspector's preferences. As the restaurant caters to the inspector's idiosyncratic preference, it deviates from the average of other inspectors' preferences. This may explain why the new inspector finds more violations, especially if the previous inspector has had a longer relationship with the restaurant.

The second explanation is based on *diminishing inspector attention*. When a new inspector inspects a restaurant for the first time, she is not familiar with the restaurant and does not know as to what areas are vulnerable to violations in the restaurant. Therefore, it is optimal for her to check out every corner with a high level of attention. However, returning to the restaurant, with the knowledge she acquired before, she may find it more efficient to devote efforts to the areas in which she found problems in the last inspection. Anticipating the inspector to reduce her attention gradually, the restaurant would focus more on previously detected violations, which leads to fewer violations reported by a repeat inspector.

Both explanations reflect strategic detection and compliance, but the key difference lies in inspector heterogeneity. In the first explanation, inspectors differ in taste and stringency. This inherent but idiosyncratic preference applies to all restaurants inspected by the same individual inspector. In contrast, diminishing inspector attention emphasizes within-inspector rather than across-inspector variations. It depends on how an inspector varies her detection effort over time, even if every inspector has exactly the same inherent preference and faces the same detection cost. As detailed below, we attempt to separate these two effects empirically by using inspector-violation type fixed effects and by allowing inspection outcomes to depend directly on the previous inspector's inherent preference.

Separating the two explanations sheds light on the design of the inspection program. In particular, if the new-repeat phenomenon is driven by heterogeneity of inspector criteria, the DHR may want to train all inspectors to ensure homogeneity. However, if diminishing inspector attention is the dominant reason, training only ensures that every inspector behaves the same way as a new inspector, but it does not address progressive shirking after the first inspection. Another policy lever is inspector rotation, which addresses both

effects, but its effectiveness depends on the relative importance of the two. Thus, using our empirical estimates, we quantitatively simulate the effect of random inspector assignment (which increases the probability of new inspection) versus the effect of making inspector homogenous (which does not change the frequency of new inspection but reduces inspector heterogeneity). Counterfactual simulations reinforce the conclusion that both inspector-targeted compliance and diminishing inspector attention contribute significantly to inspection outcomes.

There might be other possible explanations for the new-repeat gap. Among those, probably the most popular explanation is regulatory capture: maybe new inspectors are more likely to follow the books in the first visit but get captured afterwards. However, the DHR has already built in a number of institutional safeguards against inspector capture. The probability of a fine is extremely low for a regular inspection (1.6%), the fine amount (on average, $861 if fined) is determined by a separate branch not the inspector, and the restaurant can contest and appeal any inspection outcome through a formal process. These institutions suggest that the expected economic return of capture (to restaurants) is low. Even if the inspectee is determined to bribe the inspector, he should have incentives to do so in the first visit, not until the inspector returns. Another potential explanation is endogenous assignment of new inspectors: if the DHR is more likely to assign new inspectors to dirty restaurants, Figure 1 could simply reflect endogenous inspector selection. We rule out this explanation, as the same pattern remains after we use propensity score matching and instrumental variables to address endogenous inspector rotation.

Our work contributes to several strands of the economics literature. First, our empirical analysis complements a large body of literature on principal-agent theory. Second, we aim to extend the literature on inspector heterogeneity, in which there are only a few studies until now. For example, Feinstein (1989) estimated inspector fixed effects in nuclear plant inspections and found significant inspector heterogeneity, especially the presence of underperformers. In another paper, Feinstein (1991) also found substantial heterogeneity among IRS tax examiners in terms of detection rates. Macher et al. (2011) examined FDA inspections of pharmaceutical manufacturing and found considerable heterogeneity among regulators. On top of this literature, we demonstrate the new-repeat difference and clarify the role that inspector heterogeneity plays in this difference. As for the new-repeat difference itself, the closest study is Short, Toffel and Hogill (2016). They examined global

4

supply chain audits for code-of-conduct compliance with labor standards, and found that audits report more violations when auditors have not audited the establishment previously. We extend their study by quantifying heterogenous inspector criteria and diminishing attention as two separate explanations for the new-repeat difference. This distinction allows us to shed light on alternative inspection designs.[4]

The rest of the paper is organized as follows. In Section 2, we describe the data and demonstrate the new-repeat difference in inspection outcomes. Section 3 examines endogenous inspector assignment as a potential explanation. In Section 4, we first explain how heterogenous inspector criteria and diminishing inspector attention can lead to a gap between new and repeat inspectors, and then quantify the two effects empirically. We also present the results from counterfactual simulations. A brief conclusion is offered in Section 5.

2 Data Description and the New-Repeat Difference

2.1 Data Description

Our sample is constructed from three administrative data sets collected by the DHR: (1) restaurant/food service inspection files, (2) restaurant license files, and (3) restaurant disciplinary activity reports. The data include all restaurant inspections in Florida from July 2003 to March 2010. All food establishments are required to be inspected twice per fiscal year by state laws and three times by administrative rules. However, due to labor shortage, the average number of regular inspections per restaurant per year is less than 2 except for the 2008-2009 fiscal year. In fact, about 20-40% of restaurants received only one regular inspection a year.

The final sample is constructed in several steps. Starting with 600,492 regular inspections in the raw data, we exclude the first six months of each restaurant since its first appearance in our data because we use these months to define the restaurant's history. If the first six months do not cover the restaurant's first regular inspection in our data, we exclude the restaurant's history up to its first regular inspection. We also exclude any inspections conducted before March 2004 because Florida reclassified some non-critical vi-

---

[4]Our paper is different from Short et al. (2016) in many other regards. For example, we examine government inspections while they examine private-sector inspections. Also the inspections in our paper are conducted by individual inspectors, while those in their paper are done by audit teams.

olations as risk-factors at the time. As we apply restaurant fixed effects in all estimations, we also exclude the 11,819 restaurants that have only one inspection throughout the sample period.

The final sample includes 426,831 regular inspections, covering 60,976 unique restaurants and 358 individual inspectors.[5] Each year there are around 220 active inspectors. Each inspector conducts, on average, more than 200 inspections per year although this number varies greatly across inspectors. There are 55 violation items, among which 18 are critical violations, 26 are non-critical, and 11 are risk factors.[6] Throughout the paper, we refer to critical, non-critical and risk-factor as three violation categories. As shown in Table 1, the average number of violations per inspection is 7.9, of which 1.62 are critical, 2.54 are risk factors, and 3.75 are non-critical. About 96% of regular inspections are routine ones, while 3.7% are initiated by complaints and 0.1% are licensing inspections.[7]

To get a sense of inspector heterogeneity, we simply regress the total number of violations per inspection on a full set of individual inspector dummies, controlling for fiscal year, month, and restaurant fixed effects. Such a regression yields an adjusted R-square of 0.51, which is higher than that without inspector fixed effects (0.46). As shown in Appendix Figure 1, the estimated fixed effects range from -6 to 16, except few outliers. This is huge given the fact that the average number of violations is 7.9 with a standard deviation of 7.2. The large heterogeneity is consistent with those documented for nuclear inspectors, tax auditors, and pharmaceutical plant inspectors (Feinstein 1989, 1991; Macher et al. 2011).

Inspectors may have different criteria on different items. To shed light on this, we repeat the regression exercise with inspector-item fixed effects. Within each inspector, we code the item with the largest inspector-item fixed effect as the inspector's most frequently cited item (MFCI).[8] Appendix Figure 2 plots the histogram of MFCI across all 55 items. The distribution is dispersed, with relatively high frequencies in certain categories (notably, 2, 8, 14, 22, 32, 37 and 45), which proves substantial heterogeneity in inspectors' tastes while there are some focal items that most inspectors pay attention to.

Using the original DHR dataset, we define inspector i as "new" to restaurant r if the

---

[5]The original inspection files include 386 inspectors and 97,990 restaurants.

[6]See Appendix Table A1 for the details of the 55 items.

[7]The results for the sample of only routine inspections are similar to those we will present below. The results can be obtained from the authors upon request.

[8]Before determining which category is an inspector's favorite, we test the statistical significance of each inspector-category fixed effect and exclude all that are insignificant from zero by 95% confidence.

observed inspection is the first inspection conducted by i at r during the whole sample period. As shown in Table 1, about 27% of regular inspections are conducted by new inspectors, and the average number of inspections by repeat inspectors is 3.7. That is, on average, a restaurant is inspected 3 to 4 times by the same inspector until she is replaced by a new inspector. Since a majority of inspections are conducted by repeat inspectors, restaurants should on average expect a repeat inspector rather than a new inspector.

The DHR dataset provides a rich set of variables which may be correlated with restaurant or inspector efforts. We will use these variables as control variables in our regression analysis below. With no access to inspector personnel data, we proxy inspector experience by the number of regular inspections that an inspector has conducted in our data before a specific inspection. As shown in Table 1, the average inspector experience is 1,535. In addition, 44% of inspections are conducted by inspectors with less than the median level of experience (inexperienced inspectors), and 1% are conducted by inspectors with the experience of less than 30 inspections (novice inspectors). Another interesting variable is the number of inspections that the inspector has done in a day. An inspector may become tired during the day and incur higher effort costs due to fatigue. On average, an inspector has completed 1.6 inspections before coming to the inspection under study and 28% of inspections are the first one in the day. Table 1 further reports that the average time span between the current and last inspections is 184 days, and 38% of the inspections occur during lunch time (12-2pm). These two variables may affect inspection outcomes because restaurants are likely to adjust compliance effort according to when the next inspection is expected to occur and most restaurants are busy at lunch time. During the sample period, the DHR has adopted portable digital assistant (PDA) for inspectors. Since PDA usage has a significant impact on inspection outcome (Jin and Lee 2014a), we create a dummy variable indicating whether the study inspection uses PDA. We also control for the number of PDA inspections that the restaurant has experienced before the inspection, as the restaurant may adjust compliance according to PDA usage.

2.2 New-Repeat Difference

Despite universal training requirement for inspectors,[9] the first panel of Table 2 reveals significant difference between new and repeat inspectors. New inspectors on average find 9.36 violations, almost two additional violations (or 27% higher) than repeat inspectors. Such differences are found in all three categories (critical, risk-factor, and non-critical). The rest of Table 2 presents the new-repeat difference by the number of times that the previous inspector has inspected the restaurant (*Lrepeat*). The new-repeat difference remains highly significant, and the difference increases slightly with *Lrepeat*.

Figure 1 shows the average number of violations during a typical new-repeat history of a restaurant. It starts with the first new inspector for a restaurant, followed by four repetitions of this inspector and then the next round for the next new inspector. To construct the graphs, we identify the first "new" inspection for every restaurant in the sample ($1^{st}New$) and the repeat inspections conducted by the same inspector up to the fourth (*Repeat* 1, 2, 3 and 4). While all restaurants must have their first new inspection by definition, the number of restaurants differs over the order of repetition because restaurants have different numbers of repetitions by the first new inspector. Similarly, we identify the second, third, and fourth new inspections of restaurants, if they are observed in the sample, and the following repeat inspections. The graphs show the average number of violations for each new and repeat inspection.

We present two panels of graphs in Figure 1. Figure 1A plots the raw data of reported violations in total and three categories separately. Every point in Figure 1A reflects the sample mean, with a vertical line for the 95% confidence interval. In both total and category-specific graphs, we observe a spike in the first visit of a new inspector. The "new" spike reflects sharp changes on both sides: comparing with her predecessor, a new inspector reports more violations in her first visit; but comparing with herself, she reports fewer violations when she returns the second time. In all graphs, the tip of the "new" spikes increases cycle after cycle, probably because of overall time trends.

To filter out possible time trends and restaurant heterogeneity, we use the whole sample to regress violations (total and each category, separately) on year-month fixed effects,

---

[9]A newly hired inspector should receive at least 120 hours of training in her first year of employment. Also existing inspection staff receive a minimum of 20 hours of training each year. Each inspector is checked by the FDA every three years to ensure compliance with national standards. Each inspector is required to pass a certified food manager examination every five years.

restaurant age and restaurant fixed effects.[10] Based on the residuals from those regressions, we compute the average number of (residual) violations for each new-repeat cycle in Figure 1B. Spikes remain for new inspectors. The increasing trends of the tips, observed in the raw data in Figure 1A, disappear mostly. There are still some increasing trends for total, risk factor, and non-critical violations, but the increases are small in magnitude, compared to those in Figure 1A. In the meantime, the trend for critical violations is now decreasing. This is plausible if we believe that reported violations trigger compliance and some of the compliance is permanent. In contrast to the sharp decline in an inspector's first repetition, reported violations tend to decrease over repetitions, while these changes are small and insignificant.

To better quantify the new-repeat difference, we turn to the following regression analysis on the analysis sample.

$$
\begin{aligned}
y_{irt} = \ & \alpha_r + \alpha_t + \alpha_i + \beta_{new} \cdot NEW_{irt} + \beta_{Lrpt} \cdot Lrepeat_{i_{-1}rt} \\
& + \beta_{NewLrpt} \cdot NEW_{irt} \cdot Lrepeat_{i_{-1}rt} + \beta_x \cdot X_{irt} + \epsilon_{irt}
\end{aligned}
\tag{1}
$$

where the dependent variable is the number of violations (y_{irt}) for restaurant r by inspector i at time t. The key explanatory variables are a dummy indicating whether i is new to the restaurant (New), the number of times that the inspector of the last inspection has visited the restaurant ($Lrepeat$), and their interaction ($New \cdot Lrepeat$). If the current and previous inspectors are the same, New is equal to zero and $Lrepeat$ reflects the length of the relationship this inspector has had with the restaurant before this particular inspection. If the current inspector is new, $Lrepeat$ measures the repetitions of the previous inspector. Restaurant age, inspector experience, inspection time, and all the other control variables mentioned above are included (X_{rit}), in addition to restaurant fixed effects (α_r), year-month fixed effects (α_t), and inspector fixed effects (α_i). Table 3 reports regression results for critical, non-critical, and risk-factor violations, separately.

From Figure 1, we know that the "new" spike reflects two changes: an increase of reported violations on the left (when we compare a new inspector to her predecessor) and a decrease on the right (when we compare a new inspector's first visit to her second visit

---

[10]We also include a dummy indicator for missing value in restaurant age.

of the same restaurant). Our regression results capture both. Consistent with the left side of the "new" spike, Table 3 suggests that a new inspector reports more violations than the previous inspector, and this pattern is more conspicuous if the inspected restaurant has had a longer relationship with the previous inspector. For example, when $Lrepeat = 1$, a new inspector reports 0.28 more critical violations ($= 0.2657 + 0.0152 \times 1$) than the previous inspector coming back.[11] This difference is about 17.3% larger than the sample mean. Similar results are found for risk-factor and non-critical violations; the difference is 14.2% for risk-factor and 11.1% for non-critical violations. When $Lrepeat > 1$, the difference gets larger. For example, when $Lrepeat = 3$, a new inspector reports 0.31 more critical violations ($= 0.2657 + 0.0152 \times 3$) than the returning inspector.

Consistent with the right side of the "new" spike, regression results show that the number of violations declines most between the first and second visits of the same inspector. For instance, consider a restaurant that has seen four visits of its previous inspector before a new inspector comes in. For the first visit of the new inspector, we have $New = 1$ and $Lrepeat = 4$. For the second visit of the same inspector, we have $New = 0$ and $Lrepeat = 1$. For the third visit, $New = 0$ and $Lrepeat = 2$. Thus, the regression predicts that the new inspector reports 0.29 fewer critical violations from her first to second visits, but only 0.013 fewer from the second to third visits.[12] These numbers correspond to 17.8% and 0.8% of the sample mean, respectively. For risk factors and non-critical violations, the effects are similar.[13]

Put together, Table 3 confirms the cycling pattern we have seen in the raw data: the number of reported violations shoots up at every new inspection, declines sharply when the new inspector returns, followed by small and incremental declines until the next new inspector arrives. Note that our estimates do not imply that the same restaurant gets more and more citations from one new inspector to another. To see this, consider a restaurant that always receives a new inspector after the last inspector visited T times. According to Equation (1), the difference between the new inspections of the two adjacent inspectors

---

[11]More specifically, when $Lrepeat = 1$, the regression predicts β_{Lrpt} violations from the previous inspector versus $\beta_{new} + \beta_{Lrpt} + \beta_{NewLrpt}$ violations from a new inspector. Hence the new-repeat difference is $\beta_{new} + \beta_{NewLrpt} = 0.2657 + 0.0152 = 0.2809$.

[12]Everything else equal, the new inspector finds 0.2753 ($= 0.2657 + (-0.0128) \times 4 + 0.0152 \times 4$) plus a constant on her first visit, -0.0128 plus a constant on the second visit, and -0.0256 ($= -0.0128 \times 2$) plus a constant on the third visit.

[13]We estimated the reduced-form model for various subsamples, divided by inspector, restaurant, and inspection characteristics. The results are very consistent across the subsamples (Appendix Table A2). We also added a quadratic term of $Lrepeat$ in the regression but find little change in results.

is zero. Now let us compare another two new inspectors: one comes after the previous inspector visited T times and the other after the previous inspector visited only once. The difference between these two new inspectors' first inspections is $(T-1)(\beta_{Lrpt} + \beta_{NewLrpt})$. Because β_{Lrpt} and $\beta_{NewLrpt}$ are of opposite signs but similar magnitude, they largely cancel out each other. In other words, even if two adjacent new inspectors face different values of $Lrepeat$ from their predecessor, their spikes are likely to be similar.

3 Endogenous Inspector Rotation

This section examines whether non-random inspector assignment can explain the new-repeat difference. If the DHR assigns stringent inspectors to dirty restaurants, the new-repeat gap could simply reflect endogenous inspector rotation.

As a first check, in Figure 2, we plot the likelihood of a new inspector's arrival by the number of inspections that the last inspector has made for this restaurant ($Lrepeat$), after controlling for year, month and restaurant fixed effects. The curve increases steadily after $Lrepeat = 3$, suggesting that there is no obvious rotation cycle in inspector assignment, although the DHR seems reluctant to change inspectors twice consecutively within a restaurant.

A more relevant concern is that the DHR may assign new inspectors according to a restaurant's last inspection records. To check this, at the inspection level, we regress the indicator of the new inspector on the restaurant's total violations found in the previous inspection, in addition to restaurant fixed effects, year-month fixed effects, $Lrepeat$, and other attributes of the previous inspection. Results are reported in the first column of Table 4, showing that the number of total violations in the previous inspection has a small (-0.001) but statistically significant effect on the propensity of a new inspector's arrival. However, the negative sign of this coefficient estimate is actually opposite to the type of inspector selection we are concerned of. The second column of Table 4 breaks total violations in the previous inspection into critical, risk-factor, and non-critical categories. To our surprise, the number of critical violations found in the previous inspection is positively correlated with having a new inspector next time, but the number of risk-factor or non-critical violations shows an opposite effect. We do not have a good explanation for the results, but this pattern cannot explain the robust finding about the new-repeat difference

11

in our empirical analysis.

What other factors drive the DHR to send a new inspector to a restaurant? We were told that inspectors are typically assigned to territories near their residence in order to minimize transportation costs. While we do not know individual inspectors' residential locations, we do observe most inspectors' assignments clustered by no more than four zip codes.[14] Further analysis suggests two other important factors, namely inspector retirement and new hiring. Given the lack of inspector personnel data, we operationally define an inspector's retirement date as the date of her last appearance in our inspection data, and count an inspector as a new hire in the quarter of her first appearance in our data.

In the last two columns of Table 4, we include on the right hand side the proportion of "retired" inspectors and the proportion of "new hires" in the corresponding subdistrict and the previous quarter (there are 186 subdistricts in Florida). As we expect, both variables are highly correlated with whether a restaurant receives a new inspector. To the extent that inspector retirement and hiring are beyond the control of any single restaurant, we believe that the inspector assignment driven by subdivision-level retirement and hiring occurrences can be treated as exogenous.

These results suggest that retirement and new hires may serve as valid instrumental variables (IV) for New in Equation (1). In particular, we define two IVs, indicating whether there was any retirement or new hire respectively in the corresponding subdistrict and the previous quarter. Their interactions with $Lrepeat$ serve as instrumental variables for $New \cdot Lrepeat$. The IV results in the first panel of Table 5 show that the coefficients of New and $Lrepeat$ are qualitatively same as the OLS estimates in Table 3; new inspectors find significantly more violations than repeat inspectors for all three categories, and the new-repeat gap increases in the number of the previous inspector's visits.

As another test of the potential endogeneity of New, the second panel of Table 5 uses propensity score matching (PSM), where the propensity score prediction is based on Column 1 of Table 4. We conduct the PSM for each of the three violation categories, and by

---

[14]In particular, for each inspector-year, we list all the zip codes where an inspector conducted initial inspections and find that 80 to 90% of her assignments concentrate in four zipcodes. The average Herfindahl index of zip codes within each inspector in a given year is around 3,500. If we examine inspector assignment by zip-code, the Herfindahl index of inspectors within each zip-code is on average 7,000 to 9,000 each year. This suggests that a zip-code is typically served by only one or two inspectors. Moreover, comparing inspector assignment from one year to the next, we find that 57% of inspectors carry over at least 50% of her top-4 zip-code assignments to the next year. Should new inspectors be assigned mostly in an attempt to break a restaurant-inspector relationship, we should see weaker geographical concentration by inspectors and greater turnover of assignments between two consecutive years.

the exact value of *Lrepeat*. Specifically, for each value of *Lrepeat* from 1 to 7, we estimate the linear probability model where the dependent variable is New, the explanatory variables are the lagged variables of total violations and X_{rit}, and we also control for year-month and restaurant fixed effects. The predicted probability of new inspector's arrival is used as the matching variable. We used the nearest-neighbor matching without replacement with the 0.05 caliper (about 0.2 of the matching variable's standard deviation). We computed the bootstrapped 90% confidence intervals. The reported estimate is for the coefficient of New, along with its 90% bootstrapped confidence interval. The PSM estimates are qualitatively the same as the OLS results.

Lastly, should new inspector assignment be targeted (and restaurants know the assignment rule), we shall observe a larger new-repeat difference when the chance of getting a new inspector is low. To check this, we rerun the OLS regressions for a few subsamples, depending on whether there is high inspector employment turnover at the subdistrict level in the previous quarter, whether $Lrepeat \leq 3$, and whether the predicted probability of getting a new inspector is below 10%. In each of these subsamples, we obtain similar results as in the full-sample OLS regression. The key results are presented in Appendix Table A2, or refer to Jin and Lee (2014b).

In short, we conclude that inspector assignment is not random but largely driven by factors beyond the control of individual restaurants, and the realization of inspector identity is hard to predict by individual restaurants. Thus, the new-repeat gap in inspection outcomes cannot be explained by endogenous inspector rotation.

4 Two Explanations for the New-Repeat Difference

This section explores two explanations for the new-repeat gap, which we refer to as heterogenous inspector criteria and diminishing inspector attention. Both explanations imply time-varying effects beyond inspector fixed effects. Therefore, we propose a structural estimation model to address the time-varying effects empirically and check the robustness of our results in the reduced-form estimation.

4.1 Heterogenous Inspector Criteria

The heterogeneity of inspectors' tastes and stringency may explain the new-repeat gap observed in our data. Suppose that a restaurant has been inspected repeatedly by an inspector and the restaurant expects the same inspector to come back for the next inspection.[15] Also suppose that the inspector is strict on, for example, bathroom cleanliness but lenient on food temperature. Then the restaurant will devote most compliance efforts to bathroom cleanliness and pay little attention to food temperature. However, suppose that the next inspector turns out to be a new one and the new inspector focuses more on food temperature. Then the restaurant will receive many citations in food temperature but few citations in dirty bathroom. From the new inspector's report, the restaurant can infer the new inspector's type. That point onward, the restaurant expects the same inspector to return and targets compliance according to that inspector's type.

This simple example gives us a key prediction: the new-repeat gap is larger if the new inspector is of a different type from the previous inspector. If we can measure inspector type continuously, it implies that the further away the two inspectors' types, the greater the new-repeat gap. In our structural model, we attempt to measure the difference between two consecutive inspectors' types (unobservable to the econometrician) and estimate the effect of the difference on inspection outcomes.

The heterogeneity in inspector criteria does not always predict a positive gap between new and repeat inspectors, as targeted compliance tends to counter detection and reported violations are the product of these two countervailing forces. The mismatch of restaurant expectation and realized inspector identity could go either way, which may end up to be a pleasant or unpleasant surprise for the restaurant. Nevertheless, heterogeneity in inspector criteria could predict a positive new-repeat gap on average, if the surprise is more likely to be unpleasant (which depends on the distribution and assignment of inspectors), or if the exact cancellation between compliance and detection is asymmetric across the two types of surprise (which depend on the cost functions of compliance and detection). It is essentially an empirical question, which we quantify in counterfactual simulations.

---

[15]This is a simplifying assumption for this illustrative example. In reality, a sophisticated restaurant may adjust compliance to a statistical distribution of potential inspector identity. The crucial point is that there are different types of inspectors, and their different inspection criteria trigger different types of compliance reaction.

4.2 Diminishing Inspector Attention

The finding that an inspector reports fewer violations when she returns to the restaurant could also be explained by diminishing inspector attention. When a new inspector inspects a restaurant for the first time, she is not familiar with the restaurant and does not have any first-hand knowledge as to what areas are most vulnerable to violation.[16] Checking out every corner with a high level of attention, she may detect violations in many areas. However, returning to the restaurant, she is equipped with the knowledge she acquired before and may prefer to devote efforts to the areas she found problems previously. This effect could magnify itself when she returns again, because she pays less attention to the areas she recorded no problems multiple times. On the restaurant side, the restaurant learns more about the inspector's preference from previous visits and will most likely focus compliance on those areas. As a result, the same inspector will find fewer violations.

There could be various mechanisms behind diminishing inspector attention. It may be a rational decision by the inspector, because detection effort is costly and the possibility of finding a new violation in a previously compliant area is low. Within the same restaurant, diminishing inspector attention could also be driven by a behavioral bias, as humans often have confirmation bias and tend to avert self-contradiction (Jennings, Amabile and Ross 1982). Given the difficulty to distinguish rational and behavioral explanations, we are agnostic on the mechanisms behind diminishing inspector attention, but aim to separately identify its effect from the effect of heterogeneous inspector criteria.

4.3 Empirical Identification

To summarize, the hypothesis of heterogenous inspector criteria is built upon variations *across* inspectors, while the hypothesis of diminishing inspector attention is based on variation *within* the same inspector. The difference gives us empirical identification. Note that both hypotheses attempt to capture time-varying effects beyond inspector fixed effects. In particular, heterogenous inspector criteria emphasize the mismatch between restaurant compliance (to the last inspector's preference) and the preference of the new inspector. Hence it depends on the inherent difference between two adjacent inspectors. With enough data, we can derive these inherent preferences from individual inspectors' complete inspec-

---

[16]Even if she reads the previous inspector's report beforehand, the report contains mostly a check of yes or no to the 55 pre-printed violation items, with no details on the physical layout of the restaurant or how the layout relates to a particular violation.

tion records across all restaurants. Diminishing inspector attention focuses on differential detection efforts within the same inspector-restaurant pair. In this sense, diminishing inspector attention captures temporal variation within the same inspector-restaurant pair and does not depend on who the previous inspector is.

To identify the two effects separately, we extend Equation (1) to the following specification:

$$
\begin{aligned}
log(\lambda_{irct}) &= log(E(y_{irct})) \\
&= \underbrace{\beta_{new} \cdot NEW_{irt}}_{\text{new inspector's higher attention}} + \underbrace{\mu_{ic} + \beta_{hetero}(\mu_{ic} - \mu_{i_{-1}c})}_{\text{heterogeneous inspector criteria}} \quad\quad (2) \\
&+ \beta_{Lrpt} \cdot Lrepeat_{i_{-1}rt} + \beta_{newLrpt} \cdot NEW_{irt} \cdot Lrepeat_{i_{-1}rt} + \mu_{rc} + \mu_{ymc} + \zeta X_{rit}
\end{aligned}
$$

Denoting y_{irct} as the number of observed violations in category c for restaurant r by inspector i at time t, we assume y_{irct} follows a Poisson distribution with mean λ_{irct}.[17] New and $Lrepeat$ are defined as before. There are a rich set of fixed effects: μ_{rc} is the restaurant-category fixed effect; μ_{ic} is the inspector-category fixed effect for the current inspector; μ_{ymc} is the year-month-category fixed effect; $\mu_{i_{-1}c}$ the inspector-category fixed effects for the last inspector. Vector X_{irt} includes control variables such as restaurant age, inspector experience, whether the inspection is the inspector's first of the day, whether the inspection is during lunch hours, how many days since the last inspection, and so on.

This specification includes two main changes from Equation (1): first, since the number of violations is a count variable, we believe that a Poisson model is more appropriate than a linear model. In fact, running Equation (1) with Poisson or OLS gives us similar results. That is why we started with OLS in Section 2 and applied IV and PSM in Section 3. The second change is adding $\beta_{hetero}(\mu_{ic} - \mu_{i_{-1}c})$ on the right hand side. β_{hetero} captures the extent to which inspection outcomes depend on the inherent stringency and taste heterogeneity between the current and previous inspectors. Note that β_{hetero} does not apply unless the current inspector is new and differs from the previous inspector. Since we control for the inspector-category fixed effect of the current inspector, it captures the extent to which a restaurant complies with the idiosyncratic preference of the last inspector rather than that of the current inspector.

---

[17]Some potential econometric issues regarding the specification are discussed in Jin and Lee (2014b).

In comparison, β_{new} captures the effect of the new inspector's higher attention. β_{new} and β_{hetero} can be identified separately because the former compares new versus repeat inspectors no matter who the last inspector was, while the latter explores the identity and preference difference between the current and previous inspectors.

Neither heterogenous inspector criteria nor diminishing attention predict a clear effect of $Lrepeat$. Strictly speaking, if a single visit is enough to reveal the inspector's inherent taste, the restaurant has the incentive to comply to that taste for any subsequent visits. Similarly, an inspector may reduce her detection effort after the first visit and focus on the same problematic areas for every subsequent visit (which invites the restaurant to focus on the same areas for compliance). In reality, the $Lrepeat$-neutral prediction is violated if the learning of inspector taste is gradual, if the restaurant changes its expectation of inspector repetition, or if the inspector anticipates the restaurant's targeted compliance and therefore adjusts detection effort after each repeat visit. Given these ambiguities, we continue to include $Lrepeat$ and $New \cdot Lrepeat$ in the regression, but do not attempt to distinguish the two explanations from the estimates of β_{Lrpt} and $\beta_{newLrpt}$.

To summarize:

- The distribution of inspector-category fixed effects (μ_{ic}) captures inspector heterogeneity in their inherent stringency and taste (as well as the corresponding compliance response if restaurants can predict perfectly who will come next time).

- We predict $\beta_{new} > 0$ if new inspectors have higher attention than repeat inspectors, even if they have the same inherent taste and stringency.

- We predict $\beta_{hetero} > 0$, which captures the new inspector effect due to heterogenous inspector criteria when restaurants cannot perfectly predict the identity of the next inspector.

- The signs of β_{Lrpt} and $\beta_{newLrpt}$ are ambiguous, depending on how restaurants and inspectors adjust their behavior and expectation over repeat inspections.

4.4 Estimation Results

Table 6 reports the results from Equation (2).[18] We estimate the specification for critical, risk-factor, and non-critical violations separately, so that inspectors are allowed to differ in taste and stringency across the three categories.[19]

We have two kinds of coefficients for inspector heterogeneity, μ_{ic} and β_{hetero}. Figure 3 plots the histogram of the estimated μ_{ic} for critical, risk-factor, and non-critical violations, respectively. Using the most frequent inspector (in our data) as the benchmark, $\mu_{ic} = 0.5$ should be interpreted as inspector i on average reports $exp(0.5) - 1 = 65\%$ more violations than the benchmark inspector. According to this interpretation, Figure 3 shows enormous inspector heterogeneity in all three categories of violations.[20]

In the bottom part of Table 6, we show that, within each inspector, the estimated μ_{ic} has a correlation coefficient of 0.73 between critical and risk-factor violations, 0.44 between critical and non-critical, and 0.47 between risk-factor and non-critical. This suggests that inspector heterogeneity is driven by differences in both overall stringency and relative taste specific to each category.

Despite the substantial variation in μ_{ic}, β_{hetero} is small. The hypothesis of heterogenous inspector criteria predicts $\beta_{hetero} > 0$ because, if the last inspector is less stringent in category c than the current inspector ($\mu_{ic} - \mu_{i_{-1}c} > 0$) and the restaurant anticipates the last inspector to return, it should comply less and such lack of compliance will be reflected in more violations reported by the current more stringent inspector. This prediction holds for critical violations: for a new inspector who is 65% more stringent than the last (corresponding to $\mu_{ic} - \mu_{i_{-1}c} = 0.5$), she will find 1.9% more violations in addition to what she

---

[18]In the bottom panel of Table 6, we report the goodness of fit for each column. The calculation follows Cameron and Windmeijer (1996), which describes the log likelihood improvement from a constant-only model to our model as a fraction of the log likelihood improvement from the constant-only model to the perfect fit using the raw dataset itself. By definition, it is bounded within 0 and 1. The goodness of fit measure is 0.45 for critical violations, 0.55 for risk factors, and 0.60 for non-critical violations, suggesting that our model fits the raw data reasonably well. The bottom panel of Table 6 also reports the comparison of our predicted violations versus the actual violations in both mean and standard deviation. The mean is literally zero for all three columns, while the standard deviation is between 1.5 and 2.7. Most of the seemingly large standard deviation is driven by the fact that our predicted violations are continuous but actual violations are integers.

[19]All estimations maximize the conditional likelihood with restaurant fixed effects. According to HHG (1984) and Wooldridge (1999), restaurant fixed effects will drop out of the likelihood function conditional on total violations found at a restaurant. Hence, we only need to estimate μ_{ic}, β_{hetero}, β_{new}, β_{pda}, β_{Lpda1}, β_{Lpda0}, β_{new}, β_{Lrpt}, $\beta_{newLrpt}$ and ζ.

[20]After constructing robust standard errors according to Wooldridge (1999), we find that 22 inspectors have a stringency statistically different (at 95% confidence) from the reference inspector for critical violations, 8 inspectors are different for risk-factor violations, and 166 inspectors are different for non-critical violations.

would find if the restaurant fully anticipates her arrival. In comparison, we find that β_{hetero} is indifferent from zero for risk-factor violations and significantly negative for non-critical violations. The latter could occur if restaurants face time or personnel constraints in compliance so that more compliance in critical violations implies less compliance in non-critical violations.

After addressing inspector heterogeneity via μ_{ic} and $(\mu_{ic} - \mu_{i_{-1}c})$, β_{new} captures the effect of new inspectors' higher attention. Our estimates suggest that a new inspector reports 17.5% more critical violations, 14.6% more risk-factor violations, and 12.7% more non-critical violations compared to a repeat inspector coming back for her second visit to the restaurant. These percentages are computed by $(exp(\beta_{new} + \beta_{newLrpt} \times 1) - 1) \times 100\%$ because the lowest value of $Lrepeat$ is one. This large effect, combined with its increase with the length of the relationship between the last inspector and the restaurant (as indicated by the significant positive coefficient of $\beta_{newLrpt}$), suggests that new inspectors may have significantly higher attention in their first visit of a restaurant.

In comparison, the coefficient on the relationship alone, β_{Lrpt}, is small but negative, suggesting that among repeat inspectors one extra visit brings down the reported violations by 0.7-1.8%. This number could reflect a mixture of gradual compliance to the last inspector's stringency, gradual slack of repeat inspectors over time, or more compliance with the growing fear of a new inspector coming next time. The positive and significant estimate of $\beta_{newLrpt}$ (across all three categories) suggests that slackness of repeat inspectors is probably the most likely among these possibilities.

One may argue that the large effect of new inspectors' higher attention can be interpreted as restaurants catering to heterogenous inspector criteria within each of the three violation categories. To address this possibility, we rerun the model for four frequently cited items separately (items 2, 8, 22, and 23). In all four estimations, β_{new} is positive and statistically significant. The estimates mean that a new inspector, compared to a repeat inspector's second visit, reports 16.7% more violations in item 2, 10.8% more in 8, 3.8% more in 22, and 9.0% more in 23. In comparison, the inspector attention effect derived from the category-wise model is 14.6% more for risk-factor violations (items 2 and 8 are risk factors), and 12.7% more for non-critical violations (22 and 23 are non-critical). The results suggest that only a small fraction of the category-wise attention effects can be explained by heterogenous inspector criteria across specific items within the same category.

The results of some control variables may also suggest the tendency of diminishing inspector attention. To the extent that novice inspectors (whose experience is no more than 30 inspections) probably have the freshest memory of training and are still learning where to pay more attention, they find 18.4-27.1% more violations across all three groups of categories. Similarly, inexperienced inspectors whose experience is less than the median also report more detected violations, suggesting that diminishing inspector attention is applicable not only to a longer restaurant-inspector relationship but also to longer experience of inspectors. The sensitivity to inspector experience is consistent with what Macher et al. (2011) found regarding the inspection of pharmaceutical manufacturing, although we do not have detailed training and experience data as they do. Surprisingly, there is no obvious first-visit effect for the first inspection of the day, but inspectors find fewer violations throughout the day, probably due to fatigue. In addition, older restaurants tend to have more violations, as do inspections made during lunch time.

4.5 Counterfactual Simulations

Separating the effects of heterogeneous inspector criteria and diminishing attention sheds light on the design of the inspection program. In particular, if the new-repeat phenomenon is driven by heterogeneity of inspector criteria, the DHR may want to train all inspectors to ensure homogeneity. However, if diminishing inspector attention is the dominant reason, training only ensures that every inspector behaves the same as a new inspector, but it does not address progressive shirking after the first inspection. Another lever is inspector rotation, which addresses both effects, but its effectiveness depends on the relative importance of the two. Thus, using our empirical estimates, we quantitatively simulate the effects of random inspector assignment versus the effect of making inspector homogenous.

To better understand how much of the new-repeat gap is driven by heterogenous inspector criteria and diminishing attention, we perform three counterfactual simulations. The first simulation assigns inspectors randomly within each district. Specifically, we compute the frequency of each inspector in the raw data and use this as the weight of random assignment for that inspector. This way, the number of assignments for every inspector is similar to that of raw data, but the assignment itself is random. To minimize simulation error, we simulate random assignment for 100 times and compute the average predicted violations

for each inspection. The simulation results are presented in Table 7.[21] Many variables in the regression model are related to inspector assignment, but the greatest change is that over 80% of random assignments involve new inspectors as compared to 27% in the existing assignments. Consequently, random assignments on average yield 11.3% (non-critical) to 17.6% (critical) more detected violations. Further decomposition finds that most of the effects are driven by the large attention effect of new inspectors rather than inspector heterogeneity.

The second and third counterfactual simulations aim to compare the raw data with situations without any inspector heterogeneity. To do this, we keep the same inspector assignment as in the raw data but assume that every inspector is the same as either the average inspector or the most stringent inspector. Comparing the former to the raw data, inspector homogeneity leads to a lower mean and lower dispersion in the simulated distribution of detected violations. However, in the latter simulation, inspector homogeneity leads to a higher mean and higher dispersion of simulated violations. This is because more stringency leads to more detected violations in every inspection and this increase is greater in magnitude for dirtier restaurants due to the exponential functional form of the Poisson model. This also contributes to the increased dispersion of violations.

Given the simulated effects of inspector rotation and inspector training, which lever should the DHR adopt in its inspection program design? The results in Table 7 show that, compared to the raw data, the increase in the average number of violations by making every inspector the same as the most stringent one is much larger than the increase made by random assignment, 12 times larger for critical violations, 17 times for risk factors, and 13 times for non-critical violations. What does this mean for the DHR? Take critical violations as an example. Assuming that the DHR aims to detect as many critical violations as possible (which might minimize the public health risk) and it has to choose only one change, the results mean that, if the cost of training all inspectors into the most stringent type is 12 times larger than the cost of random assignment, random assignment should be a more efficient policy lever. Similar interpretation can be made on other violation categories, if we are willing to ignore correlations across categories.

---

[21]Not surprisingly, the actual violations are more dispersed but their medians are close to the predicted violations, especially in risk-factor and non-critical categories (because these two groups are less censored at zero).

5 Conclusion

Government inspections often involve repeated interactions between inspectors and inspectees. In this paper, we use restaurant hygiene inspections as an example to show that inspector assignment and repetition can have significant impacts on inspection outcomes. Our paper is motivated by the fact that new inspectors detect significantly more violations than repeat inspectors. After ruling out regulatory capture and endogenous inspector rotation, we attempt to explain the new-repeat gap by heterogeneous inspector criteria and diminishing inspector attention. We find that, after controlling for heterogenous inspector criteria, a new inspector finds 13-18% more violations than the second visit of a repeat inspector, likely due to the new inspector's higher level of attention. Given that the new-repeat gap is 27% in the raw data, our findings suggest that about 48-67% of the gap is explained by diminishing attention of repeat inspectors. The remaining gap can be attributed to restaurant heterogeneity, time trends, and heterogeneity of inspector tastes and stringency.

Using our estimates, we perform counterfactual simulations to highlight the effects of inspector training and rotation on regulatory outcomes. In reality, there should be various considerations for designing an inspection program. For example, more frequent rotation of inspectors should increase the transportation cost of inspectors, and therefore reduce the number of inspections that can be conducted per inspector per year. If the DHR aims to finish a fixed number of inspections per restaurant per year, it requires the DHR to hire more inspectors. Also it might be very expensive to create and run an effective training program. Putting all these complicated issues aside and under the simple premise that the DHR is to detect as many potential violations as possible, our simulation results suggest a tradeoff between strengthening the training program and randomly rotating inspectors. The two are comparable if training all inspectors into the most stringent type is about 12 to 17 times more costly than random rotation. Otherwise, one should be more cost-effective than the other depending on the cost ratio. There are various ways to improve inspections besides those we considered. Still we hope that our findings provide a useful guideline for policy makers.

References

Avenhans, Rudolf; von Stengel, Bernhard and Zamir, Shmuel (2002) "Inspection Games." Chapter 51, Handbook of Game Theory, Vol. 3, eds. R. J. Aumann and S. Hart, North-Holland, Amsterdam, 1947-1987.

Cameron, A. Colin and Frank A. G. Windmeijer (1996) "R-Squared Measures for Count Data Regression Models with Applications to Health-Care Utilization." Journal of Business & Economic Statistics, 14(2): 209-20.

Dranove, David and Ginger Zhe Jin (2010) "Quality Disclosure and Certification: Theory and Practice." Journal of Economic Literature 48(4): 935-63.

Dixit, Avinash (2002) "Incentives and Organizations in the Public Section: An Interpretative Review." Journal of Human Resources, 37(4): 696-727.

Feinstein, Jonathan (1989) "The Safety Regulation of U.S. Nuclear Power Plants: Violations, Inspections, and Abnormal Occurrences." Journal of Political Economy, 97: 115-154.

Feinstein, Jonathan (1991) "An Econometric Analysis of Income Tax Evasion and its Detection." RAND Journal of Economics, 22(1): 14-35.

Fudenberg, Drew and Jean Tirole (1991) Game Theory. Cambridge, MA: MIT Press. Hall.

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches (1984) "Econometric Models for Count Data with an Application to the Patents-R & D Relationship." Econometrica, 52(4): 909-938.

Jennings, Dennis L., Teresa M. Amabile, and Lee Ross (1982) "Informal Covariation Assessment: Data-based versus Theory-based Judgments." in Daniel Kahneman, Paul Slovic and Amos Tversky edited Judgment under Uncertainty: Heuristics and Biases. Cambridge: Cambridge U. Press, 1982, 211-30.

Jin, Ginger Zhe and Phillip Leslie (2003) "The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." Quarterly Journal of Economics, May 2003, 118(2), 409-51.

Jin, Ginger Zhe and Jungmin Lee (2014a) "Inspection technology, detection and compliance: evidence from Florida Restaurant inspections." RAND Journal of Economics, 45(4): 885-917.

Jin, Ginger Zhe and Jungmin Lee (2014b) "A Tale of Repetition: Lessons from Florida Restaurant Inspections." NBER Working Paper No. 20596.

Kleven, Henrik J., Martin B. Knudsen, Claus T. Kreiner, Soren Pedersen, and Emmanuel Saez (2010) "Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark." Econometrica 79(3): 651-692.

Laffont, Jean-Jacques and Jean Tirole (1993) A Theory of Incentives in Procurement and Regulation. Cambridge, Mass.: MIT Press.

Levine, David I. ; Michael W. Toffel and Matthew S. Johnson (2012) "Randomized Government Safety Inspections Reduce Worker Injuries with No Detectable Job Loss." Science, 336 (6083), 18 May 2012, pp. 907-911.

Macher, Jeffery T., John W. Mayo, and Jack A. Nickerson (2011) "Regulator Heterogeneity and Endogenous Efforts to Close the Information Asymmetry Gap." Journal of Law & Economics 54(1): 25-54.

Martimort, David (1999) "The Life Cycle of Regulatory Agencies: Dynamic Capture and Transaction Costs." Review of Economic Studies 66(4): 929-947.

Mookherjee, Dilip and Ivan Png (1989) "Optimal Auditing, Insurance and Redistribution." Quarterly Journal of Economics 104(2): 399-415.

Mookherjee, Dilip and Ivan Png (1995) "Corruptible Law Enforcers: How Should They Be Compensated?" Economic Journal 105(428): 145-159.

Office of Program Policy Analysis & Government Accountability (2005) "Division of Hotels and Restaurants Improves Operations But Not Meeting Inspection Goals." Report No. 05-51, An Office of the Florida Legislature.

Office of Program Policy Analysis & Government Accountability (2007) "Division of Hotels and Restaurants Improves Operations and Makes Progress in Meeting Inspection Goals." Report No. 07-41.

Prendergast, Canice (1999) "The Provision of Incentives in Firms." Journal of Economic Literature, 37(1): 7-63.

Prendergast, Canice (2007) "The Motivation and Bias of Bureaucrats." American Economic Review, 97(1): 180-196.

Short, Jodi L., Michael W. Toffel and Andrea R. Hugill (2016) "Monitoring Global Supply Chains." Strategic Management Journal, 37: 1878-1897.

Slemrod, Joel and Shlomo Yitzhaki (2002) "Tax Avoidance, Evasion and Administration." in the Handbook of Public Economics, No. 3, edited by Alan Auerbach and Martin Feldstein, Elsevier.

Tirole, Jean (1986) "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." Journal of Law, Economic and Organization, 2, 181-214.

Woodridge, Jeffery (1999) "Distribution-free Estimation of Some Nonlinear Panel Data Models." Journal of Econometrics 90: 77-97.

Figure 1. Average Violations by Inspection History

A. Raw Data
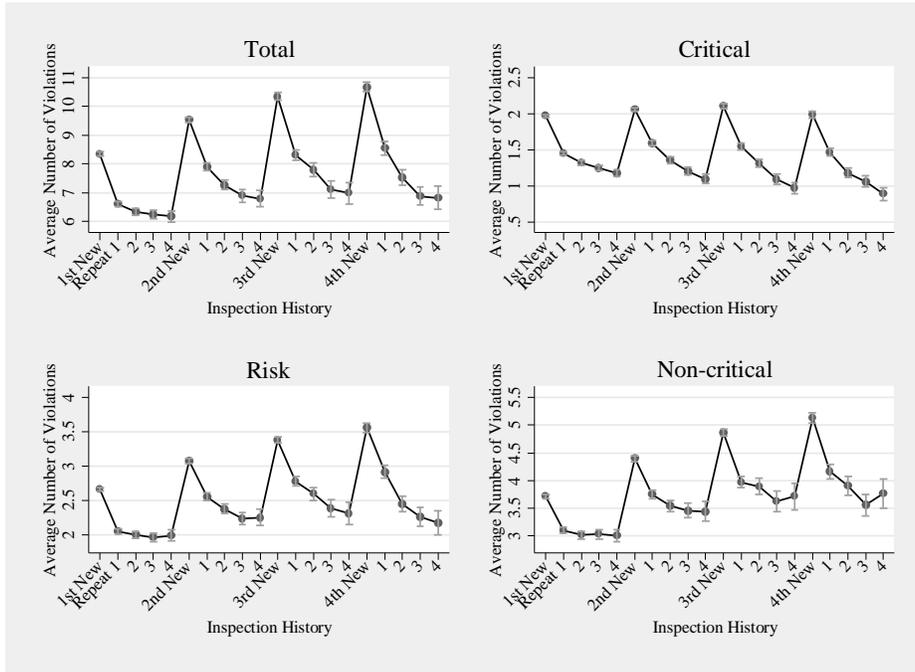
B. Regression-Adjusted Trends
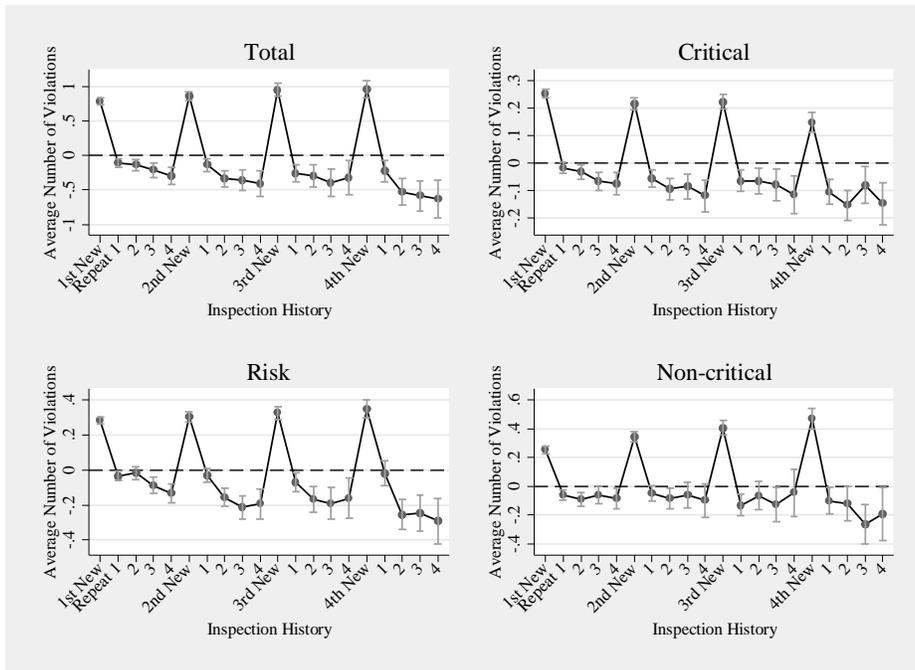
Notes: The graphs in Panel A show the trends of violations over the history of a restaurant. Every point reflects the sample mean, with a vertical line for the 95% confidence interval. Panel B shows the regression-adjusted trends after controlling for year-month fixed effects, restaurant age and restaurant fixed effects.

Figure 2. Probability of New Inspector Arrival by the Number of Inspections of the Previous Inspector
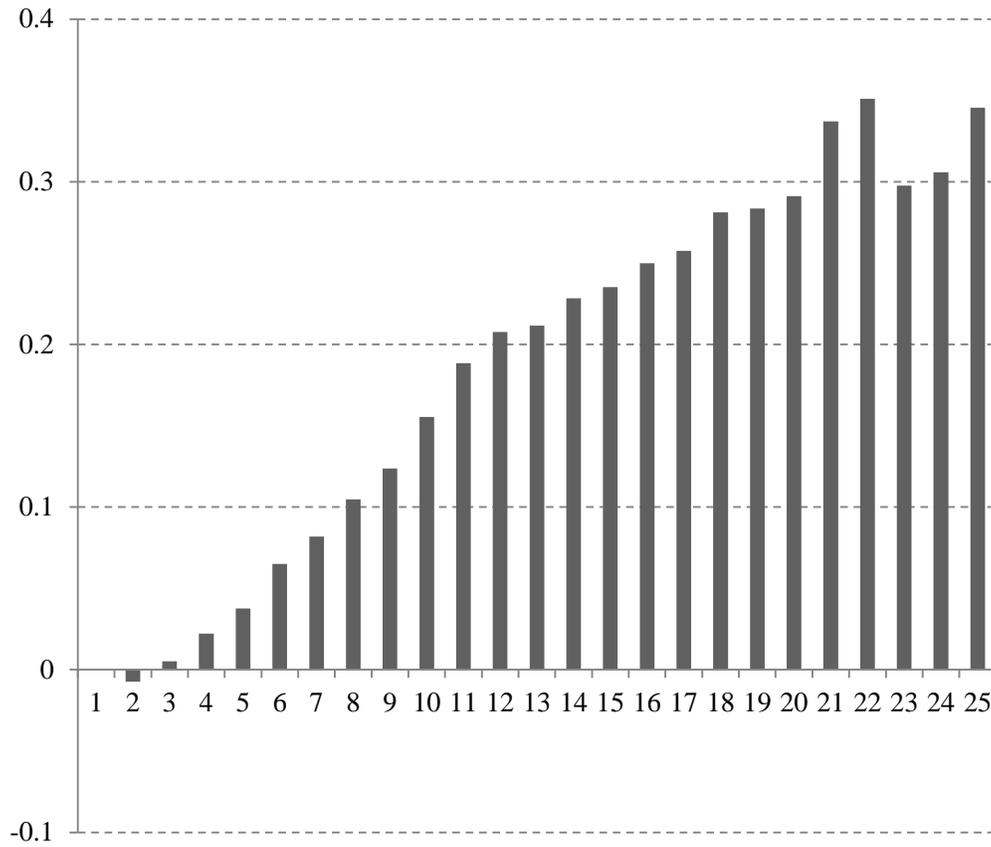
Notes: The linear probability model for a new inspector's arrival is estimated by using the specification of Column (1) of Table 4. The probability is relative to that when the number of inspection by previous inspector is one. The estimates when the number of inspections by previous inspector is greater than 25 are omitted. The sample mean number of inspections by previous inspector is 3.65, as shown in Table 1.

Figure 3. Histograms of Inspector Fixed Effects Estimated from Poisson Models
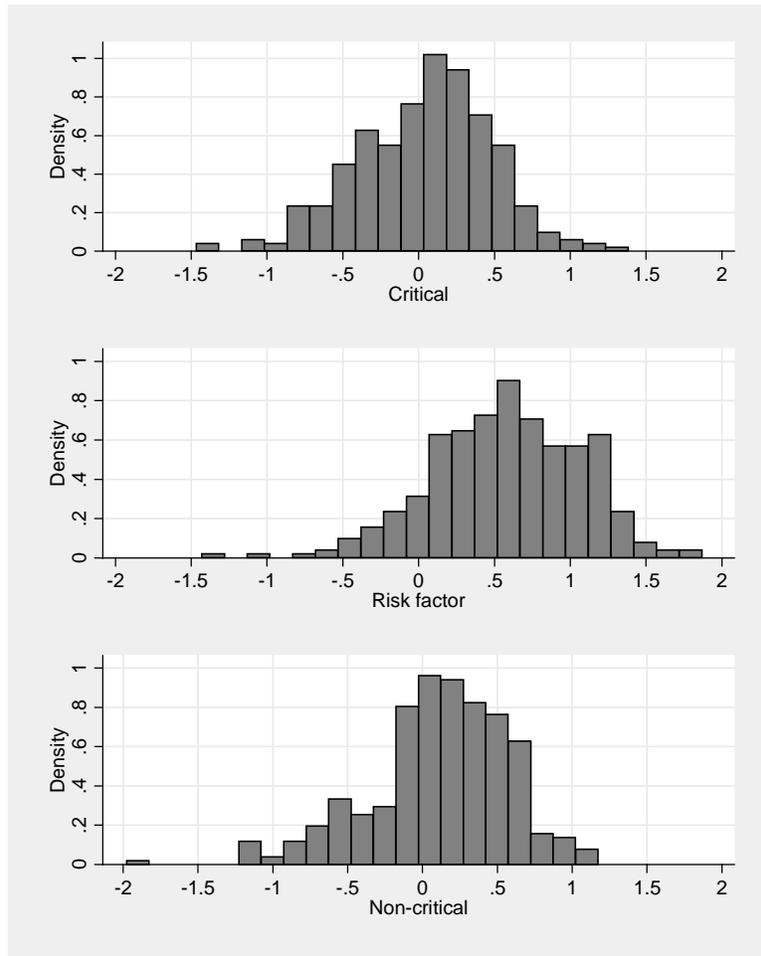
Notes: We use the most frequent inspector in our data as the benchmark. All the inspector fixed effects are relatively to this single inspector.

Table 1. Variables and Summary Statistics

	Mean	Std.Dev.	Min	Max
Number of critical violations	1.62	1.93	0	33
Number of risk factors	2.54	2.84	0	44
Number of non-critical violations	3.75	3.96	0	62
Number of all violations	7.90	7.16	0	111
New inspector	0.27	0.44	0	1
Number of inspections by previous inspector	3.65	3.07	1	38
PDA inspection	0.89	0.31	0	1
Number of previous PDA inspections	5.06	3.87	0	41
Whether the last inspection has used PDA	0.92	0.28	0	1
Restaurant age (years)*	4.10	2.74	0	14.5
Missing age	0.24	0.42	0	1
Complaint inspection	0.04	0.19	0	1
Licensing inspection	0.001	0.03	0	1
Inspector experience (# of inspections done before t)	1535	1114	0	5791
Inexperienced inspector (# of inspections less than the median)	0.44	0.5	0	1
Novice inspector (previous inspections $\leq$ 30)	0.01	0.11	0	1
First inspection of the day**	0.28	0.45	0	1
Number of previous inspections in the same day**	1.6	1.53	0	35
Missing inspection time	0.1	0.3	0	1
Time span from the last regular inspection (in days)	184	92	1	2004
Lunch time	0.38	0.49	0	1
Total # of restaurants	60,976			
Total # of inspectors	358			
Total # of inspections	426,831			
Total # of category-inspection (55 categories per inspection)	23,475,705			

* N = 326,461 obs. with age not missing. ** N = 384,198 obs. with inspection time not missing.

Table 2. New vs. Repeat Inspectors by the Number of Inspections by Previous Inspector

	(1)	(2)	(2) – (1)
	Repeat	New	New – Repeat (%)
A. All inspections			
All violations	7.38	9.36	27%***
	(6.63)	(8.27)	
Critical violations	1.47	2.01	37%***
	(1.79)	(2.24)	
Risk factors	2.36	3.05	29%***
	(2.66)	(3.23)	
Non-critical violations	3.54	4.31	22%***
	(3.74)	(4.45)	
B. By number of inspections by previous inspector			
B.1 # of inspections by previous inspector = 1 to 3			
Critical violations	1.49	1.96	32%***
	(1.81)	(2.19)	
Risk factors	2.32	2.91	25%***
	(2.68)	(3.13)	
Non-critical violations	3.44	4.14	20%***
	(3.77)	(4.32)	
B.2 # of inspections by previous inspector = 4 to 6			
Critical violations	1.43	2.13	49%***
	(1.73)	(2.32)	
Risk factors	2.35	3.35	43%***
	(2.63)	(3.40)	
Non-critical violations	3.57	4.63	30%***
	(3.70)	(4.63)	
B.3 # of inspections by previous inspector = 7 to 9			
Critical violations	1.48	2.25	52%***
	(1.76)	(2.42)	
Risk factors	2.47	3.54	43%***
	(2.66)	(3.53)	
Non-critical violations	3.79	5.03	33%***
	(3.72)	(4.96)	
B.4 # of inspections by previous inspector = 10 or more			
Critical violations	1.49	2.20	48%***
	(1.77)	(2.45)	
Risk factors	2.57	4.00	56%***
	(2.61)	(3.81)	
Non-critical violations	3.94	5.42	38%***
	(3.65)	(5.15)	

Notes: Standard deviations are presented in parentheses. *** 1% significance for inequality between new and repeat inspectors.

Table 3. OLS Results with Restaurant, Inspector, Year-Month Fixed Effects

	(1)	(2)	(3)
Violation type	Critical	Risk factor	Non-critical
Sample average of dependent variable	1.62	2.54	3.75
New inspector (β_{new})	0.2657***	0.3140***	0.3930***
	(0.0109)	(0.0148)	(0.0199)
# visits by the last inspector (β_{Lrpt})	-0.0128***	-0.0457***	-0.0186***
	(0.0018)	(0.0026)	(0.0035)
New ×# visits by the last inspector ($\beta_{newLrpt}$)	0.0152***	0.0467***	0.0243***
	(0.0030)	(0.0042)	(0.0055)
Restaurant age	0.0113***	-0.0067	0.0372***
	(0.0041)	(0.0063)	(0.0088)
Missing age	0.4948	-0.1683	-0.2660
	(0.6044)	(0.4612)	(0.8647)
Complaint inspection	-0.3553***	-0.4231***	-0.7132***
	(0.0160)	(0.0219)	(0.0320)
Licensing inspection	-0.4352***	-0.6340***	-1.2830***
	(0.0917)	(0.1222)	(0.1557)
Inexperienced inspector	0.0959***	-0.0268**	0.1443***
	(0.0097)	(0.0132)	(0.0185)
Novice inspector	0.4867***	0.4658***	1.0624***
	(0.0366)	(0.0465)	(0.0681)
Fatigue	-0.0529***	-0.0755***	-0.1125***
	(0.0024)	(0.0033)	(0.0044)
First of the Day	0.0344***	0.0398***	0.0787***
	(0.0089)	(0.0121)	(0.0164)
Missing fatigue	-0.1123***	-0.2238***	-0.2904***
	(0.0177)	(0.0240)	(0.0327)
PDA	0.1127***	0.3996***	0.5333***
	(0.0210)	(0.0292)	(0.0398)
# previous PDA inspections	-0.0864***	-0.0422***	-0.0105
	(0.0050)	(0.0075)	(0.0104)
# previous PDA inspections × PDA	-0.0233***	-0.0379***	-0.0854***
	(0.0035)	(0.0054)	(0.0074)
Time span	-0.0770***	-0.1229***	-0.1019***
	(0.0059)	(0.0078)	(0.0106)
Lunch	0.0401***	0.0485***	0.0228**
	(0.0064)	(0.0088)	(0.0115)
Restaurant FE	Yes	Yes	Yes
Inspector FE	Yes	Yes	Yes
Month-by-Year FE	Yes	Yes	Yes
Adj. R squared	0.1306	0.1579	0.1494
Number of observations	426,831	426,831	426,831

Notes: Robust standard errors are presented in parentheses. *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

Table 4. Probability of New Inspector Arrival

	(1)	(2)	(3)	(4)
Previous inspection's characteristics				
Total violations	-0.0014***		-0.0015***	
	(0.0002)		(0.0002)	
Critical violations		0.0019***		0.0016***
		(0.0005)		(0.0005)
Risk factor violations		-0.0012***		-0.0012***
		(0.0004)		(0.0004)
Noncritical violations		-0.0030***		-0.0030***
		(0.0003)		(0.0003)
Retired inspectors in previous quarter			1.0815***	1.0805***
			(0.0212)	(0.0212)
New hires in previous quarter			0.3643***	0.3632***
			(0.0169)	(0.0169)
Restaurant age	-0.0279***	-0.0278***	-0.0275***	-0.0274***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Missing age	0.0900	0.0885	0.0930	0.0916
	(0.1002)	(0.1003)	(0.1000)	(0.1001)
Complaint inspection	0.0205***	0.0207***	0.0211***	0.0214***
	(0.0041)	(0.0041)	(0.0041)	(0.0041)
Licensing inspection	-0.0057	-0.0060	-0.0091	-0.0093
	(0.0278)	(0.0278)	(0.0282)	(0.0282)
Inexperienced	-0.0666***	-0.0665***	-0.0639***	-0.0639***
	(0.0022)	(0.0022)	(0.0022)	(0.0022)
Novice	-0.0215***	-0.0219***	-0.0320***	-0.0323***
	(0.0080)	(0.0080)	(0.0079)	(0.0079)
FE of # visits by the last inspector	Yes	Yes	Yes	Yes
Restaurant FE	Yes	Yes	Yes	Yes
Month-by-Year FE	Yes	Yes	Yes	Yes
R squared	0.2732	0.2733	0.2837	0.2838
Number of observations	365,855	365,855	365,304	365,304

Notes: Linear probability models are estimated. Robust standard errors are presented in parentheses. *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

Table 5. Instrumental Variable and Propensity Score Matching Results by Violation Type

Violation type	(1) Critical	(2) Risk factor	(3) Non-critical
A. Instrumental variable results			
New inspector (β_{new})	0.2724***	0.2942***	1.1414***
	(0.0792)	(0.1049)	(0.1426)
# visits by the last inspector (β_{Lrpt})	-0.0243***	-0.0727***	-0.0579***
	(0.0046)	(0.0064)	(0.0082)
New×# visits by the last inspector ($\beta_{newLrpt}$)	0.0151	0.0764***	0.0165
	(0.0180)	(0.0246)	(0.0315)
Control variables	Yes	Yes	Yes
Restaurant FE	Yes	Yes	Yes
Inspector FE	Yes	Yes	Yes
Month-by-Year FE	Yes	Yes	Yes
Adj. R squared	0.0613	0.0814	0.0473
Number of observations	418,038	418,038	418,038
B. Propensity score matching results			
# visits by the last inspector = 1	0.2174	0.3410	0.4650
	[0.1713, 0.2440]	[0.2767, 0.3846]	[0.3758, 0.5310]
# visits by the last inspector = 2	0.4227	0.5609	0.6125
	[0.3623, 0.4576]	[0.4878, 0.6300]	[0.5182, 0.7296]
# visits by the last inspector = 3	0.5307	0.6461	0.7264
	[0.4615, 0.5970]	[0.5414, 0.7541]	[0.5723, 0.8687]
# visits by the last inspector = 4	0.5756	0.7096	0.5305
	[0.4905, 0.7169]	[0.5704, 0.8725]	[0.2545, 0.7417]
# visits by the last inspector = 5	0.6400	0.6622	1.0307
	[0.4284, 0.7908]	[0.4698, 1.0024]	[0.7300, 1.3983]
# visits by the last inspector = 6	0.4746	1.4873	1.0424
	[0.1638, 0.8142]	[0.8850, 1.9788]	[0.4304, 1.7941]
# visits by the last inspector = 7	0.6429	1.6964	1.3214
	[-0.0956, 1.3846]	[0.5686, 3.2679]	[0.1373, 3.2083]

Notes: For IV, robust standard errors are presented in parentheses. *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level. For PSM, 90% bootstrapped confidence intervals are presented in squared brackets. The propensity score is estimated based on Column (1) of Table 4.

Table 6. Poisson Estimation Results

	(1)	(2)	(3)
Dependent variable	Critical	Risk factor	Non-critical
Sample average of dependent variable	1.62	2.54	3.75
New inspector (β_{new})	0.1508***	0.1224***	0.1163***
	(0.0075)	(0.0043)	(0.0059)
# visits by the last inspector (β_{Lrpt})	-0.0180***	-0.0189***	-0.0068***
	(0.0012)	(0.0007)	(0.0011)
New×# visits by the last inspector ($\beta_{newLrpt}$)	0.0102***	0.0141***	0.0030**
	(0.0018)	(0.0011)	(0.0015)
Restaurant age	0.0074**	-0.0059***	0.0149***
	(0.0032)	(0.0020)	(0.0026)
Missing age	0.3614	0.1532	0.1273
	(0.3212)	(0.1206)	(0.2570)
Complaint inspection	-0.2017***	-0.1491***	-0.1531***
	(0.0096)	(0.0053)	(0.0077)
Licensing inspection	-0.3082***	-0.2978***	-0.4065***
	(0.0625)	(0.0479)	(0.0572)
Inexperienced inspector	0.0593***	-0.0082*	0.0437***
	(0.0080)	(0.0048)	(0.0065)
Novice inspector	0.2377***	0.1686***	0.2396***
	(0.0162)	(0.0098)	(0.0137)
# inspections before the current inspection	-0.0391***	-0.0358***	-0.0362***
	(0.0019)	(0.0011)	(0.0015)
First inspection of the day	0.0030	0.0046	0.0085*
	(0.0057)	(0.0034)	(0.0045)
Missing inspection time	-0.0830***	-0.1191***	-0.0938***
	(0.0113)	(0.0074)	(0.0102)
Time span	-0.0626***	-0.0523***	-0.0365***
	(0.0037)	(0.0022)	(0.0030)
Lunch	0.0250***	0.0211***	0.0081**
	(0.0041)	(0.0025)	(0.0032)
Inspector heterogeneity (β_{hetero})	0.0379***	0.0009	-0.0547*
	(0.0142)	(0.0114)	(0.0309)
Restaurant FE	Yes	Yes	Yes
Inspector FE	Yes	Yes	Yes
Month-by-Year FE	Yes	Yes	Yes
Log Likelihood	-1,405,660	-2,219,280	-3,350,400
Number of observations	426,831	426,831	426,831

| Correlations between Inspector FEs (μ_{ic})			
Risk factor	0.727		
Non-critical	0.442	0.474	
Goodness of fit (pseudo R-squared)	0.451	0.553	0.605
Mean of (Predicted Y - Y)	0.000	0.000	0.000
Std. dev. of (Predicted Y - Y)	1.468	1.974	2.672

Notes: The three columns are estimated separately. All columns control for whether this and previous inspections are paperless (PDA) or not. The full results on the PDA effects can be obtained from the authors. Also please refer to Jin and Lee (2014a). Robust standard errors are in parentheses. *** 1% significance; ** 5% significance; * 10% significance.

Table 7. Counterfactual Simulations

Simulations by Category and Scenario	Mean	Std. dev.	1st quartile	2nd quartile	3rd quartile
Critical					
Raw data	1.616	1.931	0.000	1.000	2.000
Predicted	1.616	1.334	0.694	1.293	2.161
Random weighted assignment	1.900	1.388	0.927	1.608	2.542
Every inspector same as average	1.486	1.179	0.674	1.215	1.989
Every inspector same as the most stringent	5.024	3.987	2.279	4.106	6.725
Risk factor					
Raw data	2.542	2.839	0.000	2.000	4.000
Predicted	2.542	2.175	1.013	1.975	3.459
Random weighted assignment	2.919	2.174	1.324	2.442	4.036
Every inspector same as average	2.527	2.178	1.020	1.963	3.428
Every inspector same as the most stringent	8.808	7.591	3.554	6.840	11.948
Non-critical					
Raw data	3.745	3.958	1.000	3.000	5.000
Predicted	3.745	3.090	1.535	2.998	5.117
Random weighted assignment	4.170	3.011	1.958	3.583	5.752
Every inspector same as average	3.498	2.737	1.539	2.911	4.789
Every inspector same as the most stringent	9.384	7.342	4.129	7.809	12.847

Notes: Counterfactual simulations are conducted by using the coefficient estimates from the Poisson model (Table 6).

Appendix Figure 1. Inspector Heterogeneity: Individual Fixed Effects
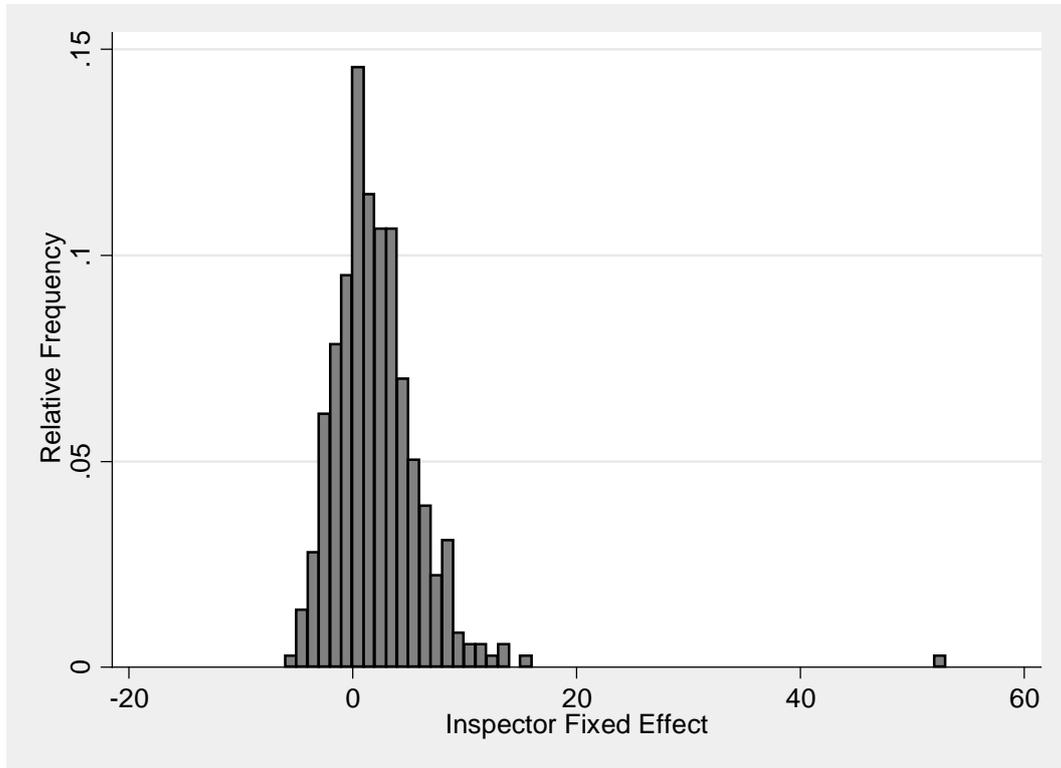
Notes: Inspector-specific fixed effects are estimated from regressing total violations on inspection month and fiscal year fixed effects as well as restaurant-specific fixed effects. There are 358 inspectors in total. By adding inspector-specific fixed effects, the R squared increases from 0.034 to 0.131. The omitted inspector's ID = 59. The outlying inspector (ID = 49233) has only 12 inspections in the final sample for all different restaurants. The average number of inspections per inspector is 1,192, the median 952 and the maximum 4,275.

Appendix Figure 2. Inspector Heterogeneity by Item

Item 22: Food contact surfaces of equipment
and utensils clean

Item 32: Restrooms with self-closing doors,
fixtures operate properly, facility clean,
supplied with hand-soap, disposable towels or
hand drying devices, tissue, covered waste
receptacles

Item 8: Food protection, cross-
contamination
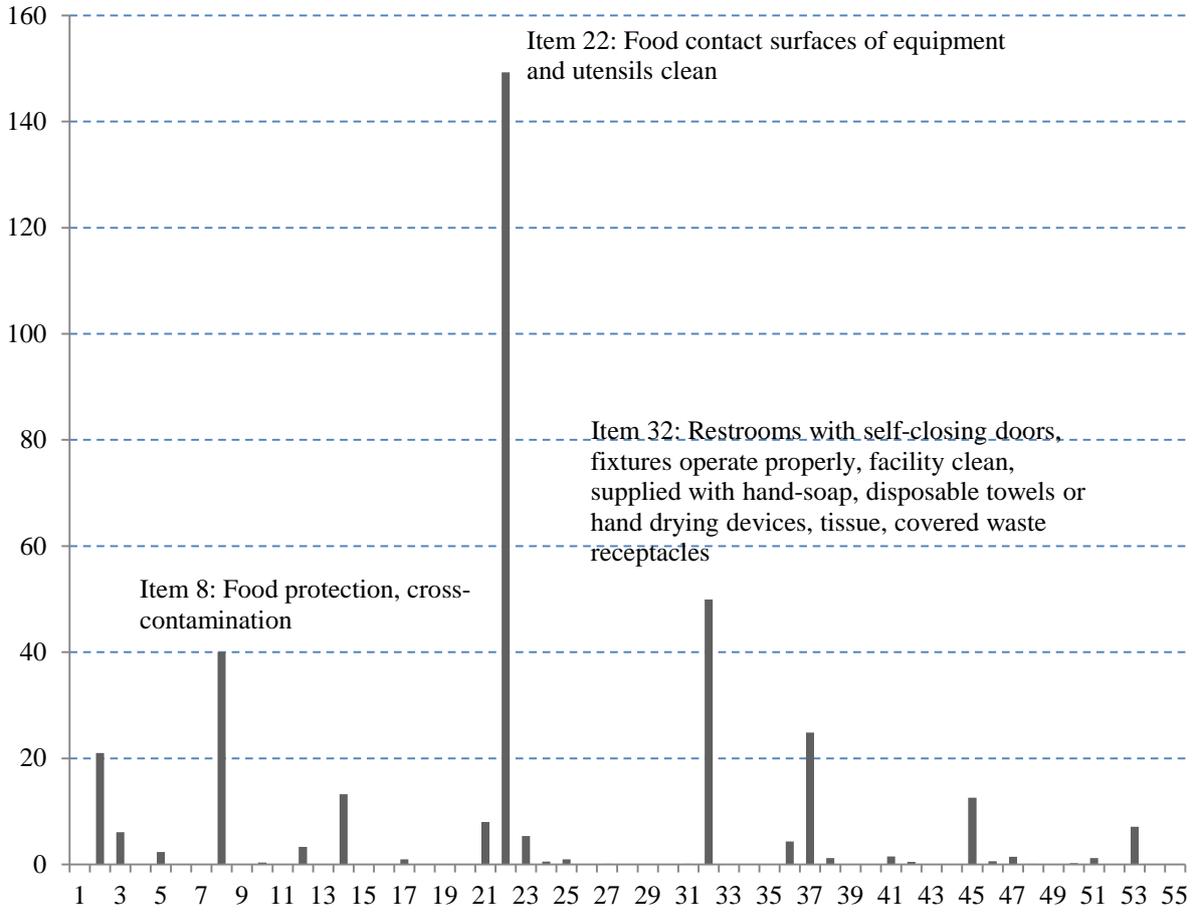
Notes: Inspector-item fixed effects are estimated by regressing item-specific violations on restaurant fixed effects, month and fiscal year fixed effects. The graph shows the frequency of inspectors at each item for which they are most likely to detect violations. There are 358 unique inspectors in the sample. See Appendix A for all items.

Appendix Table A1. Inspection Items

Item Number	Details	Category
1	Approved source	R
2	Original container: properly labeled, date marking, consumer advisory	R
3	Food Out of Temperature	R
4	Facilities to maintain product temperature	C
5	Thermometers provided and conspicuously placed	C
6	Potentially hazardous food properly thawed	C
7	Unwrapped or potentially hazardous food not re-served	R
8	Food protection, cross-contamination	R
9	Foods handled with minimum contact	R
10	In use food dispensing utensils properly stored	N
11	Personnel with infections restricted	R
12	Hands washed and clean, good hygienic practices, eating/drinking/smoking	R
13	Clean clothes, hair restraints	N
14	Food contact surfaces designed, constructed, maintained, installed, located	N
15	Non-food contact surfaces designed, constructed, maintained, installed, located	N
16	Dishwashing facilities designed, constructed, operated	C
17	Thermometers, gauges, test kits provided	C
18	Pre-flushed, scraped, soaked	N
19	Wash, rinse water clean, proper temperature	N
20	Sanitizing concentration or temperature	C
21	Wiping cloths clean, used properly, stored	N
22	Food contact surfaces of equipment and utensils clean	N
23	Non-food contact surfaces clean	N
24	Storage/handling of clean equipment, utensils	N
25	Single service items properly stored, handled, dispensed	N
26	Single service articles not re-used	N
27	Water source safe, hot and cold under pressure	C
28	Sewage and wastewater disposed properly	C
29	Plumbing installed and maintained	N
30	Cross-connection, back siphonage, backflow	C
31	Toilet and hand-washing facilities, number, convenient, designed, installed	C
32	Restrooms with self-closing doors, fixtures operate properly, facility clean, supplied with hand-soap, disposable towels or hand drying devices, tissue, covered waste receptacles	R
33	Containers covered, adequate number, insect and rodent proof, emptied at proper intervals, clean	N
34	Outside storage area clean, enclosure properly constructed	N
35	Presence of insects/rodents. Animals prohibited. Outer openings protected from insects, rodent proof	C
36	Floors properly constructed, clean, drained, coved	N
37	Walls, ceilings, and attached equipment, constructed, clean	N

38	Lighting provided as required. Fixtures shielded	N
39	Rooms and equipment - vented as required	N
40	Employee lockers provided and used, clean	N
41	Toxic items properly stored, labeled and used properly	R
42	Premises maintained, free of litter, unnecessary articles. Cleaning and maintenance equipment properly stored. Kitchen restricted to authorized personnel	N
43	Complete separation from living/sleeping area, laundry	N
44	Clean and soiled linen segregated and properly stored	N
45	Fire extinguishers - proper and sufficient	C
46	Exiting system - adequate, good repair	C
47	Electrical wiring - adequate, good repair	C
48	Gas appliances - properly installed, maintained	C
49	Flammable/combustible materials - properly stored	C
50	Current license properly displayed	C
51	Other conditions sanitary and safe operation	N
52	False/misleading statements published or advertised relating to food/beverage	N
53	Food management certification valid / Employee training verification	R
54	Florida Clean Indoor Air Act	N
55	Automatic Gratuity Notice	N

Notes: In the third column, C represents critical violation, R risk factors, and N non-critical violation. There are 17 critical violation items, 11 risk factor items, and 27 non-critical items. On the actual inspection form, there is an extra column that further divides items into sub-items by classification code. All the information contained in this table can be downloaded from the website of the Division of Hotels and Restaurants of Florida Department of Business & Professional Regulation (www.myfloridalicense.com/dbpr/hr/index.html).

Appendix Table A2. Robustness Checks

Violation type	(1) Critical	(2) Risk factor	(3) Non-critical
A. High inspector turnover at previous quarter			
New inspector	0.1480***	0.2265***	0.3208***
	(0.0406)	(0.0524)	(0.0657)
# visits by the last inspector	-0.0592***	-0.0320***	-0.0320**
	(0.0081)	(0.0107)	(0.0136)
New × # visits by the last inspector	0.0529***	0.0307*	-0.0010
	(0.0120)	(0.0159)	(0.0183)
R squared	0.6984	0.7192	0.7451
Number of observations	65,616	65,616	65,616
B. # visits by the last inspector ≤ 3			
New inspector	0.1850***	0.2593***	0.3491***
	(0.0198)	(0.0264)	(0.0355)
# visits by the last inspector	-0.0735***	-0.1335***	-0.0726***
	(0.0052)	(0.0071)	(0.0096)
New × # visits by the last inspector	0.0557***	0.0841***	0.0558***
	(0.0107)	(0.0143)	(0.0192)
R squared	0.4935	0.5728	0.6030
Number of observations	262,333	262,333	262,333
C. Prob(new inspector) < 10%			
New inspector	0.2140***	0.2084**	0.4848***
	(0.0662)	(0.0992)	(0.1336)
# visits by the last inspector	-0.0261***	-0.0421***	-0.0236***
	(0.0039)	(0.0050)	(0.0069)
New × # visits by the last inspector	0.0279**	0.0538**	0.0387
	(0.0132)	(0.0221)	(0.0263)
R squared	0.5538	0.6513	0.6808
Number of observations	112,412	112,412	112,412

Notes: All regressions include all control variables and fixed effects of our main specification in Table 3. Robust standard errors are presented in parentheses. Panel A restricts the sample to those where more than 10% of inspectors in the same subdistrict retired or more than 10% of inspectors in the same subdistrict were newly hired in the previous quarter. For Panel C, the probability of new inspector's arrival is predicted from Column (1) of Table 4. *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.