

How Scheduling Can Bias Quality Assessment

Online Supplement to

How Scheduling Can Bias Quality Assessment:

Evidence from Food Safety Inspections

Maria R. Ibanez

Kellogg School of Management at Northwestern University, Evanston, IL 60208,
maria.ibanez@kellogg.northwestern.edu

Michael W. Toffel

Harvard Business School, Boston, MA 02163, mtoffel@hbs.edu

Appendix A. Supplemental Descriptive Statistics and Correlations

**Table A1. Inspection Sequence
within the Day**

1st inspection of the day	5,328
2nd inspection of the day	3,618
3rd inspection of the day	1,971
4th inspection of the day	763
5th inspection of the day	248
6th inspection of the day	61
7th+ inspection of the day	28
Total number of inspections	12,017

**Table A2. Number of Inspector-days
by Inspection Workload**

Inspector-days with 1 inspection	1,790
Inspector-days with 2 inspections	2,226
Inspector-days with 3 inspections	1,637
Inspector-days with 4 inspections	801
Inspector-days with 5 inspections	295
Inspector-days with 6 inspections	83
Inspector-days with 7+ inspections	48
Total number of inspector-days	6,880

An *inspector-day* refers to a particular day during which a given inspector conducts at least one inspection.

Table A3. Inspections by Hour Begun and Corresponding Meal Period

7 am or earlier	39	3,856 during <i>breakfast period</i> (inspection began midnight to 10:59 am)
8 am	222	
9 am	972	
10 am	2,623	
11 am	1,986	7,888 during <i>lunch period</i> (inspection began 11:00 am–3:59 pm)
12 pm	1,331	
1 pm	2,331	
2 pm	1,653	
3 pm	587	273 during <i>dinner period</i> (inspection began 4:00 pm–11:59 pm)
4 pm	171	
5 pm	59	
6 pm or later	43	
Total number of inspections:		12,017

Table A4. Extensions: Summary Statistics

Variable	Obs	Mean	SD	Min	Max
Potential extent of shift-prolonging (in hours)	12,017	0.25	0.52	0	1.79
Potential extent of shift-prolonging is >0.5 & <=1 hours (binary)	12,017	0.07	0.25	0	1
Potential extent of shift-prolonging is >1 & <=1.5 hours (binary)	12,017	0.05	0.21	0	1
Potential extent of shift-prolonging is >1.5 & <= max of 1.79 hours (binary)	12,017	0.07	0.25	0	1
Penultimate inspected establishment's violations	12,014	2.10	2.60	0	25
Antepenultimate inspected establishment's violations	12,008	2.07	2.57	0	25
Prior inspected establishment's violations for the first inspection of the day	12,017	0.96	2.05	0	25
Prior inspected establishment's violations for the second+ inspection of the day	12,017	1.15	2.21	0	23
Overnight break length	12,017	2.88	7.36	0	76.40
Within-day break length	12,017	0.04	0.08	0	1.10
Break length	12,017	2.92	7.35	0	76.40
Critical violations	12,017	0.93	1.27	0	11
Noncritical violations	12,017	1.49	1.94	0	16
Inspection duration (minutes)	12,017	62.87	32.01	5	270
Log inspection duration	12,017	4.01	0.55	1.61	5.60
Violation citations per hour	12,017	2.35	2.66	0	42
Log (violation citations per hour + 1)	12,017	0.96	0.71	0	3.76

Table A5. Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Violations	1.00									
(2) Prior inspected establishment's violations (H1)	0.18*	1.00								
(3) Prior inspected establishment's violation trend (H2)	0.11*	0.56*	1.00							
(4) Prior inspected establishment saliently improved (H3)	0.01	-0.17*	-0.37*	1.00						
(5) Prior inspected establishment saliently deteriorated (H3)	0.12*	0.60*	0.69*	-0.29*	1.00					
(6) Time inspecting earlier today (H4)	0.01	0.10*	0.08*	-0.02	0.08*	1.00				
(7) Potentially shift-prolonging (H5)	-0.01	0.03*	0.01	0.00	0.01	0.38*	1.00			
(8) Inspector experience	-0.05*	-0.01	0.02	-0.04*	0.01	-0.03*	-0.02*	1.00		
(9) Returning inspector	-0.14*	-0.03*	-0.01	-0.01	-0.01	-0.05*	0.10*	0.33*	1.00	
(10) Establishment's <i>n</i> th inspection (2nd through 10th)	0.05*	0.01	0.02*	-0.01	0.02*	-0.02*	0.06*	0.48*	0.35*	1.00
(11) Lunch period (11:00 am–3:59 pm)	-0.01	0.02*	0.02*	-0.01	0.02	0.39*	0.34*	-0.01	0.01	-0.01
(12) Potential extent of shift-prolonging (in hours)	-0.03*	0.02	0.00	0.00	0.00	0.35*	0.8*	-0.05*	0.07*	0.05*
(13) Potential extent of shift-prolonging is >0.5 & <=1 hours (binary)	-0.01	0.01	0.01	-0.01	0.00	0.15*	0.44*	0.01	0.05*	0.02*
(14) Potential extent of shift-prolonging is >1 & <=1.5 hours (binary)	0.00	0.00	0.01	0.02	0.00	0.16*	0.37*	0.00	0.05*	0.04*
(15) Potential extent of shift-prolonging is >1.5 & <= max of 1.79 hours (binary)	-0.03*	0.01	-0.01	0.00	0.00	0.23*	0.46*	-0.06*	0.02*	0.02*
(16) Penultimate inspected establishment's violations	0.15*	0.17*	0.07*	0.05*	0.1*	0.03*	-0.01	-0.02*	-0.04*	0.01
(17) Antepenultimate inspected establishment's violations	0.15*	0.16*	0.07*	0.05*	0.08*	0.03*	-0.01	-0.02*	-0.03*	0.01
(18) Prior inspected establishment's violations for the first inspection of the day	0.11*	0.58*	0.31*	-0.1*	0.33*	-0.39*	-0.15*	-0.05*	-0.02*	0.00
(19) Prior inspected establishment's violations for the second+ inspection of the day	0.12*	0.65*	0.38*	-0.11*	0.41*	0.47*	0.17*	0.03*	-0.01	0.02
(20) Overnight break length	0.00	0.00	-0.01	0.00	-0.01	-0.32*	-0.08*	-0.12*	-0.04*	-0.05*
(21) Within-day break length	0.01	-0.02*	0.00	-0.02*	-0.01	0.26*	0.27*	0.03*	0.01	0.01
(22) Break length	0.00	0.00	-0.01	0.00	-0.01	-0.32*	-0.07*	-0.11*	-0.04*	-0.05*
(23) Critical violations	0.76*	0.13*	0.08*	0.01	0.09*	0.00	0.01	-0.08*	-0.16*	-0.01
(24) Noncritical violations	0.91*	0.17*	0.10*	0.01	0.11*	0.01	-0.02*	-0.01	-0.09*	0.07*
(25) Inspection duration (minutes)	0.43*	0.02	0.01	-0.01	0.01	-0.03*	-0.03*	-0.19*	-0.14*	-0.01
(26) Log inspection duration	0.40*	0.00	0.01	-0.02*	0.01	-0.05*	-0.04*	-0.16*	-0.11*	-0.03*
(27) Violation citations per hour	0.75*	0.18*	0.11*	0.02	0.12*	0.03*	0.02	0.03*	-0.06*	0.04*
(28) Log (violation citations per hour + 1)	0.78*	0.17*	0.10*	0.01	0.12*	0.02*	0.00	0.00	-0.09*	0.01

N = 12,017 inspections

(continued on next page)

Table A5. Correlations (continued)

	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)
(11)	1.00																	
(12)	0.21*	1.00																
(13)	0.19*	0.26*	1.00															
(14)	0.15*	0.43*	-0.06*	1.00														
(15)	0.05*	0.79*	-0.07*	-0.06*	1.00													
(16)	-0.02*	-0.01	0.01	-0.01	-0.01	1.00												
(17)	0.00	-0.01	0.00	-0.01	0.00	0.16*	1.00											
(18)	-0.25*	-0.13*	-0.07*	-0.07*	-0.06*	0.1*	0.09*	1.00										
(19)	0.26*	0.13*	0.08*	0.06*	0.07*	0.12*	0.11*	-0.25*	1.00									
(20)	-0.17*	-0.05*	-0.04*	-0.04*	-0.01	0.01	-0.01	0.22*	-0.20*	1.00								
(21)	0.26*	0.27*	0.12*	0.1*	0.19*	-0.02*	0.01	-0.24*	0.20*	-0.20*	1.00							
(22)	-0.16*	-0.04*	-0.03*	-0.04*	-0.01	0.01	-0.01	0.22*	-0.20*	1.00*	-0.19*	1.00						
(23)	-0.02*	0.00	0.00	-0.01	0.01	0.11*	0.11*	0.09*	0.07*	0.04*	0.01	0.04*	1.00					
(24)	0.00	-0.04*	-0.01	0.00	-0.04*	0.14*	0.14*	0.09*	0.12*	-0.03*	0.01	-0.03*	0.42*	1.00				
(25)	-0.09*	-0.04*	-0.03*	-0.02*	-0.02*	0.03*	0.02*	0.09*	-0.07*	0.10*	-0.04*	0.10*	0.41*	0.33*	1.00			
(26)	-0.10*	-0.05*	-0.02*	-0.03*	-0.03*	0.01	0.01	0.09*	-0.07*	0.09*	-0.04*	0.09*	0.37*	0.32*	0.92*	1.00		
(27)	0.05*	0.00	0.01	0.02*	-0.02	0.15*	0.15*	0.05*	0.16*	-0.04*	0.03*	-0.04*	0.51*	0.72*	-0.03*	-0.04*	1.00	
(28)	0.04*	-0.01	0.00	0.01	-0.02*	0.14*	0.14*	0.06*	0.15*	-0.03*	0.03*	-0.03*	0.57*	0.72*	0.08*	0.10*	0.89*	1.00

Appendix B. Interpretation of Results

To illustrate the magnitude of the estimated effects, we consider interventions that exploit outcome effects and ameliorate daily schedule effects, both of which would lead inspectors to cite violations that currently go underreported. In particular, we consider various scenarios that both (a) *amplify the outcome effects* in order to more routinely trigger the heightened inspector scrutiny that ensues after inspections reveal many violations and worsening compliance trends, and (b) *mitigate the daily schedule effects* in order to attenuate the reduced scrutiny that accompanies successive inspections and potentially shift-prolonging inspections. We estimate the effects of such interventions on the average inspection based on our sample, scale up the results to estimate the impact across the entire United States, and translate how such an increase in cited violations would translate to fewer foodborne illness cases and their associated healthcare costs.

In the best-case scenario, outcome effects (which increase scrutiny) would be fully triggered all the time and daily schedule effects (which erode scrutiny) would be entirely eliminated. The full consequence of these biases is reflected by the difference in inspection outcomes between this best-case scenario and the status quo, which quantifies the number of unreported violations and excess illnesses and costs that could be avoided if steps were taken to address these biases. Our discussions with inspectors suggest that some interventions are feasible—such as limiting or smoothing the number of inspections each inspector conducts per day—often without imposing any additional costs. We estimate a range of scenarios that consider the impacts associated with the daily schedule effects being attenuated by—and the outcome effects being actuated by—varying amounts.

We first consider the average impact on violations cited per inspection. Specifically, we compare the status quo (that is, the current practice with its associated scheduling effects) with alternative scenarios that consider various percentage changes (10% to 100% in 10% increments) of the effects we identified that would increase inspectors' detection rate (that is, decrease by 10% the daily schedule effects and increase by 10% the outcome effects). We make all these comparisons based on Model 1 in Table 2. Specifically, we calculate average predicted values for each scenario based on the model's estimates after recoding the estimated coefficients on *time inspecting earlier today*, *potentially shift-prolonging*, *prior inspected establishment's violations*, and *prior inspected establishment's violation trend* by the percentage specified; we report results in Column 1 of Table B1. We obtain equivalent results if we preserve the estimated coefficients and instead recode the values of the variables by that same percentage; thus, the results can be interpreted as altering the per-unit bias represented by the estimated coefficients (for example, raising inspectors' detection rates to the heightened levels associated with the identified outcome effects) or as altering the factors that generate the bias (for example, reducing the number of prior inspections conducted by the inspector per day to reduce the *time inspecting earlier*

today). For the status quo, we use the model's estimates to calculate the average predicted number of violations per inspection, based on actual values of all variables, to be 2.42365. Column 2 reports the percent change (from the status quo) in the average predicted violations for each of these scenarios. This shows the average percent change in violations cited per inspection compared to the status quo.

For example, consider the very conservative "10% scenario" depicted in the second row of Table B1. In this scenario, we estimate the effects of (1) amplifying the outcome effects by increasing by 10% the actual values of *prior inspected establishment's violations* and *prior inspected establishment's violation trend* while also (2) mitigating the daily schedule effects by decreasing by 10% the actual values of *time inspecting earlier today* and *potentially shift-prolonging*. Applying these recoded values to the coefficient estimates from Model 1 of Table 2, we calculate the average predicted number of violations to be 2.448 (Column 1). This indicates that this "10% scenario" would result in 1.01% more violations being cited per inspection than the status quo of 2.42365 (Column 2).¹

We then estimate the potential nationwide implications of our calculations, based on the assumptions that the estimated one million food establishments that are monitored by state, local, and tribal agencies in the United States (US Food and Drug Administration 2016) are each inspected annually and that our sample of inspections is representative of those conducted across the country. To calculate a nationwide figure, we take the difference between the average predicted values from each scenario and the status quo (that is, the Column 1 figure minus 2.42365) and multiply that by the one million inspections conducted annually across the country; results are reported in Column 3. Those figures can be interpreted in the context of an estimated 2.4 million violations cited in the status quo scenario.² Continuing the 10% scenario, we scale the difference in average predicted violations per inspection that arise in this scenario compared to the status quo (2.44819 - 2.42365) by the one million inspections conducted annually nationwide to estimate that this scenario would yield 24,536 additional violations (currently undetected) being cited nationwide per year (Column 3).

Citing violations leads establishments to improve their food safety practices, which in turn mitigates foodborne illness cases and resulting hospitalizations. To calculate the health impacts if these undetected violations were being cited, we translate the estimated nationwide changes in violation counts into health outcomes and their associated costs. We attempt to be as conservative as possible but acknowledge that there are uncertainties associated with these conversions.³

¹ The 1.01% figure is calculated as $(2.44819 - 2.42365) / 2.42365$.

² The 2.4 million figure is calculated by multiplying 2.424 violations cited per inspection in the status quo scenario by the one million inspections conducted annually nationwide.

³ The estimates we construct should be considered as an illustration of the possible implications of the biases. We acknowledge the possibility that our estimates might *overestimate* the effects if the conversion factors we use overestimate the benefits of citing a particular violation and that they might *underestimate* the effects because we do not incorporate the spillovers and system-wide benefits of citing a particular violation, as each citation may encourage establishments to improve health practices more broadly. Failing to cite one violation thus not only carries the health risks associated with that violation but may also encourage noncompliance—an effect similar to the broken window phenomenon. That said, while developing a more

First, we consider how the increased violations cited per inspection beyond the status quo translates into fewer hospitalizations for foodborne illness (Columns 4 and 5). We multiply the percent change in the average predicted number of violations between the scenario and the status quo (Column 2) by the ratio of a 20% decrease in hospitalizations per 5% improvement in restaurant compliance scores based on prior research on Los Angeles restaurants (Jin and Leslie 2003).⁴ To calculate the impact of the 10% scenario, we multiply the 1.01% increase in the number of violations cited per inspection (Column 2) by the ratio of 20% decrease in hospitalizations per 5% improvement in restaurant compliance scores, which indicates that hospitalizations would decrease by 4.05% (Column 4). This would correspond to 5,183 fewer hospitalizations for foodborne illness each year across the United States (Column 5), based on applying the 4.05% decline to the estimated 128,000 annual US hospitalizations for foodborne illness under the status quo (Scallan et al. 2011).

We also estimate the impact of citing more violations on annual nationwide foodborne illness cases (Column 6). Based on the ratio of Scallan et al.'s (2011) two nationwide annual estimates of 47.8 million foodborne illness cases and 128,000 annual foodborne illness hospitalizations, there are 373.4 foodborne illness cases per foodborne illness hospitalization. Therefore, the 10% scenario, estimated earlier to reduce foodborne illness hospitalizations by 5,183, would also reduce foodborne illness cases by 1.94 million cases (calculated as $5,183 * 373.4$).

Finally, we estimate the impact of citing more violations on the costs associated with foodborne illness cases based on two alternative estimates of the average cost per foodborne illness case of \$747 (Minor et al. 2015) and \$1,626 (Scharff 2012), which we use to construct the lower and upper bounds of our cost estimates (Columns 7 and 8). In the 10% scenario, applying these figures to the estimated 1.94 million fewer foodborne illness cases compared to the status quo yields a \$1,446-million-to-\$3,147-million drop in the annual costs associated with foodborne illness cases nationwide.

As noted, there are many assumptions and caveats associated with these analyses and one can consider alternative scenarios. Our estimations above assume that mitigating bias would yield citations of violations that are as correlated with foodborne incidents as the violations currently cited. But what if newly cited violations are less “important,” meaning they impose less health risk? For example, suppose

comprehensive methodology to estimate the health impacts of citing more food safety violations is a necessary and worthy endeavor, it is beyond the scope of this paper.

⁴ We are aware of little research that has estimated the effect of each food safety violation on health outcomes and we rely on Jin and Leslie (2003), which we believe presents the best estimate. They show that introducing restaurant grade cards—signs posted outside restaurants that report the establishment's letter grade based on its most recent food safety inspection results—affects food safety inspection violation scores and health outcomes, so restaurant grade cards can be viewed as an instrument that reveals the relationship between violations cited and health outcomes. Because violations are supposed to be corrected when cited, we assume that the new citations resulting from reducing the bias translate into fewer actual violations. (To be conservative, we are not accounting for how citations motivate compliance more broadly.) The relationship Jin and Leslie (2003) identified between compliance and health outcomes applies to our setting because it is based on a similar type of inspection and a compliance measure based on total violations, which implicitly controls for the heterogeneous effects of different types of violation on health.

that remediating a newly cited violation would prevent half as many foodborne incidents as remediating a currently cited violation. Estimating the health impacts would then require adjusting Jin and Leslie's (2003) finding that a 5% improvement in restaurant compliance yields a 20% decline in foodborne illness hospitalizations to a 10% decline. In that scenario, if the drivers of outcome effects were doubled (that is, amplified by 100%) and the drivers of daily schedule effects were fully mitigated (that is, reduced by 100%), the 11.03% increase in citations (last row, Column 2) would translate into a 22.06% decline in hospitalizations [=11.03*(-10/5)] (compared to our original estimate of a 44.12% decline, calculated as 11.03*(-20/5) and reported in the last row of Column 4), which nationwide would result in 28,235 fewer foodborne-illness-related hospitalizations and 10.54 million fewer foodborne illness cases, saving \$7.88 billion to \$17.14 billion in foodborne illness costs.

Table B1. Estimates of Nationwide Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Impact on citations of violations			Impact on health and associated costs				
Bias-reduction scenario	Average predicted number of violations cited per inspection	Percent change in average predicted number of violations cited compared to the status quo	Change in nationwide annual number of violations cited compared to the status quo	Percent change in foodborne illness hospitalizations compared to the status quo	Change in nationwide annual number of foodborne illness hospitalizations compared to the status quo	Change in nationwide annual number of foodborne illness cases compared to the status quo, in millions	Change in nationwide annual costs of foodborne illness cases compared to status quo, in \$millions	
							Lower estimate	Upper estimate
0% (status quo)	2.424	0.00%	0	0.00%	0	0.00	\$0	\$0
10%	2.448	1.01%	24,536	-4.05%	-5,183	-1.94	-\$1,446	-\$3,147
20%	2.473	2.04%	49,528	-8.17%	-10,463	-3.91	-\$2,919	-\$6,353
30%	2.499	3.09%	74,987	-12.38%	-15,841	-5.92	-\$4,419	-\$9,619
40%	2.525	4.16%	100,924	-16.66%	-21,320	-7.96	-\$5,947	-\$12,946
50%	2.551	5.25%	127,352	-21.02%	-26,903	-10.05	-\$7,505	-\$16,336
60%	2.578	6.37%	154,283	-25.46%	-32,593	-12.17	-\$9,092	-\$19,790
70%	2.605	7.50%	181,732	-29.99%	-38,391	-14.34	-\$10,710	-\$23,311
80%	2.633	8.65%	209,711	-34.61%	-44,302	-16.54	-\$12,358	-\$26,900
90%	2.662	9.83%	238,234	-39.32%	-50,327	-18.79	-\$14,039	-\$30,559
100%	2.691	11.03%	267,315	-44.12%	-56,471	-21.09	-\$15,753	-\$34,290

Each row represents a bias-reduction scenario. For example, the 10% scenario depicted in the second row illustrates the results of reducing bias if the outcome effects (which increase scrutiny) were amplified by 10% and the daily schedule effects (which erode scrutiny) were mitigated by 10%.

Column 1 is the average predicted number of violations per inspection, based on Model 1 of Table 2, under each scenario.

Column 2 is calculated as the percent change in the average predicted number of violations per inspection, comparing each scenario (Column 1) to the status quo value of 2.42365.

Column 3 is calculated as the difference in the average predicted number of violations per inspection, comparing each scenario (Column 1) to the status quo value of 2.42365 and multiplying this by the one million food safety inspections conducted nationwide each year.

Column 4 is calculated by multiplying the percent change in average predicted number of violations compared to the status quo (Column 2) by the ratio of the change in hospitalizations to the change in compliance (derived from the 20% decline in hospitalizations per 5% improvement in restaurant compliance relationship reported by Jin and Leslie (2003); that is, $-20\%/5\% = -4$).

Column 5 is calculated as the difference in hospitalizations between (a) the estimated number that would have occurred under each scenario and (b) the 128,000 that actually occurred (Scallan et al. 2011). Specifically, we multiply the percent change in hospitalizations (Column 4) by the 128,000 nationwide annual hospitalizations.

Column 6 is calculated by multiplying the change in nationwide annual number of foodborne illness hospitalizations compared to the status quo (Column 5) by 373.4, the number of illness cases per hospitalization (calculated as the ratio between Scallan et al.'s (2011) two estimates of the 47.8 million annual foodborne illnesses and the resulting 128,000 hospitalizations).

Columns 7 and 8 are calculated by multiplying the estimated change in illness cases (Column 6) by \$747 (the weighted average from Minor et al. (2015)) and by \$1,626 (the enhanced model estimate from Scharff (2012)) in estimated costs per illness case, respectively.

Figure B1. Estimated nationwide increase in food safety violations being cited as biases are attenuated

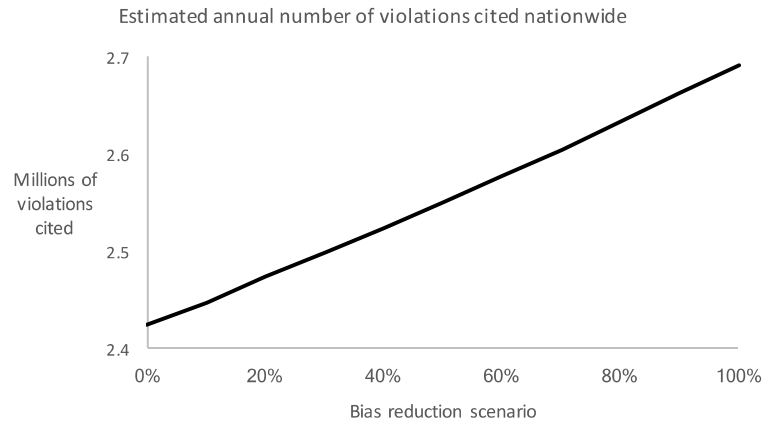


Figure B2. Estimated reduction in healthcare cases associated with more food safety violations being cited as biases are attenuated

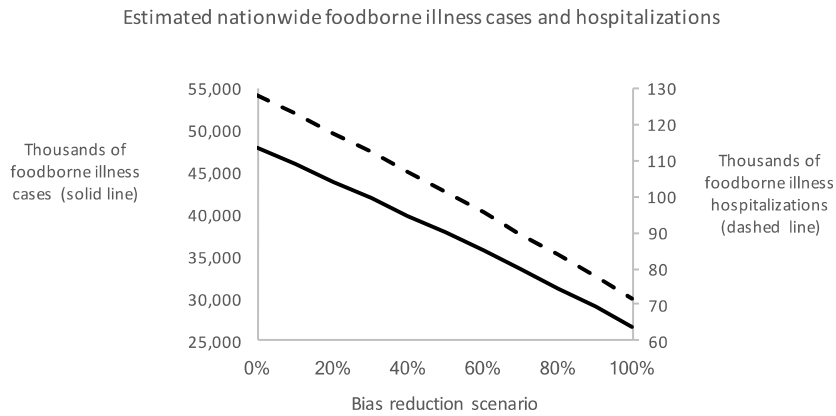
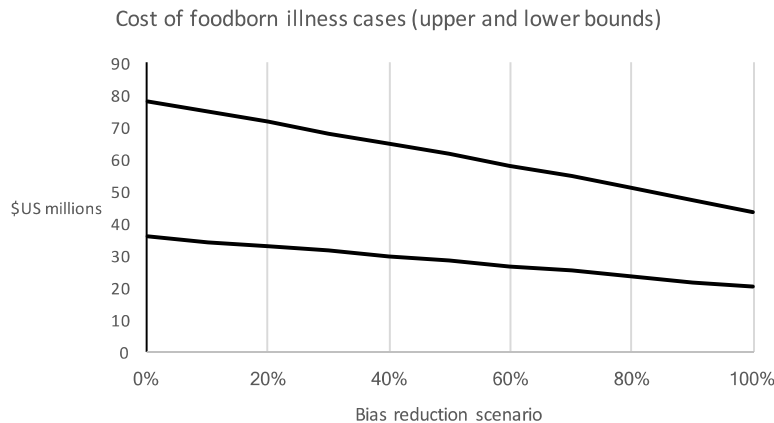


Figure B3. Estimated cost reductions associated with improved health impacts resulting from more food safety violations being cited as biases are attenuated



Notes: These figures graph data from Table B1, which is based on the methodology described in Appendix B. The horizontal axes represent different bias-reduction scenarios. For example, the 20% scenario illustrates the results of reducing bias by amplifying by 20% the outcome effects (which increase scrutiny) and mitigating by 20% the daily schedule effects (which erode scrutiny).

Appendix C. Supplemental Analysis: Alternative Measures of Potentially Shift-prolonging (H5) Construct

Our primary results test the effects on inspector scrutiny of conducting inspections that risk prolonging the inspector's shift (H5). As discussed in the paper (Section 4.7, "Extensions"), we also investigate whether the extent to which the shift might be prolonged matters. Results are reported in Table C1 below. First, we use a continuous variable instead of the binary *potentially shift-prolonging* to examine how the magnitude of shift-prolonging affects inspector scrutiny (Columns 3–4 of Table C1). We also pursued a more flexible approach by creating a series of dummy variables denoting ranges of *potential extent of shift-prolonging* (Columns 5–6 of Table C1). These specifications yield the same inferences as our main models, which are reported in the paper in Table 2 and are reproduced in Columns 1–2 of Table C1 as the basis of comparison for the following robustness tests.

.

Table C1. Alternative Measures of Potentially Shift-prolonging (H5) Construct

Measure of potentially shift-prolonging (H5): Dependent variable:	<i>Potentially shift-prolonging (binary)</i> (our primary specification)		<i>Potential extent of shift-prolonging</i> (in hours)		<i>Potential extent of shift-prolonging</i> indicators	
	<i>violations</i>	<i>violations</i>	<i>violations</i>	<i>violations</i>	<i>violations</i>	<i>violations</i>
	(1)	(2)	(3)	(4)	(5)	(6)
Prior inspected establishment's violations (H1)	0.017*** (0.004)	0.015*** (0.004)	0.017*** (0.004)	0.015*** (0.004)	0.017*** (0.004)	0.015*** (0.004)
Prior inspected establishment's violation trend (H2)	0.013** (0.006)		0.013** (0.006)		0.013** (0.006)	
Prior inspected establishment saliently improved (H3)		0.011 (0.023)		0.012 (0.023)		0.012 (0.023)
Prior inspected establishment saliently deteriorated (H3)		0.079*** (0.027)		0.079*** (0.027)		0.079*** (0.027)
Time inspecting earlier today (H4)	-0.038*** (0.011)	-0.039*** (0.011)	-0.034*** (0.011)	-0.035*** (0.011)	-0.036*** (0.011)	-0.036*** (0.011)
Potentially shift-prolonging (binary) (H5)	-0.051** (0.025)	-0.050** (0.025)				
Potential extent of shift-prolonging (in hours)			-0.068*** (0.023)	-0.068*** (0.023)		
Potential extent of shift-prolonging is >0.5 & <=1 hours (binary)					-0.050 (0.040)	-0.051 (0.040)
Potential extent of shift-prolonging is >1 & <=1.5 hours (binary)					-0.078* (0.045)	-0.078* (0.045)
Potential extent of shift-prolonging is >1.5 & <= max of 1.79 hours (binary)					-0.109** (0.044)	-0.109** (0.044)
Inspector experience	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Returning inspector	-0.114*** (0.035)	-0.117*** (0.035)	-0.115*** (0.035)	-0.117*** (0.035)	-0.116*** (0.035)	-0.118*** (0.035)
Lunch period (11:00 am–3:59 pm)	-0.049** (0.024)	-0.049** (0.024)	-0.054** (0.024)	-0.053** (0.024)	-0.055** (0.024)	-0.054** (0.024)
Month fixed effects	Included	Included	Included	Included	Included	Included
Year fixed effects	Included	Included	Included	Included	Included	Included
Establishment's <i>n</i> -th-inspection (2nd through 10th or more) fixed effects	Included	Included	Included	Included	Included	Included
Inspection-type fixed effects	Included	Included	Included	Included	Included	Included
Inspector-establishment dyad fixed effects	Included	Included	Included	Included	Included	Included
Number of observations (inspections)	12,017	12,017	12,017	12,017	12,017	12,017

Notes: Poisson regression coefficients with robust standard errors clustered by establishment. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix D. Supplemental Analysis: Persistence of Outcome Effects

We conduct additional analysis to examine the persistence of some of our outcome effects (Table D1; see Section 4.7, “Extensions”).

Table D1. Persistence of Outcome Effects

Dependent variable:	<i>violations</i>		
	(1)	(2)	(3)
Prior inspected establishment’s violations	0.017*** (0.004)	0.017*** (0.004)	
Penultimate inspected establishment’s violations	0.010*** (0.003)	0.010*** (0.003)	
Antepenultimate inspected establishment’s violations		0.007* (0.004)	
Prior inspected establishment’s violations for the first inspection of the day			0.016*** (0.006)
Prior inspected establishment’s violations for the second+ inspection of the day			0.018*** (0.005)
Prior inspected establishment’s violation trend	0.013** (0.006)	0.013** (0.006)	0.014** (0.006)
Time inspecting earlier today	-0.038*** (0.011)	-0.039*** (0.011)	-0.040*** (0.012)
Potentially shift-prolonging	-0.055** (0.024)	-0.056** (0.024)	-0.051** (0.025)
Inspector experience	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Returning inspector	-0.118*** (0.035)	-0.120*** (0.034)	-0.114*** (0.035)
Lunch period (11:00 am–3:59 pm)	-0.046* (0.024)	-0.046* (0.024)	-0.051** (0.025)
Month fixed effects	Included	Included	Included
Year fixed effects	Included	Included	Included
Establishment’s-nth-inspection (2nd through 10th or more) fixed effects	Included	Included	Included
Inspection-type fixed effects	Included	Included	Included
Inspector-establishment dyad fixed effects	Included	Included	Included
Number of observations (inspections)	12,011	12,000	12,017

Notes: Poisson regression coefficients with robust standard errors clustered by establishment. Columns 1 and 2 have fewer observations than our main estimation sample because inspectors’ *Penultimate inspected establishment’s violations* is missing for the inspectors’ first and second inspections and inspectors’ *Antepenultimate inspected establishment’s violations* is missing for the inspectors’ first, second and third inspections. Most of those observations were already dropped by other sample restrictions.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix E. Supplemental Analysis: Breaks

We conduct additional analysis to explore how breaks could affect inspector scrutiny (Table E1; see Section 4.7, “Extensions”).

Table E1. Breaks

	Dependent variable:		
	<i>violations</i>		
	(1)	(2)	(3)
H1 Prior inspected establishment’s violations (H1)	0.017*** (0.004)	0.014*** (0.005)	0.017*** (0.004)
H2 Prior inspected establishment’s violation trend (H2)	0.014** (0.006)	0.018** (0.007)	0.014** (0.006)
H4 Time inspecting earlier today (H4)	-0.034*** (0.012)	-0.034*** (0.012)	-0.036*** (0.012)
H5 Potential extent of shift-prolonging (H5)	-0.055** (0.025)	-0.055** (0.025)	-0.056** (0.025)
Inspector experience	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Returning inspector	-0.117*** (0.035)	-0.117*** (0.035)	-0.117*** (0.035)
Lunch period (11:00 am–3:59 pm)	-0.051** (0.025)	-0.050** (0.025)	-0.049** (0.024)
Overnight break length	0.003* (0.002)	0.002 (0.002)	
Within-day break length	0.089 (0.124)	0.051 (0.148)	
Prior inspected establishment’s violations (H1) × × Overnight break length		0.000 (0.001)	
Prior inspected establishment’s violations (H1) × × Within-day break length		0.032 (0.054)	
Prior inspected establishment’s violation trend (H2) × × Overnight break length		-0.001 (0.001)	
Prior inspected establishment’s violation trend (H2) × × Within-day break length		-0.061 (0.093)	
Break length			0.003* (0.002)
Time inspecting earlier today (H4) × Break length			0.062 (0.095)
Month fixed effects	Included	Included	Included
Year fixed effects	Included	Included	Included
Establishment’s- <i>n</i> th-inspection (2nd through 10th or more) fixed effects	Included	Included	Included
Inspection-type fixed effects	Included	Included	Included
Inspector-establishment dyad fixed effects	Included	Included	Included
Number of observations (inspections)	12,017	12,017	12,017

Notes: Poisson regression coefficients with robust standard errors clustered by establishment.

*** p < 0.01, ** p < 0.05, * p < 0.10.

Appendix F. Supplemental Analysis: Critical versus Noncritical Violations

To assess whether our hypothesized relationships differentially influence inspectors' behavior across different types of violation, we estimated our models on two subsets of violations. First, we predict the number of *critical violations*, which are related to food preparation practices and employee behaviors that more directly contribute to foodborne illness or injury. These factors are prioritized in Alaska and in Camden County by being displayed on the first page of the inspection report and in Lake County by being tagged in the reports. Second, we estimated our models on the number of *noncritical violations* (that is, violations of procedures often referred to as "good retail practices"). While less risky than the other type, these are also important for public health and include overall sanitation and preventative measures to protect foods, such as proper use of gloves. Inspections averaged 0.93 *critical violations* and 1.49 *noncritical violations*.

Fewer noncritical violations are cited in inspections conducted during the lunch period than in other periods, but the results yield no evidence that time of day affects critical violations (see Table F1). The latter finding is consistent with critical violations being related to longer-term establishment practices that are insensitive to the number of customers being served or to the staff's busyness and thus ability to respond to the inspector's presence.

Outcome effects are ubiquitous, affecting critical and noncritical violations alike. Each additional violation cited at the inspector's prior inspected establishment is associated with 1.92% more critical violations (Column 1: $\beta = 0.019$, $p < 0.01$) and 1.61% more noncritical violations (Column 3: $\beta = 0.016$, $p < 0.01$) cited in the focal inspection.

As with total violations, there is no evidence of critical and noncritical violations being affected when the *prior inspected establishment saliently improved*. When the *prior inspected establishment saliently deteriorated*, inspections yield, on average, 7.57% more critical violations (Column 2: $\beta = 0.073$, $p < 0.10$) and 8.22% more noncritical violations (Column 4: $\beta = 0.079$, $p < 0.05$).

Turning to daily schedule effects, we find that fatigue affects inspectors' ability to discover and report both types of violations. Specifically, the estimated coefficients on *time inspecting earlier today* indicate that each additional hour conducting prior inspections during the day results, on average, in 2.86% fewer critical violations cited (Column 1: $\beta = -0.029$, $p < 0.10$) and 4.30% fewer noncritical violations cited (Column 3: $\beta = -0.044$, $p < 0.01$).

These results also indicate that the *potentially shift-prolonging* effects identified in our primary results are driven by noncritical violations rather than critical ones. In particular, *potentially shift-prolonging* inspections result in 6.29% fewer citations (Column 3: $\beta = -0.065$, $p < 0.01$). However, we find no evidence that citations of critical violations are affected by whether the inspection risks

prolonging the shift: the coefficient on *potentially shift-prolonging* is not statistically significant when predicting critical violations (Columns 1 and 2). This suggests that avoiding prolonging the shift does not affect inspectors' ability to discover and report critical violations.

Overall, these results indicate that inspectors' schedules have somewhat different effects on citing critical versus noncritical violations. Citing noncritical violations appears to be influenced by all types of daily schedule effects and outcome effects, while citing critical violations appears to be influenced by all but the *potentially shift-prolonging* effects.

Table F1. Critical and Noncritical Violations

Dependent variable:		<i>critical violations</i>		<i>noncritical violations</i>	
		(1)	(2)	(3)	(4)
H1	Prior inspected establishment's violations	0.019*** (0.005)	0.016*** (0.006)	0.016*** (0.005)	0.015*** (0.005)
H2	Prior inspected establishment's violation trend	0.014* (0.008)		0.013* (0.007)	
H3	After salient improvement		-0.017 (0.031)		0.028 (0.028)
H3	After salient deterioration		0.073* (0.039)		0.079** (0.031)
H4	Time inspecting earlier today	-0.029* (0.015)	-0.030* (0.015)	-0.044*** (0.013)	-0.045*** (0.013)
H5	Potentially shift-prolonging	-0.026 (0.034)	-0.025 (0.034)	-0.065** (0.029)	-0.064** (0.029)
	Inspector experience	0.001*** (0.000)	0.001*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
	Returning inspector	-0.113** (0.049)	-0.115** (0.049)	-0.102** (0.041)	-0.104** (0.041)
	Lunch period (11:00 am–3:59 pm)	-0.038 (0.032)	-0.037 (0.032)	-0.057** (0.028)	-0.056** (0.028)
	Month fixed effects	Included	Included	Included	Included
	Year fixed effects	Included	Included	Included	Included
	Establishment's- <i>n</i> th-inspection (2nd through 10th or more) fixed effects	Included	Included	Included	Included
	Inspection-type fixed effects	Included	Included	Included	Included
	Inspector-establishment dyad fixed effects	Included	Included	Included	Included
	Number of observations (inspections)	10,298	10,298	10,624	10,624

Notes: Poisson regression coefficients with robust standard errors clustered by establishment.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix G. Supplemental Analysis: Inspector Speed and Citation Pace

Our primary results show how inspections of prior establishments and daily schedules are associated with the number of *violations* cited. To assess whether such results might be driven by inspectors spending more or less time and exhibiting more or less scrutiny in the subsequent (focal) inspection, we estimate our primary models on the log of *inspection duration*, the number of minutes between an inspection's start time and end time. Moreover, to assess the net effect of the changes in *violations* cited and *inspection duration*, we explore the inspector's citation pace—a measure of productivity in this setting—and estimate our primary models on the log (after adding 1) of *violation citations per hour*. The results are reported in Table G1.

Considering potential outcome effects, we find that inspectors spend only slightly more time conducting inspections succeeding inspections in which more violations were cited (Column 1: *prior inspected establishment's violations* $\beta = 0.004$, $p < 0.10$) and find no evidence that the violation trend of the inspector's prior inspection affects inspection duration (Columns 1 and 2: the estimated coefficients on the *prior inspected establishment's violation trend*, *prior inspected establishment saliently improved* and *prior inspected establishment saliently deteriorated* are not statistically significant). Recall that our primary results found that more violations or worsening trends at an inspector's prior establishment predicted more violations cited at the focal inspection. Results in Column 3 indicate that citation pace increases by 1.0% for each additional violation at the prior establishment ($\beta = 0.010$, $p < 0.01$) and by 1.9% for each one-standard-deviation increase in the *prior inspected establishment's violation trend* ($\beta = 0.012$, $p < 0.05$). Column 4 indicates that, as was the case with the number of violations, this effect is asymmetric and driven by negative trends: whereas we find no change in citation pace after inspecting an establishment with salient improvement, it does increase by 3.9% after inspecting an establishment with salient deterioration (Column 4: $\beta = 0.039$, $p < 0.10$). This indicates that our main outcome-effect findings—that more violations and worsening trends at an inspector's prior establishment increase the inspector's citations at his or her next inspection—result mostly from inspectors increasing their citation pace rather than spending more time onsite.

We next consider potential daily schedule effects. We find that inspectors conduct inspections more quickly as they progress through their shift: *inspection duration* decreases by 3.1% for each additional hour already spent conducting inspections that day (Column 1: *time inspecting earlier today* $\beta = -0.031$, $p < 0.01$). For context, recall that our primary results indicate that each additional hour inspecting during the day cites an average of 3.73% fewer violations. The model reported in Column 3 indicates that the net effect is that inspector citation pace decreases by 1.9% for each subsequent inspection of the day (*time inspecting earlier today* $\beta = -0.019$, $p < 0.05$).

Turning to *potentially shift-prolonging* inspections, recall that our primary results indicated that these had 5.0% fewer citations. Column 1 reveals that inspectors conduct such inspections 4.3% more quickly (*potentially shift-prolonging* $\beta = -0.044$, $p < 0.01$). Column 3 reveals that the effect of *potentially shift-prolonging* on citation pace is not statistically significant. These results jointly suggest that the diminishment in citations results from shorter inspection duration rather than slower inspector speed, with inspectors' citation pace remaining largely unaffected by the risk of working late. This, in turn, suggests that our earlier finding that *potentially shift-prolonging* inspections result in fewer violations is likely due to inspectors' desire to avoid working late, rather than to fatigue eroding their citation pace.

Table G1. Effects of Inspectors' Schedules on Speed and Citation Pace

Dependent variable:	Inspector speed		Inspector citation pace	
	log inspection duration		log (violation citations per hour + 1)	
	(1)	(2)	(3)	(4)
Prior inspected establishment's violations	0.004* (0.002)	0.003* (0.002)	0.010*** (0.003)	0.010*** (0.003)
Prior inspected establishment's violation trend	-0.000 (0.003)		0.012** (0.005)	
Prior inspected establishment saliently improved		0.009 (0.009)		-0.017 (0.016)
Prior inspected establishment saliently deteriorated		0.006 (0.012)		0.039* (0.020)
Time inspecting earlier today	-0.031*** (0.005)	-0.031*** (0.005)	-0.019** (0.008)	-0.020*** (0.008)
Potentially shift-prolonging	-0.044*** (0.011)	-0.044*** (0.011)	-0.016 (0.018)	-0.015 (0.018)
Inspector experience	0.000 (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000*** (0.000)
Returning inspector	0.082*** (0.016)	0.082*** (0.016)	-0.100*** (0.024)	-0.100*** (0.024)
Lunch period (11:00 am–3:59 pm)	-0.034*** (0.010)	-0.034*** (0.010)	0.010 (0.016)	0.010 (0.016)
Month fixed effects	Included	Included	Included	Included
Year fixed effects	Included	Included	Included	Included
Establishment's- <i>n</i> th-inspection (2nd through 10th or more) fixed effects	Included	Included	Included	Included
Inspection-type fixed effects	Included	Included	Included	Included
Inspector-establishment dyad fixed effects	Included	Included	Included	Included
Number of observations (inspections)	12,017	12,017	12,017	12,017
R-squared	0.45	0.45	0.20	0.20

Notes: Ordinary least squares coefficients with robust standard errors clustered by establishment.
 *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

References

- Jin G.Z., Leslie P. 2003. The effect of information on product quality: Evidence from restaurant hygiene grade cards. *Quarterly Journal of Economics* **118**(2) 409-451.
- Scallan E., Griffin P.M., Angulo F.J., Tauxe R.V., Hoekstra R.M. 2011. Foodborne illness acquired in the United States—unspecified agents. *Emerging Infectious Disease journal* **17**(1) 16.
- Scharff R.L. 2012. Economic burden from health losses due to foodborne illness in the United States. *Journal of Food Protection* **75**(1) 123-131.