American Economic Journal: Economic Policy 2019, 11(4): 429–443
https://doi.org/10.1257/pol.20180543

New Evidence on Information Disclosure through Restaurant Hygiene Grading: Reply†

By Ginger Zhe Jin and Phillip Leslie*

Ho, Ashwood, and Handan-Nader (forthcoming) replicates table VI of Jin and Leslie (2003) but questions its research design. Robustness checks support our original conclusion—foodborne hospitalizations, as defined in JL, declined in Los Angeles County (LA) relative to the rest of California (CA) after LA adopted restaurant hygiene grade cards in 1998. More precisely, the decline in LA is pronounced against central and Northern CA, but insignificant when compared with the rest of Southern CA. One possible explanation is that the LA regulation has generated spillovers in Southern CA. (*JEL* D83, H75, I12, I18, L83, L88)

We are glad that Ho, Ashwood, and Handan-Nader (forthcoming, henceforth refered to as HAH) is able to replicate Jin and Leslie (2003, henceforth refered to as JL) table VI, despite using different data. In particular, they use California hospitalization records by quarter and five-digit zip code, while our data are the same records but by month and three-digit zip code. When we requested hospitalization data from California's Office of Statewide Health Planning and Development (OSHPD) in the early 2000s, OSHPD required us to choose between the two versions. We chose the monthly version, as our other data (restaurant inspection and revenue) are from a short time window (1995–1999), and we know the exact dates by which the LA county (and cities within) adopted their disclosure regulation. Monthly data give us more observations and allow a better match to the timing of regulation. As a trade-off, we had to accept a cruder definition of geography, as HAH points out. We believe the noise introduced by the three-digit zip code is not crucial because the zip code is tied to the patient's residence, not necessarily the place he/she got the tainted food. Since many people commute across zip codes in Southern California and we do not know the source of the tainted food (which could be restaurant, grocery store, party, or home), more precision on the patient's residence may not be very helpful. We are glad that this belief is confirmed by HAH: they are able

* Jin: Department of Economics, University of Maryland, College Park, MD 20742 (email: jin@econ.umd.edu); Leslie: Department of Economics, University of Maryland, College Park, MD 20742 (email: pleslie@fastmail.com). Kate Ho was coeditor for this article. We are grateful to officials from the Los Angeles County health department for providing us data for the original study and for commenting on an early draft of this note. Three anonymous referees have helped us improve the writing and statistical analysis. All errors are ours. This note responds to Ho, Ashwood, and Handan-Nader's comments on our 2003 publication at the *Quarterly Journal of Economics*, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards."

† Go to https://doi.org/10.1257/pol.20180543 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

to replicate our results using the other version of the data, though their estimate on the key coefficient has a larger magnitude (~ -0.3) than ours (~ -0.2).

After replication, HAH questions our research design, namely the variables we include on the right-hand side of the econometric specification, the control group we used in our estimation, and the way we constructed our dependent variable. We will address these concerns one by one and then conclude.

I. Missing an Interaction Term?

The specification we used in JL is

$$(1) \qquad \ln(a_{ijt}) = \alpha_{ij} + \tau_t + \beta_1 m_{it} + \beta_2 v_{it} + \gamma_1 food_j m_{it} + \gamma_2 food_j v_{it} + \varepsilon_{ijt},$$

where a_{ijt} is one plus the number of hospital admissions for digestive disorders[1] in three-digit zip code i, of type j (either food related or non-food related), in month t. The variable $m \in [0,1]$ measures the share of the three-digit zip code that is subject to mandatory disclosure of Los Angeles County (LA) restaurant hygiene grade cards in that month, and $v \in [0,1]$ measures the share of the zip code subject to voluntary disclosure in that month. The dummy variable *food* equals one for food-related digestive disorders and zero otherwise. The equation includes zip code-illness-type fixed effects (α_{ij}) as well as year fixed effects and month fixed effects (τ_t). Errors are clustered by zip code-illness type.

As articulated in JL, "identification is based on time-series and cross-sectional variations provided by the presence of two control groups: California outside of LA and nonfood-related digestive disorders" (page 439). The definition of $food_j$ does not vary by time hence the stand-alone term *food* is absorbed in the zip code-illness-type fixed effects (α_{ij}). By this specification, Table 1 shows that we can express the predicted log admissions by eight groups.

After we control for α_{ij} and τ_t, it is clear that β_1 is identified by comparing group 6 to the average of groups 1, 4, 7, and 8 after accounting for different levels of fixed effects (α_{ij}); β_2 is identified by comparing group 5 to the average of groups 1, 4, 7, and 8; γ_1 is identified by comparing group 3 to group 6; and γ_2 is identified by comparing group 2 to group 5. In other words, β_1 and β_2 capture the general difference between LA and the rest of CA in the time with hygiene grade cards, and γ_1 and γ_2 capture the extra difference between LA and the rest of CA for foodborne hospitalizations on top of β_1 and β_2.

We never refer to the JL specification as triple difference-in-differences (DID), but we agree that the JL specification is more restrictive than triple DID. In particular, it assumes all groups follow the same time trend (τ_t) in the absence of the LA regulation. This assumption can be relaxed in multiple ways: one is allowing food and non-food hospitalizations to differ after 1998, which is equivalent to adding $food_j \cdot post1998_t$ as HAH suggests; the other is estimating the model for food

[1] JL used one plus hospitalization counts to make sure that log is defined for every observation. We continue this practice for consistency.

Table 1—Identification Implied by JL Specification
Specification: $\ln(a_{ijt}) = \alpha_{ij} + \tau_t + \beta_1 m_{it} + \beta_2 v_{it} + \gamma_1 food_{ijt} m_{it} + \gamma_2 food_{ijt} v_{it} + \varepsilon_{ijt}$

Group	Three-digit zip codes (i)	Food (j)	Regulation (m and v)	Predicted $\ln(a_{ijt})^{\dagger}$
1	LA zip codes	Food related	0	$\alpha_{ij} + \tau_t$
2	LA zip codes	Food related	$m = 0, v = 1$	$\alpha_{ij} + \tau_t + \beta_2 + \gamma_2$
3	LA zip codes	Food related	$m = 1, v = 0$	$\alpha_{ij} + \tau_t + \beta_1 + \gamma_1$
4	LA zip codes	Non-food related	0	$\alpha_{ij} + \tau_t$
5	LA zip codes	Non-food related	$m = 0, v = 1$	$\alpha_{ij} + \tau_t + \beta_2$
6	LA zip codes	Non-food related	$m = 1, v = 0$	$\alpha_{ij} + \tau_t + \beta_1$
7	Rest of CA zip codes	Food related	Always 0	$\alpha_{ij} + \tau_t$
8	Rest of CA zip codes	Non-food related	Always 0	$\alpha_{ij} + \tau_t$

†Zip code-illness-type fixed effects (α_{ij}) vary across groups because each group has its own i and j.

Table 2—Replicate JL Table VI with Alternative Variables and Alternative Samples

	Replicate JL table VI	Use HAH specification	Food-borne only	Non-food only	LA only
LA mandatory	0.0271	−0.0257	(absorbed)	−0.0205	0.1799
	(0.0239)	(0.0246)		(0.0251)	(0.1228)
LA voluntary	0.0736	0.0188	(absorbed)	0.0145	0.2396
	(0.0238)	(0.0237)		(0.0239)	(0.1238)
Foodborne × post1998		−0.1007			(absorbed)
		(0.0332)			
Foodborne × LA mandatory	−0.2217	−0.1161	−0.1471	(dropped)	−0.1654
	(0.0431)	(0.0545)	(0.0484)		(0.0327)
Foodborne × LA voluntary	−0.2181	−0.1085	−0.0855	(dropped)	−0.2061
	(0.0314)	(0.0481)	(0.0414)		(0.0353)
Observations	6,840	6,840	3,420	3,420	1,440
R^2	0.9811	0.9811	0.4986	0.9809	0.9828

Notes: Covariates not shown include fixed effects for food-related illnesses in each three-digit zip code, fixed effects for non-food-related illnesses in each three-digit zip code, year dummies, and month dummies. We also include three-digit zip code-illness-type random effects (i.e., we cluster the standard errors by three-digit zip code and illness type with Huber-White standard errors). Standard errors are in parentheses. We attempted to use the same data and program as JL table VI. In that process, we discovered a minor coding error. Correcting that error makes the coefficients change in the second or the third decimal points. The original JL table VI reported the key coefficients as −0.2243 (0.0426) for food × LA mandatory and −0.2055 (0.0350) for food × LA voluntary.

and non-food hospitalizations separately, thus allowing time fixed effects to differ completely between the two types. The latter uses weaker assumptions than triple DID. Table 2 shows how JL table VI changes when we use these alternative specifications.

Since the new interaction term ($food_j \cdot post1998_t$) has a significant coefficient, we agree with HAH that our original assumption of common time fixed effects might be violated. Including this interaction reduces the estimates from −0.2217 and −0.2181 to −0.1161 and −0.1085, but they remain statistically significant at the 95 percent level. When we further relax the assumption by estimating food and non-food hospitalizations separately, we find the effect on foodborne hospitalizations remains significantly negative (−0.1471 and −0.0855) while the effect on non-food is not different from zero (−0.0205 and 0.0145). In the last column, we focus on LA county only. Again, results suggest that foodborne hospitalizations decline 16.54–20.61 percent after the LA regulation, relative to non-food hospitalizations in the same county.

II. Alternative Geographic Control Group?

HAH also criticizes our geographic definition of the control group. As mentioned above, JL uses two control groups: one is California (CA) outside LA, and the other is hospitalizations due to non-food-related digestive disorders. HAH argues that we should have used Southern CA only rather than the whole CA outside LA because they find LA and the other Southern CA counties follow more similar pretreatment trends before the start of our data.

In particular, HAH plots table V of JL in two graphs for food and non-food hospitalizations separately.

Since food and non-food hospitalizations appear in logs within the same regression, our Figure 2 puts them in the same picture with logs. We do not know how HAH translates hospitalization counts into hospitalization rate, so we stick with the original hospitalization counts reported in JL table V.

By using different scales on the vertical axis and not taking log, HAH figure 1 magnifies how hospitalization counts trend in logs. That been said, we agree it is more rigorous to add $food_j \cdot post1998_t$ to control for potentially different trends in food and non-food hospitalizations. With this addition, there is still a significant decline in foodborne hospitalizations after LA adopted hygiene grade cards in 1998 (Table 2).

To further address HAH's concern on pretreatment trends, we perform pretreatment lead tests in our data. Since our before-regulation data ranges from 1995 to 1997, we take 1995 as the default year and estimate the (placebo) impact of 1996 and 1997 on LA county, relative to three geographic control groups: the rest of CA, the other Southern CA counties, and the rest of CA excluding Southern CA. For complete transparency and comparison, we use the following specification for pretreatment and treatment tests:

$$\begin{aligned}(2)\qquad \ln(a_{ijt}) = {} & \alpha_{ij} + \tau_t + \theta_1 \cdot Treated_{ij} \cdot 1996 + \theta_2 \cdot Treated_{ij} \cdot 1997 \\ & + \theta_3 \cdot Treated_{ij} \cdot 1998 + \theta_4 \cdot Treated_{ij} \cdot 1999 + \varepsilon_{ijt},\end{aligned}$$

where *Treated* is defined as the interaction of food-related disorders (*food*) and percent of a three-digit zip code in LA county (*LActy*) in the full sample, or *LActy* alone in the food-only or non-food-only sample. The coefficients, $\{\theta_1, \theta_2, \theta_3, \theta_4\}$, capture how the level of the dependent variable differs each year away from the base year of 1995. This way, in the full sample, we contrast LA's food-related hospitalizations versus a big control group that includes foodborne hospitalizations in the control areas as well as non-food hospitalizations everywhere. In the subsamples (of food-only or non-food-only), we compare LA versus control areas within each illness group separately. Specification (2) does not include the stand-alone term of *Treated* because it is absorbed in the zip code-illness-type fixed effects (α_{ij}). Nor does it include the interaction $food_j \cdot post1998_t$ because we want to test pretreatment patterns in our

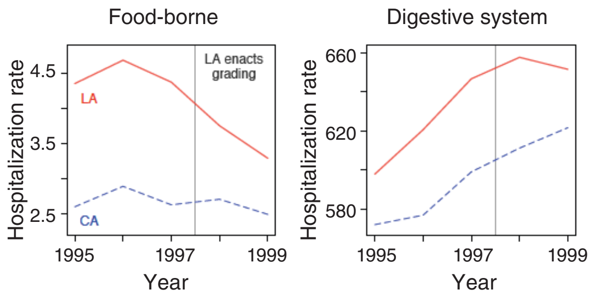

FIGURE 1. REPLICATION OF HAH

Notes: Trends in LA (solid) and the rest of CA (dashed) of hospitalizations for foodborne and (non-foodborne) digestive system disorders are in the left and right panels, respectively, from 1995–1999. The adoption of grading is denoted by the vertical line. This figure plots the same data as table V in JL (p. 437).

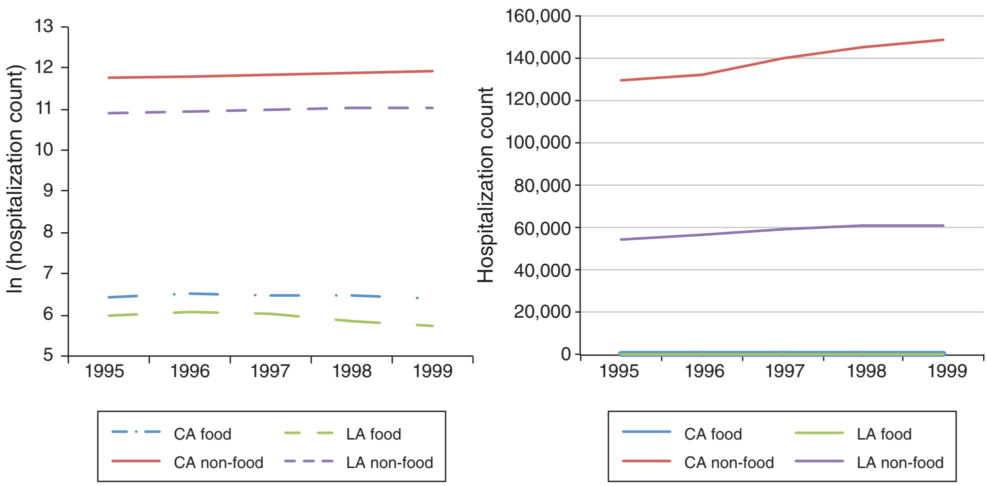

FIGURE 2. LN(HOSPITALIZATION COUNT) AS REPORTED IN JL TABLE V

original specification. However, pretreatment tests conducted in the food-only and non-food-only subsamples restrict comparison within each illness group and therefore are more comprehensive than adding the interaction in the full sample.

HAH argues that the 1994 salmonella outbreak in Southern CA continued its impact beyond 1994, making LA comparable to Southern CA but not to the rest of CA. If so, we should see $\{\theta_1, \theta_2\} = 0$ when we use Southern CA as the control, but $\{\theta_1, \theta_2\} \neq 0$ when we use other geography as controls. This HAH hypothesis is tested by F-statistics. For comparison, we also test $\{\theta_3, \theta_4\} = 0$ for changes after LA adopted grade cards. These tests are reported in Table 3.

According to Table 3, we never reject the null $\{\theta_1, \theta_2\} = 0$ on the estimate of θ_1, the estimate of θ_2, or the two jointly. This suggests that no control group has a significantly different pretreatment trend from foodborne hospitalizations in LA, no matter how we define the geographic control group or whether we use full sample

TABLE 3—PRETREATMENT AND TREATMENT TESTS BY DIFFERENT CONTROL GROUPS

Geographic control groups	Rest of CA (as in JL)	Other Southern CA counties (as suggested by HAH)	Rest of CA excluding Southern CA counties
Panel A. Full sample, "Treated" = food × percent of a three-digit zip code belonging to LA County			
1996 × treated (θ_1)	0.0034	0.0176	0.0058
	(0.0352)	(0.0367)	(0.0356)
1997 × treated (θ_2)	−0.029	−0.0342	−0.0255
	(0.0375)	(0.0414)	(0.0373)
1998 × treated (θ_3)	−0.1827	−0.1347	−0.2069
	(0.0352)	(0.0421)	(0.0349)
1999 × treated (θ_4)	−0.2113	−0.1794	−0.2252
	(0.0448)	(0.0521)	(0.0440)
F-test of $\mathrm{H_0}$: $\{\theta_1, \theta_2\} = 0$	0.4	0.72	0.35
p-value of F-test	0.6728	0.4883	0.7075
F-test of $\mathrm{H_0}$: $\{\theta_3, \theta_4\} = 0$	18.9	7.57	24.27
p-value of F-test	0	0.0011	0
R^2	0.9809	0.9801	0.9823
Observations	6,840	3,960	5,280
Panel B. Food-only, "Treated" = percent of a three-digit zip code belonging to LA County			
1996 × treated (θ_1)	−0.0234	0.0089	−0.0261
	(0.0421)	(0.0543)	(0.0462)
1997 × treated (θ_2)	0.0055	0.0117	0.0011
	(0.0439)	(0.0561)	(0.0481)
1998 × treated (θ_3)	−0.1375	0.0324	−0.2104
	(0.0491)	(0.0610)	(0.0538)
1999 × treated (θ_4)	−0.1327	−0.0241	−0.1727
	(0.0516)	(0.0727)	(0.0535)
F-test of $\mathrm{H_0}$: $\{\theta_1, \theta_2\} = 0$	0.23	0.03	0.21
p-value of F-test	0.7941	0.9727	0.8098
F-test of $\mathrm{H_0}$: $\{\theta_3, \theta_4\} = 0$	5.3	0.3	9.22
p-value of F-test	0.0078	0.7414	0.0005
R^2	0.4986	0.5318	0.559
Observations	3,420	1,980	2,640
Panel C. Non-food-only, "Treated" = percent of a three-digit zip code belonging to LA County			
1996 × treated (θ_1)	−0.0048	−0.0073	−0.0003
	(0.0193)	(0.0293)	(0.0192)
1997 × treated (θ_2)	−0.0109	−0.022	0.0067
	(0.0275)	(0.0582)	(0.0234)
1998 × treated (θ_3)	−0.0024	−0.0001	0.0123
	(0.0330)	(0.0651)	(0.0299)
1999 × treated (θ_4)	−0.0326	−0.0181	−0.0225
	(0.0376)	(0.0755)	(0.0355)
F-test of $\mathrm{H_0}$: $\{\theta_1, \theta_2\} = 0$	0.08	0.07	0.06
p-value of F-test	0.9243	0.9298	0.9437
F-test of $\mathrm{H_0}$: $\{\theta_3, \theta_4\} = 0$	1	0.51	0.78
p-value of F-test	0.3744	0.6056	0.4644
R^2	0.9809	0.9818	0.9842
Observations	3,420	1,980	2,640

Notes: All regressions control for year dummies, month dummies, and the fixed effects of three-digit zip codes by illness type (which absorb the dummy of "treated"). We count a three-digit zip code as Southern CA if it has any overlap with any Southern CA county other than LA. Clustered standard errors are in parentheses.

Table 4—Robustness Check of JL Table VI by Different Control Groups

Geographic control groups	Rest of CA (as in JL)	Other Southern CA counties (as suggested by HAH)	Rest of CA excluding Southern CA counties
Panel A. JL specification			
LA mandatory	0.0271	0.0965	0.0043
	(0.0239)	(0.0406)	(0.0230)
LA voluntary	0.0736	0.1491	0.0462
	(0.0238)	(0.0388)	(0.0238)
Food × LA mandatory	−0.2217	−0.2217	−0.2217
	(0.0431)	(0.0427)	(0.0433)
Food × LA voluntary	−0.2181	−0.2181	−0.2181
	(0.0314)	(0.0310)	(0.0320)
R^2	0.9811	0.9802	0.9823
Observations	6,840	3,960	5,280
Panel B. JL specification plus the interaction suggested by HAH			
LA mandatory	−0.0257	−0.0131	−0.0199
	(0.0246)	(0.0449)	(0.0232)
LA voluntary	0.0188	0.0354	0.0211
	(0.0237)	(0.0435)	(0.0224)
Food × after1998	−0.1007	−0.2089	−0.0461
	(0.0332)	(0.0595)	(0.0326)
Food × LA mandatory	−0.1161	−0.0027	−0.1734
	(0.0545)	(0.0744)	(0.0541)
Food × LA voluntary	−0.1085	−0.0093	−0.1679
	(0.0481)	(0.0724)	(0.0477)
R^2	0.9811	0.9804	0.9824
Observations	6,840	3,960	5,280

Notes: All regressions control for year dummies, month dummies, and the fixed effects of three-digit zip codes by illness type (which absorb the dummy of "treated"). We count a three-digit zip code as Southern CA if it has any overlap with any Southern CA county other than LA. Clustered standard errors are in parentheses.

or subsamples. Maybe more data will yield different conclusions, but these pretreatment tests validate the control groups used in JL and reject the HAH hypothesis.

In contrast, the treatment coefficients on 1998 and 1999 $\{\theta_3, \theta_4\}$ are of similar significance and magnitude as in JL, except in the food-only sample that uses Southern CA as the control area. And there is no treatment effect in the non-food-only sample across all control areas, as we would expect.

Table 4 repeats JL table VI by different geographic controls. In particular, we first repeat the JL specification and then report results including the interaction term suggested by HAH across the three geographic controls. Consistent with Table 3, we find similar treatment effects, except when the control area is Southern CA only and the model includes $food \cdot post1998$. This confirms HAH's finding that LA and the rest of Southern CA are similar in food-related hospitalizations even after LA adopted grade cards. However, what HAH does not show is that LA has seen a significant decline in foodborne hospitalizations than the rest of CA excluding Southern CA, and this decline only occurred after LA adopted disclosure regulation. Since the two areas are statistically comparable in 1995–1997, these results still support the conclusion in JL.

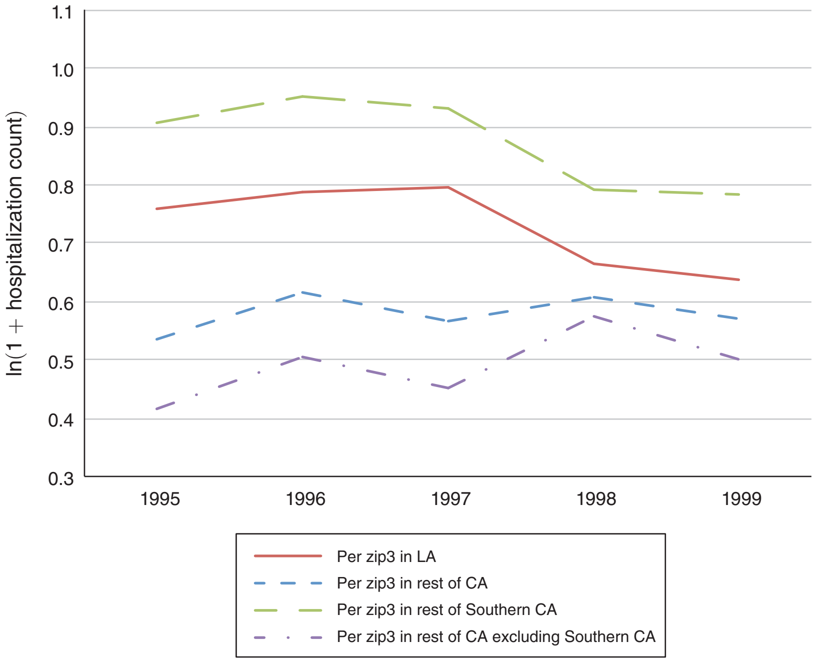

Figure 3. ln(1 + Foodborne Hospitalization Count) per Three-Digit Zip Code per Month Averaged by Area-Year

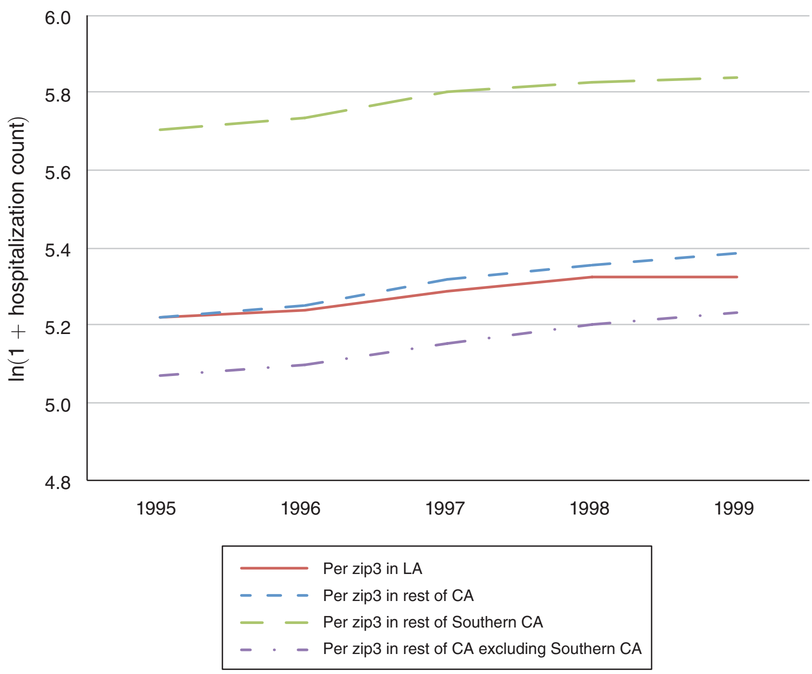

Figure 4. ln(1 + Non-Food Hospitalization Count) per Three-Digit Zip Code per Month Averaged by Area-Year

Results of Table 4 are also consistent with the raw data, even before we run any regressions. Figure 3 plots log of one plus foodborne hospitalization count per three-digit zip code per month. Figure 4 plots log of one plus non-food hospitalization count per three-digit zip code per month. They are essentially the dependent variable used in regressions. For an easy read, we average them by year and area. As confirmed in pretreatment tests, before 1998, there is no obvious trend difference between LA, the rest of CA, the rest of Southern CA, and the rest of CA excluding Southern CA. After 1998, LA witnessed a significant decline of foodborne hospitalizations relative to the rest of CA or central and Northern CA, but not relative to the rest of Southern CA. In comparison, there is no significant difference among these areas in non-food hospitalizations.

The remaining question is why LA and the rest of Southern CA are similar in foodborne hospitalizations, both before and after LA adopted grade cards. HAH argues that it is because Southern CA experienced a salmonella outbreak in 1994. This explains why LA and the rest of Southern CA look more similar in pretreatment trends if we include the outbreak year (1994) in pretreatment data, but it does not explain why the pretreatment patterns are similar *across all areas* in 1995–1997, but there is a significant difference between Southern CA and the rest of CA in 1998 and 1999. If the impact of the outbreak lingers after 1994, one would expect it gradually dissipates over time, not suddenly ramps up four years later.

Nevertheless, HAH raises a good question as to how long researchers should span the pretreatment data before estimating a treatment effect. Statistically speaking, the longer the better. Ideally, treatment and control groups behave identically all the time before the treatment, and the treatment is the only factor that makes them different thereafter. Unfortunately, this ideal does not exist in reality. The longer we accumulate data before treatment, the more likely we would find an outlier like the 1994 outbreak. Then should economists forego a research question until it fits the ideal econometric assumption? Should economists dig out all the historical data, at whatever costs, before conducting serious research?

We are glad that the research community has started to ponder these questions (Ruhm 2018), but here is why we chose to focus on 1995–1999 in JL. As HAH acknowledges, foodborne hospitalization is only a small part of JL. Our main focus was studying how grade card affects restaurant hygiene-inspection outcomes and how consumers respond to LA grade cards. For these questions, we have to rely on the data from the LA health department, which started in 1995 and ended in 1999. For consistency, we are most interested in how foodborne hospitalizations changed in the same time frame. Furthermore, despite our close contact with the LA health department, we were not aware of the 1994 outbreak until HAH raised the question. Should the outbreak have a long-lasting impact as HAH claims, we would expect LA officials to mention it during the study, given that they were the main force fighting against the outbreak and the main agency that implemented the grade cards. In hindsight, we still believe there is nothing wrong to choose 1995–1999 as the study period; and to evaluate the 1998 policy, comparable patterns between 1995–1997 are more important than a regional outbreak that occurred four years earlier.

It is puzzling why LA and the rest of Southern CA enjoyed the similar decline of foodborne hospitalizations after 1998 (relative to the rest of CA). One potential explanation is that the rest of Southern CA receives spillovers from LA. Admittedly, JL used the geographic definition of LA county and did not consider spillovers. But that does not rule out the possibility of spillovers, especially given the fact that there is little difference between voluntary and mandatory disclosure (which varies by geography and time).

Many factors could generate spillovers between LA and Southern CA. For example, the supply chain that supplies Southern CA could make their products cleaner in response to the demand of LA restaurants and grocery stores; consumers that reside in Southern CA could pay more attention to food hygiene after they read/watch news about LA's grade cards; the effect may spread by consumers traveling in and out of LA everyday;[2] Southern CA food retailers may already be alerted by the 1994 salmonella outbreak and the disclosure regulation reinforces that alert; or other Southern CA counties saw the LA regulation and started to educate their own retailers about food hygiene, even though some had introduced the same regulation before and others did not until years after.[3]

In online Appendix H, HAH tests for one type of the spillover effects and find the effect does not differ by the percent of area covered by KCBS-TV in each three-digit zip code in Southern CA. Based on this result, they dismiss the concern of spillover[4] and insist that LA is only comparable to Southern CA because Southern CA had a salmonella outbreak in 1994.

According to HAH, the aftermath of the outbreak had created a unique time trend in Southern CA. Thus, we should find LA significantly different from the rest of CA (excluding Southern CA) even before the disclosure regulation (in 1996 and 1997). The pretreatment tests shown in Table 3 do not support this hypothesis. In addition, for HAH's explanation to work, it must be that (i) the supply chain problem that drove the 1994 outbreak continued to influence the whole Southern CA throughout 1999, but (ii) the information and behavior changes triggered by the LA regulation have no impact on any common factors in Southern CA, including supply chains. We are not sure how to reconcile these two assumptions.

III. Redefine Foodborne Illnesses?

Neither of us were trained in public health, so we relied on an independent medical researcher to identify ICD-9 codes that relate to foodborne illnesses in a hospitalization. The researcher identified foodborne conditions in three categories: hospitalizations that are food related in over 90 percent of cases in 50–90 percent of

[2] The Census Bureau reports, "471,000 workers commute into Los Angeles County, Calif., each day" (https://www.census.gov/newsroom/press-releases/2013/cb13-r13.html).

[3] As HAH points out, a couple of Southern CA counties already had a similar disclosure regulation before 1998 (footnote 19).

[4] We have no evidence for or against the other potential spillover effects, but it is not logical to dismiss the possibility of spillovers based on the lack of evidence on one particular type of spillover. Though the regulation was triggered by the initial coverage by KCBS-TV in November 1997, many other media outlets (including *Los Angeles Times*) covered the progress of the regulation and market response to the regulation months after November 1997.

cases,[5] and in 10–50 percent of cases. In JL, we defined foodborne hospitalizations as food related in over 90 percent of cases and pool all the other incidents of digestive disorders as non-food hospitalization.

HAH argues that we should not exclude campylobacter from foodborne hospitalization. While different medical researchers may have different definitions, we believe our helper is a valid expert: he/she classifies campylobacter hospitalization as food related in 50–90 percent of cases and salmonella hospitalization as food related in over 90 percent of cases. This classification is consistent with HAH Table 1, which shows that campylobacter exceeds salmonella in the count of foodborne illnesses, but hospitalizations for campylobacter are less than half of those for salmonella. Nationwide, the hospitalization-to-illness ratio is 0.46:1 for salmonella, but only 0.19:1 for campylobacter, based on HAH table 1.

Our choice to treat campylobacter differently from salmonella is also confirmed by a recent publication by two public health experts at the University of Minnesota and the Minnesota Integrated Food Safety Center of Excellence (Firestone and Hedberg 2018). They show that after New York City (NYC) adopted restaurant hygiene grade cards in 2010, the rate of salmonella infections decreased 5.3 percent per year in NYC versus the rest of New York State during 2011–2015, as compared to 2006–2010. When commenting on JL's focus of hospitalization, they write, "Hospitalizations represent a subset of foodborne illnesses that may be caused by a variety of agents, such as *campylobacter*, another leading cause of foodborne illness in the United States. However, *campylobacter* rarely causes outbreaks in restaurant settings because its biology limits transmission to inadequate cooking of contaminated poultry or meats or cross-contamination from raw to ready-to-eat foods. As a result, improvements in restaurant sanitary conditions are unlikely to affect *campylobacter* transmission in restaurants."

HAH also points out that our helper's coding included rare diseases that could be caused by food problems, but some rare diseases do not have matches in the hospital discharge data. This might be true, but including them only makes the coding more thorough. It does not affect the aggregated data we use in the regressions.

In Table 5, we repeat JL table VI using four different definitions of foodborne hospitalization: (i) the conditions that our helper designated as food related in over 90 percent of cases, this is the definition in JL; (ii) the conditions that our helper designated as food related in over 50 percent of cases; (iii) the conditions that our helper designated as food related in over 10 percent of cases; and (iv) campylobacter only. In the first three definitions, we classify other digestive disorders as non-food related; in the last definition, we compare campylobacter with everything else except for those conditions that are food related in over 90 percent of cases.

Like Table 3, panel A of Table 5 reports pretreatment and treatment tests by specification (2). Given what we know about geographic control groups, Table 5 presents both the results of LA against the rest of CA and the results of LA against the rest of CA excluding Southern CA.

[5] According to our helper, diseases that are food related in 50–90 percent of cases include ICD-9 codes A001, A002, A004, A008.43, A008.44, A014, A041.4, A070.43, and A070.53. According to http://www.icd9data.com/2012/Volume1/001-139/001-009/008/008.43.htm, A008.43 refers to intestinal infection due to campylobacter.

TABLE 5—ROBUSTNESS CHECKS BY DIFFERENT DEFINITIONS OF FOODBORNE HOSPITALIZATION

Definition of foodborne hospitalization:	(i) Conditions that are food related in over 90 percent of the cases (as in JL)		(ii) Conditions that are food related in over 50 percent of the cases		(iii) Conditions that are food related in over 10 percent of the cases		(iv) *Campylobacter*-only	
Definition of non-food hospitalization:	All conditions that are food related in less than 90 percent of the cases		All conditions that are food related in less than 50 percent of the cases		All conditions that are food related in less than 10 percent of the cases		All the others except for conditions that are food related in over 90 percent of the cases	
Geographic control group:	Rest of CA	Rest of CA excluding Southern CA	Rest of CA	Rest of CA excluding Southern CA	Rest of CA	Rest of CA excluding Southern CA	Rest of CA	Rest of CA excluding Southern CA
Panel A. Pretreatment and treatment tests by specification (2), *"Treated" = food × percent of a three-digit zip code belonging to LA County*								
1996 × treated (θ_1)	0.0034	0.0058	−0.015	−0.0235	−0.073	−0.0657	−0.0185	−0.0086
	(0.0352)	(0.0356)	0.0438	(0.0439)	(0.0441)	(0.0446)	(0.0309)	(0.0322)
1997 × treated (θ_2)	−0.029	−0.0255	−0.067	−0.0721	0.0006	0.0056	−0.0794	−0.0709
	(0.0375)	(0.0373)	(0.0431)	(0.0428)	(0.0517)	(0.0521)	(0.0279)	(0.0284)
1998 × treated (θ_3)	−0.1827	−0.2069	−0.208	−0.2378	−0.1316	−0.1379	−0.0874	−0.0823
	(0.0352)	(0.0349)	(0.0464)	(0.0453)	(0.0593)	(0.0602)	(0.0344)	(0.0361)
1999 × treated (θ_4)	−0.2113	−0.2252	−0.2062	−0.2198	−0.2428	−0.2609	−0.1073	−0.0974
	(0.0448)	(0.0440)	(0.0539)	(0.0535)	(0.0464)	(0.0451)	(0.0452)	(0.0466)
F-test of $\mathrm{H_0}$: $\{\theta_1, \theta_2\} = 0$	0.4	0.35	1.57	1.71	1.87	1.58	4.31	3.66
p-value of F-test	0.6728	0.7075	0.2127	0.1863	0.1585	0.2114	0.0158	0.0299
F-test of $\mathrm{H_0}$: $\{\theta_3, \theta_4\} = 0$	18.9	24.27	12.96	16.76	14.12	17.37	3.54	2.81
p-value of F-test	0	0	0	0	0	0	0.0324	0.0657
R^2	0.9809	0.9823	0.9774	0.9787	0.9679	0.9696	0.9904	0.9901
Observations	6,840	5,280	6,840	5,280	6,840	5,280	6,840	5,280
Panel B. JL specification								
LA mandatory	0.0271	0.0043	0.0687	0.0566	0.0423	0.0175	0.0712	0.0807
	(0.0239)	(0.0230)	(0.0257)	(0.0263)	(0.0253)	(0.0246)	(0.0230)	(0.0254)
LA voluntary	0.0736	0.0462	0.1215	0.1027	0.0152	0.0018	0.1088	0.1181
	(0.0238)	(0.0238)	(0.0249)	(0.0251)	(0.0329)	(0.0326)	(0.0228)	(0.0255)
Food × LA mandatory	−0.2217	−0.2217	−0.2235	−0.2235	−0.2329	−0.2329	−0.11	−0.11
	(0.0431)	(0.0433)	(0.0448)	(0.0448)	(0.0586)	(0.0589)	(0.0293)	(0.0295)
Food × LA voluntary	−0.2181	−0.2181	−0.2969	−0.2969	−0.116	−0.116	−0.1793	−0.1793
	(0.0314)	(0.0320)	(0.0464)	(0.0465)	(0.0575)	(0.0578)	(0.0317)	(0.0315)
R^2	0.9811	0.9823	0.9774	0.9787	0.9679	0.9696	0.9904	0.9901
Observations	6,840	5,280	6,840	5,280	6,840	5,280	6,840	5,280
Panel C. JL specification plus the interaction suggested by HAH								
LA mandatory	−0.0257	−0.0199	−0.0277	−0.0204	0.0098	0.0112	−0.0208	−0.0152
	(0.0246)	(0.0232)	(0.0247)	(0.0233)	(0.0252)	(0.0239)	(0.0246)	(0.0231)
LA voluntary	0.0188	0.0211	0.0214	0.0228	−0.0185	−0.0047	0.0132	0.0185
	(0.0237)	(0.0224)	(0.0233)	(0.0222)	(0.0342)	(0.0309)	(0.0248)	(0.0233)
Food × after1998	−0.1007	−0.0461	−0.1839	−0.1468	−0.0618	−0.012	−0.1755	−0.183
	(0.0332)	(0.0326)	(0.0332)	(0.0345)	(0.0335)	(0.0345)	(0.0249)	(0.0279)
Food × LA mandatory	−0.1161	−0.1734	−0.0306	−0.0695	−0.1681	−0.2204	0.0741	0.0818
	(0.0545)	(0.0541)	(0.0563)	(0.0573)	(0.0682)	(0.0692)	(0.0391)	(0.0413)
Food × LA voluntary	−0.1085	−0.1679	−0.0967	−0.1371	−0.0487	−0.103	0.0118	0.0199
	(0.0481)	(0.0477)	(0.0591)	(0.0599)	(0.0691)	(0.0700)	(0.0409)	(0.0427)
R^2	0.9811	0.9824	0.9777	0.9788	0.9679	0.9696	0.9906	0.9903
Observations	6,840	5,280	6,840	5,280	6,840	5,280	6,840	5,280

Notes: All regressions control for year dummies, month dummies, and the fixed effects of three-digit zip codes by illness type (which absorb the dummy of "treated"). We count a three-digit zip code as Southern CA if it has any overlap with any Southern CA county other than LA. Clustered standard errors are in parentheses.

Panel A suggests that all three definitions—(i), (ii), and (iii)—passed the pretreatment test of $\{\theta_1, \theta_2\} = 0$, though the stand-alone estimate of θ_2 is marginally significant under definition (iii). In contrast, definition (iv) suggests that campylobacter alone follows a quite different trend from the other non-food hospitalizations. This is consistent with figure 8 in an earlier version of HAH, which shows different time trends of campylobacter in LA and CA.[6] Because campylobacter is one of the largest groups in definitions (ii) and (iii), the campylobacter-specific trend may explain why the pretreatment tests in definitions (ii) and (iii) have higher *F*-statistics than in definition (i).

Putting aside the pretreatment tests, panel B of Table 5 repeats the JL specification by different definitions of foodborne hospitalization. Panel C further adds the interaction term suggested by HAH. While results of the JL specification are consistent across the board, we acknowledge HAH's point about the interaction term and believe results in panel C are more trustworthy.

Let us read panel C backwards from definition (iv). The whole CA appears to witness an overall declining trend of campylobacter since 1997. Once we control for this overall trend, disclosure regulation has no negative effects on the incidence of campylobacter. The coefficient of $food \cdot LA mandatory$ is even positive (and marginally significant). Probably because of this, when we include campylobacter as food related in definitions (ii) and (iii), some coefficients corresponding to the disclosure regulation lose significance, though all of them remain negative and the coefficient of $food \cdot LA mandatory$ is significant under definition (iii). This is largely consistent with what HAH gets when they add campylobacter to our original definition of foodborne hospitalization.

However, if we exclude Southern CA from the control area (because it may receive spillovers from LA), more coefficients restore their statistical significance. And their magnitude ($-0.0695 \sim -0.2204$) is similar to what we get using the original definition of foodborne hospitalization, namely $-0.1085 \sim -0.1161$ if we compare LA county against the rest of CA and $-0.1679 \sim -0.1734$ if we compare LA county against the rest of CA excluding Southern CA. These coefficients suggest that foodborne hospitalizations declined 6.95 percent to 22.04 percent in LA relative to central and Northern CA, even after we expand foodborne hospitalization to include many other diseases.

IV. Conclusion

Overall, we are grateful that HAH replicates and extends our original study. Their criticism prompts us to conduct extra robustness checks. These checks suggest that our original conclusion still holds—foodborne hospitalizations, as defined in JL, have declined in LA relative to the rest of CA after LA adopted restaurant hygiene grade cards in 1998. More precisely, the decline in LA is pronounced when we compare it against central and Northern CA, but there is little difference between LA and other Southern CA counties.

[6] This figure is no longer available in the current version of HAH.

HAH (2018) suggests that the 1994 salmonella outbreak in Southern CA may account for the post-1998 decline of foodborne hospitalizations in that region. We have no direct evidence to rule out this possibility, but it does not explain why foodborne hospitalizations followed similar patterns across CA in 1995–1997 but declined significantly in Southern CA in 1998–1999. If the impact of the 1994 outbreak lasts for years, it should be more pronounced in 1995–1997 than in 1998–1999. Another possible explanation is that the 1998 LA regulation has generated spillovers to Southern CA.

HAH also questions the generalizability of JL. Our study may have motivated policy discussion in other jurisdictions or other contexts, but we never argue that our results should apply to any other local government equally if it adopts the same disclosure policy. HAH also suggests that table VI of JL persuaded other jurisdictions to adopt similar disclosure policies in restaurant hygiene. This is likely an overstatement. Jin testified in Maryland but the state decided against her recommendation, mostly because of industry objection. In a project funded by the Sloan Foundation, we collected online posting of hygiene inspection results from over 100 jurisdictions in the United States (including the 30 biggest cities). Only a small fraction (~10 percent) adopts any kind of score categorization (letter grade, pass-fail, etc.).

As for the importance and limit of simple disclosure, some works after JL are worth mentioning. In Jin and Leslie (2005), we emphasize that the changes we observe in LA are not only driven by printing a letter on a piece of paper, but also by media attention, political pressure, the LA-specific assessment system, and the corresponding educational efforts from the LA health department. In Jin and Leslie (2009), we present evidence that the LA disclosure regulation had different impacts on different restaurants and the pattern is consistent with reputation incentives existing before the regulation. Our recent publication (Bederson et al. 2018) shows that voluntary disclosure of restaurant hygiene in Maricopa, Arizona, is far from complete. In fact, those that chose nondisclosure include both the very best restaurants and the very worst ones. This result, together with many papers reviewed in Dranove and Jin (2010), highlights the pros and cons of disclosure policy. In two working papers (Jin, Luca, and Martin 2018a, 2018b), Jin and her coauthors conduct a series of lab experiments, investigating why human subjects may hide non-favorable information and why the traditional disclosure theory fails when disclosure involves complexity.

Other researchers have also studied restaurant hygiene disclosure. Kang et al. (2013) uses machine learning to demonstrate that one can uses Yelp review of restaurants to predict restaurant hygiene-inspection results. Dai and Luca (2018) reports two field experiments on Yelp. They find that digital posting of restaurant hygiene scores leads to a 12 percent decrease in purchase intentions for restaurants with low scores (as predefined by the City of San Francisco) and a more salient "hygiene alert" leads to a further 7 percent decrease in purchase intentions. The alert also reduces the restaurant's likelihood of getting a second alert.

In the public health literature, Serapiglia et al. (2007) finds the pass-fail system in Toronto associated with a highly significant decrease in the incidence of salmonella, *campylobacter*, and other foodborne diseases. NYC adoption of a restaurant hygiene letter grade has been studied by NYCDHMH (2012), Wong et al. (2015), and Firestone and Hedberg (2018). NYCDHMH (2012) finds that salmonella infections have declined since NYC implemented letter grading, as compared to the rest of New York, Connecticut, and New Jersey. Firestone and Hedberg (2018) confirms this finding by comparing NYC to the rest of New York State. Wong et al. (2015) associates the NYC letter grade with sustained improvements in sanitary conditions in restaurants, including several factors related to outbreaks. They also show that after 18 months, 81 percent of adults in NYC had seen letter-grade placards, and 88 percent of them have considered letter grades in their dining decisions. We hope our research will continue to encourage more researchers to work in this area.

REFERENCES

Bederson, Benjamin B., Ginger Zhe Jin, Phillip Leslie, Alexander J. Quinn, and Ben Zou. 2018. "Incomplete Disclosure: Evidence of Signaling and Countersignaling." *American Economic Journal: Microeconomics* 10 (1): 41–66.

Dai, Weijia (**Daisy**)**, and Michael Luca.** 2018. "Digitizing Disclosure: The Case of Restaurant Hygiene Scores." Harvard Business School Working Paper 18-088.

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

Firestone, Melanie J., and Craig W. Hedberg. 2018. "Restaurant Inspection Letter Grades and Salmonella Infections, New York, New York, USA." *Emerging Infectious Diseases* 24 (12): 2164–68.

Ho, Daniel E., Zoe C. Ashwood, and Cassandra Handan-Nader. 2019. "New Evidence on Information Disclosure through Restaurant Hygiene Grading." *American Journal of Economics: Economic Policy* 11 (4): 404–28.

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

Jin, Ginger Zhe, and Phillip Leslie. 2005. "The Case in Support of Restaurant Hygiene Grade Cards." *Choices* 20 (2): 97–102.

Jin, Ginger Zhe, and Phillip Leslie. 2009. "Reputational Incentives for Restaurant Hygiene." *American Economic Journal: Microeconomics* 1 (1): 237–67.

Jin, Ginger Zhe, and Phillip Leslie. 2019. "New Evidence on Information Disclosure through Restaurant Hygiene Grading: Reply: Dataset." *American Economic Journal: Economic Policy*. https://doi.org/10.1257/pol.20180543.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin. 2018a. "Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure." NBER Working Paper 21099.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin. 2018b. "Complex Disclosure." NBER Working Paper 24675.

Kang, Jun Seok, Polina Kuznetsova, Michael Luca, and Yejin Choi. 2013. "Where Not to Eat? Improving Public Policy by Predicting Hygiene Inspections Using Online Reviews." In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, edited by David Yarowsky, Timothy Baldwin, Anna Korhonen, Karen Livescu, and Steven Bethard, 1443–48. Stroudsburg: Association for Computational Linguistics.

New York City Department of Health and Mental Hygiene (NYCDHMH). 2012. *Restaurant Grading in New York City at 18 Months*. New York: NYCDHMH.

Ruhm, Christopher J. 2018. "Shackling the Identification Police?" NBER Working Paper 25320.

Serapiglia, Tino, Erin Kennedy, Sylvanus Thompson, and Ron de Burger. 2007. "Association of Food Premises Inspection and Disclosure Program with Retail-Acquired Foodborne Illness and Operator Noncompliance in Toronto." *Journal of Environmental Health* 70 (1): 54–59.

Wong, Melissa R., Wendy McKelvey, Kazuhiko Ito, Corinne Schiff, J. Bryan Jacobson, Daniel Kass. 2015. "Impact of a Letter-Grade Program on Restaurant Sanitary Conditions and Diner Behavior in New York City." *American Journal of Public Health* 105 (3): e81–7.