

Designing Information Disclosure

By Daniel E. Ho*

Information disclosure is a central tool of regulatory policy. While academic research highlights how consumers comprehend information, it largely glosses over the institutional design of information-disclosure schemes. Empirical evidence from over 700,000 restaurant health inspections in 10 jurisdictions shows that the poster child of information disclosure—restaurant letter grading—fails when institutional design is ignored.

Information Disclosure

As any administrative lawyer can attest, mandated information disclosure is pervasive pervades in the regulatory state. Decades of research, however, also show that such disclosures are often grossly ineffective. Safe Drinking Water Act reports are notoriously complex—disclosing a litany of contaminant test results (e.g., 5–67 ppb of barium)—and hence unintelligible to the typical consumer. Caloric disclosures can induce consumers to eat a higher quantity of low-caloric foods—the so-called “Snackwell effect.” And cardiac surgery report cards can affect the willingness of physicians to treat sicker patients.

The emerging consensus in social science centers on the idea of a “targeted nudge.” Disclosures should take the form of simplified signals that summarize the underlying complex information and are disseminated to consumers at the time of decisionmaking. Such targeted nudges enable consumers to act in an informed fashion.

Restaurant letter grading is a prime example. Such letter grades (*A*, *B*, or *C*) in principle summarize the health risk of a restaurant, as assessed by inspectors during unannounced on-site inspections

that typically last one to two hours. Grades are posted in the entryway of restaurants enabling consumers to choose restaurants based on their sanitation risk. In one study of Los Angeles, researchers found that letter grading reduced hospitalizations for food-borne illness by 20%. Research comparing disclosure regimes across regulatory areas, based on the Los Angeles findings, concludes that restaurant grading is “highly effective” due to the “brief, simple, easy disclosures.” *Forbes* magazine dubs it as “[t]he most effective regulatory disclosure ever.”

Yet is it really effective?

Empirically Grading Restaurant Grading

In the United States, restaurant food safety is typically regulated at the local level. The institutional design features of local environmental health agencies—such as the degree of inspector specialization, the rules for scoring food code violations, the timing of inspections, and the thresholds for letter grades where they exist—vary widely across jurisdictions.

To assess the effectiveness of letter grading, my research team collected data from over 700,000 restaurant inspections of over 100,000 restaurants in 10 jurisdictions. For simplicity, we focus here only on San Diego and New York, but the results generalize. Our main findings are threefold.

Grade Inflation

Jurisdictions that grade exhibit rampant grade inflation. In San Diego, inspectors score restaurants on a 0–100 scale, with 90 or more points resulting in an *A* grade. Out of some 9,000 restaurants, all but eight received *A*’s, with many falling just above the cutoff.

While it is possible that restaurants are cleaning up to precisely target the *A* threshold, there are reasons to doubt this. First, inspectors exercise considerable discretion in the scoring of individual violations. A typical requirement, for example, is that “food contact surfaces

are clean and sanitized,” which may be assessed in different ways. Such discretion likely leads some inspectors to bump up borderline restaurants to an *A* grade. As one San Diego inspector noted: “Some inspectors will give out a *B* for an 89. I usually warn somebody at that point. It’s a judgment call.”

Second, San Diego allows restaurants to pay roughly \$140 to be reinspected and regraded within days. Nearly a quarter of San Diego restaurants scoring below 90 achieve an *A* within one day of the original inspection and 80% do so within a month. Third, when New York introduced grading in July 2010, scores similarly began exhibiting sharp differences at the cutoff, with many more restaurants barely receiving *A*’s. And contrary to the notion that restaurants were cleaning up in response to the grades, average scores, if anything, got worse.

In short, grade inflation appears rampant, suggesting that grades provide little meaningful information that would allow consumers to distinguish between restaurants.

Consistency

For grading to be effective, the grade must minimally convey some information about the sanitation level of the restaurant. To assess this, we examined the consistency of underlying inspection scores across repeat, unannounced inspections of the same establishment. In most jurisdictions, these scores exhibit some degree of substantive consistency. A San Diego restaurant receiving a score of 90, for example, will most likely also receive a score around 90 in the next inspection cycle. Overall, roughly 25% of the variation in San Diego’s scores is explained by scores in the prior inspection cycle. This makes sense if inspection scores measure sanitation levels that are a persistent attribute of the specific restaurant, a predicate assumption of grading.

New York, however, is an outlier. While New York restaurants exhibit substantial

continued on next page

* Professor, Stanford Law School. Email: dho@law.stanford.edu. This article provides an overview of the author’s full article, *Regulatory Fudge: The Promise of Targeted Transparency and the Practice of Restaurant Grading*, 122 YALE L.J. (forthcoming). Thanks to Patrick Leahy and Laura Trice for helpful comments.

variation in grades (roughly two thirds receive *A*'s), the underlying inspection scores have no substantial predictive power over future scores. Prior scores predict less than 2% of the score variation. New York grades, in other words, do not appear to summarize any persistent attribute of restaurant sanitation.

Why does New York exhibit such inconsistency in its inspections? The best answer may lie in institutional design. First, New York's cadre of roughly 180 inspectors does not specialize exclusively in restaurant inspections. Inspectors are responsible for dozens of different types of inspections, such as for senior centers and summer camps. Second, the set of rules governing the scoring process is more complex than in any other jurisdiction we encountered. New York inspectors score more violations, with a wider point range (relative to the grade threshold), than elsewhere. San Diego, for example, records a single violation for vermin. New York records separate violations for (i) "[e]vidence of rats or live rats," (ii) "[e]vidence of mice or live mice," (iii) "[l]ive roaches," and (iv) "flies," scored at 5, 6, 7, 8, or 28 points, depending on the evidence. (New York's cut-off to receive an *A* is less than 14 points.) Thirty "fresh mice droppings in one area" result in six points, but thirty-one droppings result in seven points.

The result is that inspectors implement restaurant inspections in sharply different ways. As one public health official (albeit not based out of New York) noted: "it is the norm for every single person to do an inspection differently." A 2009 audit by the City Comptroller, which documented considerable data errors and a failure to supervise and monitor inspectors, showed that of 67 inspectors who conducted at least 100 inspections per year, one inspector averaged 15 points and another 50. And in a raucous oversight hearing in March 2012, N.Y. City Council Speaker Christine Quinn focused on the "inconsistent" implementation of restaurant inspections.

In short, New York's grades appear to have little meaning.

Resource Allocation

Perhaps the most pernicious feature of institutional design is that letter grading

appears to shift inspection resources away from the worst offenders.

Prior to grading, New York's Department of Health would conduct "compliance inspections" of restaurants posing the greatest risks—those now in the *C*-range. While some compliance inspections still occur, the grading system introduced "reinspections" for restaurants scoring less than an *A* to have a second chance to obtain a better grade. (When the Department first proposed grading, it did not include provisions for reinspections, but ultimately succumbed to restaurateur pressure.) Such reinspections occur roughly a month after the initial inspection and restaurants do not have to post a grade during the interim period. Some 14,000 inspections in the first year and a half of the grading system have been devoted to reinspections for grade resolution of *B*-range restaurants.

Grading thereby consumes resources that could have been deployed at restaurants with more egregious violations.

Institutional Design

Design and implementation matters. The underlying idea of letter grading may be well conceived. Indeed, letter grades are the rare example of a disclosure that appears to have a substantial effect on consumers. But targeted nudges fail without a well-designed public institutions to generate the underlying information.

How could we do better?

First, rules can be too detailed. Although New York's scoring system was designed to make the inspection process more objective and informative, the opposite has occurred. As has also been shown in the context of nursing home inspections, highly detailed rules (in lieu of standards) can impede the consistency and reliability of inspections. (The irony is that while proponents of grading prefer simplification for consumers, they ignore it for information producers.) In jurisdictions like New York, rules need to be simplified to reduce differences across inspectors.

Second, reinspections, which typically accompany grading systems, are ineffective. Restaurateurs know with relative precision when such reinspections occur: in New York, roughly one month after

the initial inspection; in San Diego, within days upon payment. Strategic cleanups in anticipation of such reinspections undermine the meaningfulness of sanitation scores and reinspections improperly draw resources away from the worst offenders. Grades should be based on truly random inspections to provide a meaningful measure of a restaurant's sanitation level.

Third, jurisdictions should fight grade inflation. One way to do so is by setting higher cut-offs for *A* grades and disclosing the proportion of restaurants achieving each grade.

Fourth, letter grade disclosures should account for the uncertainty in the inspection process. Well-known statistical adjustments can account for factors, such as inspector differences, as well as aid agencies in training, supervising, and monitoring inspectors. And some measure of uncertainty—akin to the margin of error in a presidential poll—should be disclosed to consumers.

Last, agencies should make full inspection data freely available. The only way to assess grading is to enable researchers to study its operation. Comprehensive data disclosure will also empower information intermediaries to find more effective ways to disseminate information to consumers. A website like Yelp, which aggregates information and ratings of local businesses and reaches roughly 66 million unique visitors per month, could, for example, include health inspection data in its restaurant characteristics, even in jurisdictions that currently do not mandate letter grade posting.

The Risk of False Promises

Our study shows that even the perceived paragon of information disclosure can be seriously flawed in implementation. Targeted nudges cannot solve or avoid the core administrative law issues—the institutional design of inspection agencies, the development of administrable rules and standards, and the accountability and oversight of expert agents. Without such elements in place, health inspections cannot generate meaningful information and targeted nudges risk turning into a facile mantra of regulatory reform. 