ARE THERE "RATATOUILLE" RESTAURANTS? ON ANTICORRELATION OF FOOD QUALITY AND HYGIENE

Hisayuki Yoshimoto and Andriy Zapechelnyuk

Abstract. We study the empirical relationship between restaurants' hygiene standards and their food quality scores, as evaluated by professional reviewers. By using data from the UK high-end restaurants, we show that this relationship is negative, observed across several econometric specifications and food quality measurements. We report that 3% of Michelin-starred restaurants have poor hygiene, while the same is true for 2.5% of high-end guidebook-listed restaurants in our dataset. We highlight two possibilities for this observed negative association: a strategic hypothesis (capturing restaurants' choices trading off hygiene for quality), and a selection hypothesis (reflecting restaurants' differential survival rates under competition). Our results indicate that the latter has more support. Our findings also illuminate potential channels through which the anticorrelation between hygiene and food quality could be mitigated and can be informative for hygiene inspectors in order to prioritize restaurants in their inspection schedule based on observable characteristics.

Keywords: Food quality, restaurants, consumer reviews, hygiene standards, certification

JEL Classification: L15, D22, I18

Date: October 22, 2024

Yoshimoto: Adam Smith Business School, University of Glasgow, University Avenue, Glasgow G12 8QQ, UK. *E-mail:* Hisayuki.Yoshimoto@glasgow.ac.uk.
Zapechelnyuk: School of Economics and Finance, University of Edinburgh, 31 Buccleuch Place, Edinburgh EH8 9JT, UK. *E-mail:* andy.zapechelnyuk@ed.ac.uk.

The authors express their special gratitude to George Deltas and an anonymous referee whose insightful comments helped to significantly improve the paper. The authors are grateful to Sven Anders, Toman Barsbai, Mark Daniel Bernhardt, Xiaogang Che, Dakshina De Silva, Elissa Philip Gentry, David Jaeger, Margaret Leighton, Marco Manacorda, David Rietzke, and Arjunan Subramanian for helpful comments and suggestions. The authors also thank participants of International Industrial Organization Society 2020 Conference in Philadelphia (virtual), Econometric Society 2021 Asian Meeting in Malaysia (virtual), European Association for Research in Industrial Economics 2021 Conference in Bergen (virtual), Royal Economic Society 2023 Conference in Glasgow, and Auctions Competition Regulation and Public Policy 2023 Conference in Lancaster, as well as at various seminars, where this paper has been presented. Financial support from the Economic and Social Research Council Grant ES/N01829X/1 is gratefully acknowledged.

1. Introduction

All across the globe, when choosing where to have a meal, diners rely on food quality reviews. These reviews come from a variety of sources, including popular guidebooks such as Michelin and Good Food Guide. Diners are after a quality meal experience, so the last thing they want is to end up sick with food poisoning. It is reported that, in the UK, "Food hygiene when eating out" tops the list of public concerns related to food safety, according to the Biannual Public Attitudes Tracker Survey 2019 published by the UK Food Standards Agency.[1]

Diners who rely on food restaurant review scores would expect these scores to be informative about hygiene standards, and thus about the likelihood of contracting a food-borne disease. In the current study, we examine the validity of this expectation. Using data from the UK high-end restaurants, across various econometric specifications, we show that there is a statistically significant conditional negative correlation between hygiene scores and food quality ratings, in which the food quality is measured by professional reviewers. In particular, we find that there are *Ratatouille* restaurants: highly rated establishments with poor hygiene practices, a concept we name after a rat-infested gourmet restaurant featured in the animated movie "Ratatouille" (Pixar & Walt Disney Pictures, 2007). Among all UK restaurants listed in the Good Food Guide, 3% of Michelin-starred restaurants have poor hygiene conditions categorized as "Improvement Necessary" or worse, while the same is true for 2.5% of guidebook-listed restaurants in our dataset.[2] The observation is that, under several econometric specifications, restaurant review scores from the well-known guidebooks are indicative of food hygiene standards, but, unfortunately, in a negative way, which is likely to be misleading for consumers.

To address the above research question, we use the data on UK-based restaurants to analyze how the hygiene conditions are related to food quality as rated by professional reviewers. The hygiene certification data comes from the UK Food Standards Agency. The professional quality review data is obtained from two popular restaurant guidebooks, *The Michelin Guide* and *The Good Food Guide*. Other restaurant characteristics are obtained from *Tripadvisor*.

To investigate the empirical relationship between hygiene and food quality, we use the Poisson, Negative Binomial (NB), hurdle, and selection models for count-dependent variable regressions, examining various possibilities of market environments. We show that a robustly negative conditional correlation between hygiene and food quality. Our estimation results also provide a quantitative basis for the analysis of the relationship between the control variables and the hygiene generating process. These results can be informative for

[1]Available here: https://www.food.gov.uk/about-us/biannual-public-attitudes-tracker

[2]As an anecdotal example, see https://www.businessinsider.com/michelin-restaurants-with-c-health-inspection-rating-2013-12.

inspection authorities in understanding the market background of restaurants and can also be useful for designing hygiene inspection policies, such as the prioritization of inspections based on observable restaurant characteristics.

We investigate two main possibilities that can lead to a negative association between food quality and hygiene in restaurants. One is that restaurant hygiene is a choice made by the restaurants in response to the market environment they find themselves in, which we refer to as the *strategic hypothesis*. The other is that restaurant food quality and hygiene are part of the restaurant's "DNA", its intrinsic capabilities, that are fundamentally uncorrelated. However, the market competition environment yields different survival probabilities for restaurants of different food quality and hygiene combinations, and in equilibrium we observe a negative association between these two restaurant attributes. We refer to this as the *selection hypothesis*. Our estimation results are consistent with the selection hypothesis, as we discuss in Section 4.7.

This paper makes a few contributions to the literature. First, we empirically study the various possibilities of market environments affecting hygiene standards, notably their relations to food quality. Our empirical analysis highlights potential channels, involving the competitive market background, such as segmented markets, as well as selections to be operated in a competitive area. Particularly, related to the anticorrelation between hygiene and food quality, we report largely different empirical findings without considering the competitive location selection, shedding light on the restaurant-side economic behavior related to hygiene generating processes. Second, we report several channels through which the anticorrelation between hygiene and food quality could be mitigated, which can be helpful for hygiene inspection design. Third, our empirical approaches address the issue of restricted quality (in our case, hygiene) data availability. Ideally, a study should get access to the raw quality inspection scores (e.g., 0–100 hygiene scores). However, due to data protection regulations and privacy laws, these raw scores are often unavailable. Instead, researchers may only have access to coarse quality grades available to general public (e.g., 0–5 or A–E hygiene grades, as well as 0-5-star discretized product or service quality ratings). Our study reports empirical approaches, such as hurdle and selection models, to overcome such a restricted data environment to have a better understanding of quality-generating processes with various market backgrounds.

The rest of the paper is organized as follows. Section 2 reviews the literature. Section 3 describes the data, reporting the key variables used for our empirical analysis, as well as the industry background. Section 4 reports our empirical analysis and discusses of findings and their interpretations, as well as policy implications. Section 5 concludes. Lastly, in the

Appendix, we report the details of variable constructions, as well as provide the tables of our empirical analysis.

2. Related Literature

This paper is related to and contributes to the literature on food hygiene and safety regulation, and more broadly, to the literature on quality assurance through information disclosure and certification. A related paper is Jin and Leslie (2009) (henceforth, JL), an empirical study of how reputation-related factors affect restaurants' incentives to provide good hygiene, with the focus on the reputation for chain-affiliated restaurants (e.g., KFC, and Burger King). JL estimate a linear regression using restaurant data from Los Angeles between 1995 and 1998. While JL include Zagat guide ratings of restaurants in the list of the control variables of their model, they do not make any conclusive inference from that, thus leaving the connection between food quality and hygiene effort an open question. Unlike JL, under several econometric specifications (including the selection specification), we obtain conclusive results that relate food quality rating and hygiene.[3]

In addition, the food hygiene literature, which emphasizes asymmetric product-quality information between restaurants and customers, has its foundations in the literature on verifiable disclosure and certification (for overview, see Dranove and Jin, 2010). Most of the related literature is concerned with the effect of disclosure rules of hygiene grades on incentives to provide good hygiene. Jin and Leslie (2003) show that the enforcement of a mandatory hygiene rating disclosure policy in Los Angeles in 1998 was effective for enhancing hygiene scores and curbing foodborne illness (see also Jin and Leslie, 2005). This is also corroborated by Simon et al. (2005) who report the reduction of foodborne illness rates in Los Angeles as compared the rest of California, as a consequence of introduction of publicly posted hygiene grade cards. In contrast, Ho et al. (2019), who also investigate the same effect in California, report no statistically significant effect. Wong et al. (2015) show a significant improvement in sanitary conditions in New York City as a result of the introduction of letter grading of hygiene combined with unannounced inspections. In contrast, Ho (2012) documents serious flaws in the implementation of the policy of the mandatory hygiene information disclosure in New York and San Diego: the hygiene certificates are not valuable informative signals

[3]Another related paper is a technical report of Salis et al. (2015). By using regional data, it investigates the effect of the introduction of a uniform country-wide food safety regulatory framework in England, Northern Ireland, and Wales on the restaurants' compliance with food safety regulations, as well as on regionally aggregated food disease statistics. It reports mostly positive results for both compliance and deterring foodborne illness.

but often just a "window dressing". Zhu (2020a) also notes the window dressing and score bunching behavior of hygiene information disclosure in New York City.[4]

Several papers tackle food safety policy design and related research questions. Focusing on the informational contents and responsive behavior of hygiene regulations, Kang et al. (2013) provide an algorithmic approach that extracts information from social media to discover restaurants with poor sanitary conditions accurately. Jin and Lee (2014) study the effectiveness of a new hygiene inspection technology introduced in Florida State. Bederson et al. (2018) study restaurants' incentives to disclose their food hygiene grades under a voluntary disclosure policy. Makofske (2019, 2021) investigates the effect of anticipation of hygiene inspections on restaurants' compliance with regulations in Arizona. Specifically, Makofske (2019) reports the importance of random (un-anticipated) inspection using the inspection records from Los Angeles, and Makofske (2021) complements the former study by documenting the effect of anticipated inspections in Las Vegas. The effect of repeated inspections on hygiene performance is examined by Jin and Lee (2018), Kovács et al. (2020), and Zhu (2020b) using the data from Florida, Los Angeles, and Seattle, respectively. Makofske (2020a) studies the role of repeated violations in Las Vegas, which stipulates that identical violations of food safety in consecutive inspections result in an automatic score-letter downgrade.[5]

Our paper is also related to the growing and actively ongoing literature on quality evaluation information, information provision, transmission, certification, intermediary, and regulation (see, e.g., Belleflamme and Peitz, 2015, Ch. 23). If we interpret the guidebook food quality score as costly but credible signals (see Figueroa and Guadalupi (2023)) to increase a restaurant's demand, as well as reducing search cost, our research shares commonalities with Hong (2022), who investigates the role of the intermediary (i.e. guidebooks and Tripadvisor in the context of this study). Considering the influential role of restaurant guidebooks and Tripadvisor as informational advisors, this research is also related to Cuevas and Bernhardt (2023), who study the role of an influential financial advisor in Chile. In this literature,

[4]The impact of hygiene information disclosure on Yelp.com on hygiene scores in Louisville is investigated by Makofske (2020b), and the impact of the same factor on consumer restaurant demand in the San Francisco is studied by Dai and Luca (2020). Focusing on the multitasking nature of hygiene inspection and evaluation score, De Silva et al. (2023) estimate the multitasking model of hygiene generation, reporting the counterfactual policy analysis, such as partially mandatory information disclosure.

[5]Another strand of literature studies consumer's attitude and habits, which is an important factor for the design of effective designing policy rules. Based on collected survey data in Ontario, Henson et al. (2006) investigate how consumers perceive restaurant safety issues in practice. Lehman et al. (2014) study consumers' attitude towards restaurants' hygiene and authenticity, and provide empirical evidence that authenticity considerations tend to take precedence. Bailey and Garforth (2014) and Adalja and Lichtenberg (2018) document and analyze the burden of using food safety practices and complying with hygiene regulations from the perspective of food farmers.

TABLE 1. The FSA 6-Tier Food Hygiene Rating Scheme.

Description of Hygiene Rating	Score Y_i
Very Good (Tier 5)	0
Good (Tier 4)	1
Generally Satisfactory (Tier 3)	2
Improvement Necessary (Tier 2)	3
Major Improvement Necessary (Tier 1)	4
Urgent Improvement Necessary (Tier 0)	5

much attention is directed to consumer-side choices, such as the effects of ratings and recommendations on sales (Belleflamme and Peitz, 2020), platform intermediation (Baye and Morgan, 2001; Loginova and Mantovani, 2019), and certification (Lizzeri, 1999). Recently, the roles of public information and optimal disclosure, which are linked to food quality and hygiene information in this study, are studied by Lemus and Temnyalov (2023) in the contest design context. Belleflamme and Peitz (2014) and Zapechelnyuk (2019) investigate seller-side decisions to improve product quality, which consumers perfectly or imperfectly observe. Lastly, the producers' incentives to implement quality control in response to different factors, such as organizational changes and consumer pressure, are investigated in, e.g., Anton et al. (2004) and Khanna et al. (2009). Our paper adds to this literature by reporting empirical approaches to study various seller-side incentives and market backgrounds in order to generate policy implications.

3. Data

Our study focuses on the UK high-end restaurant market. We use data from four sources: the Food Standards Agency, two restaurant guidebooks, and Tripadvisor.

3.1. **Food Standards Agency Data.** The hygiene data for each of the restaurants comes from hygiene reports published online by the Food Standards Agency (henceforth, FSA). The FSA reports a 6-tier hygiene rating, which can be looked up by consumers by searching for the restaurant's name.[6,7] With each of 6 tiers, we associate a score $Y_i \in \{0, 1, 2, 3, 4, 5\}$, as shown in Table 1. The decreasing order, i.e., 0 is the best and 5 is the worst, is used for consistency with the Poisson and Negative Binomial model specifications.

From the FSA website we also collected the number of food premises per each hygiene inspector, which allows us to calculate the inspector's workload in each local authority area.

[6]See http://www.food.gov.uk.

[7]Zapechelnyuk (2019) suggests that the disclosure of grades on a fine scale may be socially undesirable; in fact, even the 6-tier rating scheme is inferior to the 2-tier rating (pass/fail) under a wide range of assumptions.

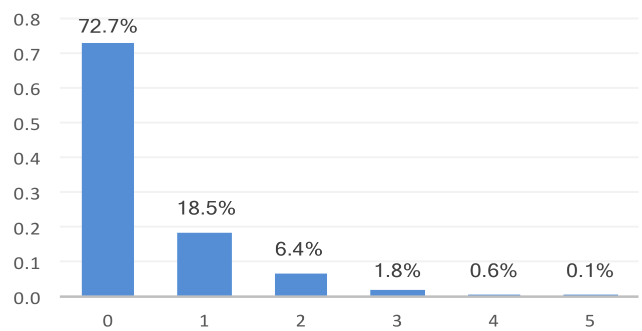


FIGURE 1. Histogram of Hygiene Scores Y_i.

While we focus on the guidebook-listed restaurants, we note that the FSA emphasizes uniformity and equality in hygiene inspections. To the best of our knowledge, there is no evidence that the hygiene inspections of high-end restaurants are conducted differently from those of other restaurants. Also, restaurants can submit appeals to the FSA, which deters any discriminatory inspection conduct.

We note that, since we use the Poisson model, variable Y_i has the domain of nonnegative integers. Yet, the data restricts Y_i to be between 0 and 5. The theoretically possible values of $Y_i > 5$ correspond to hygiene standards even worse than "Urgent Improvement Necessary" ($Y_i = 5$). So, it is reasonable to assume that such restaurants are required to close down and, thus, are unobservable. We argue that this top censoring is not a major concern. As shown in Figure 1, the frequency of $Y_i = k$ in our dataset decreases fast as k goes up. In fact, there is a single observation of $Y_i = 5$. So, if there was a possibility to add $Y_i = 6, 7, \ldots$, the likely effect on the results would be negligible.

3.2. **Restaurant Guidebook Data.** In our empirical analysis, we consider food quality as evaluated by professional reviewers. The professional review scores come from two most popular periodic restaurant guidebooks in the UK: *The Michelin Guide* (henceforth, MG) and *The Good Food Guide* (henceforth, GFG), resulting in individual records of 1,004 restaurants in England, Northern Ireland, and Wales.[8] All of these restaurants are listed in either MG or GFG (or both). Vital to this study, neither MG nor GFG include restaurants' hygiene conditions in their evaluations, as expressed in the description of their evaluation criteria.

MG is a local version of the globally recognized Michelin-brand guidebook series. One of the notable characteristics of this guidebook is its exclusive star-rating system. It lists 'star'-rated restaurants, from one to three stars, according to the criteria described in Table 2.

[8]Scotland has a different hygiene scoring system, which is difficult to compare with that for England, Northern Ireland and Wales. We thus exclude Scotland from our study.

Table 2. Michelin Star Ratings.

Rating	Description
Three Stars	Exceptional cuisine, worth a special journey
Two Stars	Excellent cooking, worth a detour
One Star	High quality cooking, worth a stop

Table 3. Good Food Guide Scores.

Score	Description
10	Perfect dishes showing faultless technique at every service. An extremely rare accolade.
9	Cooking that has reached a pinnacle of achievement. A hugely memorable experience.
8	Highly individual with impressive artistry. There is little room for disappointment here.
7	High level of ambition. Attention to the smallest detail. Accurate and vibrant dishes.
6	Exemplary cooking skills, innovative ideas, impeccable ingredients and an element of excitement.
5	Exact cooking techniques, balance and depth of flavour. A degree of ambition.
4	Dedicated, focused approach to cooking. Good classical skills, high-quality ingredients.
3	Good cooking, showing sound technical skills and using quality ingredients.
2	Decent cooking, good technical skills, interesting combinations and flavours. Occasional inconsistencies.
1	Capable cooking with simple food combinations and clear flavours, but some inconsistencies.

There are under 200 Michelin starred restaurants in the UK in 2019. It is widely recognized among restaurant industry participants that obtaining a Michelin star is a prestigious accomplishment for restaurant owners and chefs.

GFG is another very popular restaurant guidebook in the UK, which rates about 1000 high-end restaurants per year across the UK. For each restaurant, this guidebook reports a score of 10 (maximum) to 1 (minimum), which is awarded by the Good Food Guide editorial office, as well as the typical price of a three-course meal, number of seats, cuisine type, opening days and hours, and other information. Descriptions of GFG scores are provided in Table 3.

For each GFG-listed restaurant, the following individual-level data are collected: name, address, postcode, listed price, food-quality score, number of seats, wine list availability, and number of chefs. For illustration, the locations of the restaurants listed in the guidebooks are mapped in Figure 2. Michelin star ratings are obtained from MG. Out of the total of 166 Michelin-starred restaurants in its 2019 edition, 155 are also listed in the GFG. In addition, for each of the GFG-listed restaurants, we obtained two more restaurant-specific variables from Tripadvisor: the number of consumer reviews, and the service rating.

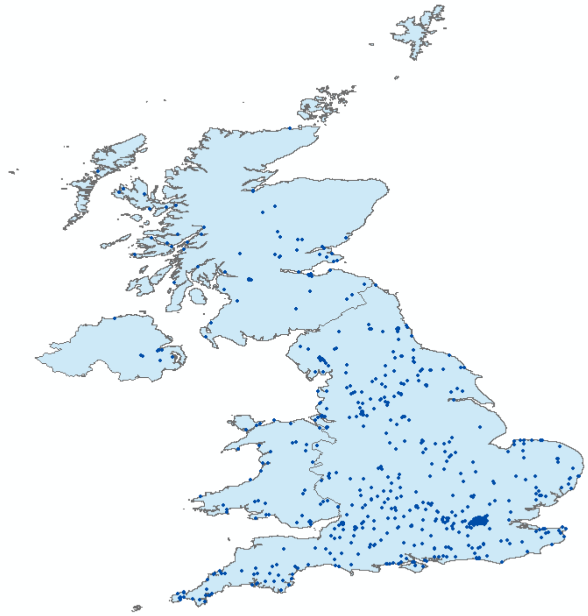
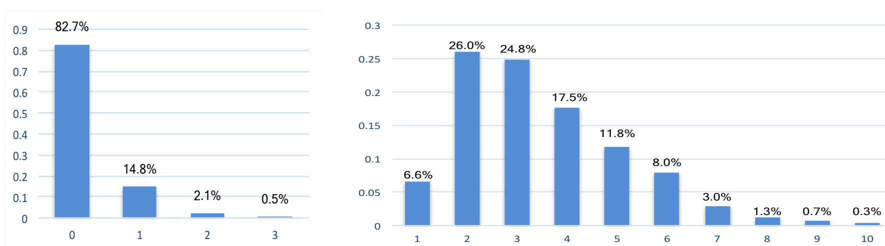


FIGURE 2. Map of restaurants in the dataset



Michelin Stars GFG Scores

FIGURE 3. Histograms of Michelin Stars and GFG Scores.

Histograms of the MG stars and GFG scores in the data are shown in Figure 3. In our analysis, we denote a restaurant i's food quality score from guidebooks by F_i. We consider the combined score[9]

$$F_i = 0.5F_i^{GFG} + 0.5F_i^{MG}, \tag{1}$$

where F_i^{GFG} is the GFG review score in $\{1, ..., 10\}$, and F_i^{MG} is the MG score defined by

$$F_i^{MG} = \frac{10}{3}(\#\text{ of stars}).$$

So, Michelin stars are mapped to the GFG scoring system proportionally: 3 stars corresponds to the GFG score 10, 2 stars to the GFG score 6.7, etc.

[9]As a robustness check, we will also estimate the regression with other scores, in particular, using the GFG score F_i^{GFG} instead of the combined score F_i.

Table 4. Amazon Bestseller Book Ranking (As of December 2018).

Guidebook Name	Amazon.co.uk Bestseller Book Ranking	
	General Category	Restaurants, Bars & Cafés for Travelers
The Michelin Guide: Great Britain and Ireland, 2019	870th	2nd
The Good Food Guide, 2019	1,265th	3rd
Harden's UK Restaurant Survey, 2019	37,102nd	54th
The AA Restaurant Guide, 2019	64,294th	193rd

A possible concern is that the MG and GFG are not the only restaurant guidebooks published for the UK market. There are others, namely Harden's UK Restaurant Survey and the AA Restaurant Guide. These guidebooks, however, are not as popular: the popularity gap between GFG and MG and the other two can be seen from the Amazon bestseller book ranking of these guidebooks from December 2018, soon after the publication of their 2019 issues (Table 4). This justifies our focus on MG and GFG.

It should be mentioned that, in this study, we also collected relevant variables from *Tripadvisor* website in August 2019 for the restaurants that are included in MG and GFG. Such variables from *Tripadvisor* include the number of online reviews and service score evaluations. See the Appendix for details.

3.3. **Sample Size.** In our dataset, we originally had 1,004 restaurants listed in GFG. However, we excluded 83 observations that had no FSA hygiene rating (awaiting inspection, closed, or missing information). We further excluded 69 observations whose GFG score or other variables were missing (e.g., new chef or awaiting review). After these exclusions, the sample size for this subset is 852.

4. Empirical Analysis

In this section, we explore various econometric specifications of count data regressions with relevant descriptions in the order of: Poisson, Negative Binomial (NB), generalized Poisson, truncated Poisson/NB, hurdle Poisson/NB, and Poisson with potential endogenous selection/survival. Specifically, we apply the pseudo-maximum likelihood estimation method for count data, which is robust to model misspecifications. The regressand is the hygiene score Y_i of restaurant i (see Section 3.1). We specify the mean λ_i as

$$\lambda_i = \exp\left(b_F F_i + b_{FC}(F_i \cdot C_i) + b_C C_i + \mathbf{b}_{cha}\mathbf{X}_i^{cha} + \mathbf{b}_{cui}\mathbf{D}_i^{cui} + \mathbf{b}_{reg}\mathbf{D}_i^{reg} + \alpha\epsilon_i\right). \quad (2)$$

In the above specification, the variables are: F_i is a food quality score (see Section 3.2); C_i is the competition dummy; $\mathbf{X}_i^{cha}$ is the vector of restaurant-level characteristic and other control variables, including a constant term; $\mathbf{D}_i^{reg}$ is the vector of regional dummies; $\mathbf{D}_i^{cui}$ is the vector of cuisine dummies (see the Appendix for the description of these variables). In addition, ϵ_i is the heterogeneity term that summarizes any factors that are not explained by the observables, and α is a parameter. We assume that the exponent of the heterogeneity term, $\exp(\alpha\epsilon_i)$, is gamma-distributed.

Below, for these empirical analyses, we primarily use the food-quality variable (F_i), consisting of the weighted combination of the Micheline and Good Food Guide food scores, as described in Section 3.2. However, as a robustness check, we also use the Good Food Guide food-quality score without combining it with the Micheline food-quality score. In the remaining part of this section, we report the results of our empirical analysis in four parts: (1) basic analysis, (2) two-part hurdle model with potential market segmentations, (3) potential location and survival selection, (4) comparison of econometric specifications with standard information criteria. In addition, we follow the pseudo-/quasi maximum likelihood (PPLM) approach (see: Cameron and Trivedi, 2005; Cameron and Trivedi, 2013), and all reported standard errors are calculated by the robust variance-covariance matrix or bootstrapping.

4.1. **Basic Specification.** We start with the basic count-data regression specifications of Poisson, Negative Binomial (NB), and generalized Poisson, which are reported in Table 5. First, in the left column of Table 5, the Poisson specification count-data regression result is reported. The estimate of the food quality coefficient (0.1192) is positive, indicating anticorrelation.[10] However, this coefficient estimate is not significant at the traditional statistical significance level, although it comes with the p-value of 0.174. Second, the result of Negative Binomial regression, which allows the possibility of overdispersion, is reported in the middle column of Table 5, indicating a similar positive estimate of 0.1198 (i.e., anticorrelation) but again without significance. Third, the regression outcome of the generalized Poisson regression, which allows both underdispersion and overdispersion, is reported in the right column of Table 5, suggesting a positive but insignificant estimate of the food quality coefficient.

However, our dataset consists of the hygiene quality count-dependent variable with many zeroes, which indicates the possibility of a segmented market: the guidebook-listed restaurants could be segmented in two groups, one for focusing on the hygiene aware consumers (who pay relatively high attention to food hygiene), and the other for food-quality aware consumers (who may pay relatively less attention to food hygiene).

[10] A positive coefficient means anticorrelation, because higher Y_i means lower hygiene (see Section 3.1).

4.2. **Truncated Regressions.** To have an initial check for a market segmentation possibility described above, we run truncated Poisson and Negative Binomial regressions among the restaurants with positive hygiene score Y_i (i.e., restaurants with positive count-dependent variables with the truncation at zero). The results are reported in Table 6. The left column of Table 6 reports the truncated Poisson regression outcome, while the right column reports that of the truncated Negative Binomial regression. Both regression results indicate a statistically significant positive coefficient of food quality, 0.5041 and 0.4958, respectively (i.e., anticorrelation) with the 1 percent significance level. These truncated regression results could imply the segmented market, warranting further investigation, which is explored next.

4.3. **Hurdle Model.** Guided by the truncated regression results, we consider the exponential mean hurdle model. Here, we examine the difference in economic behavior between the restaurants with the highest and non-highest hygiene scores. The first part of the hurdle model is a logit regression with the binary dependent variable that indicates whether a restaurant's hygiene is the highest (i.e., $Y_i = 0$) or not (i.e., $Y_i > 0$). The second part is described either by a truncated Poisson or Negative Binomial regression, conditional on the non-best hygiene quality ($Y_i > 0$) in the first part.

Table 7 reports the results of the hurdle models. The left column of Table 7 describes the hurdle-Poisson specification, in which the binary ($Y_i = 0$ or $Y_i > 0$) first-part hurdle equation is specified as a logit regression, while the second part is specified as a conditional truncated Poisson distribution. In addition, the second-from-the-left column of Table 7 summarizes the hurdle-NB specification, which employs a truncated Negative Binomial distribution in the second part. Both hurdle-Poisson and hurdle-NB model estimation results in Table 7 indicate that the food quality coefficients in the non-zero truncated dependent variable parts are significantly positive, 0.4974 and 0.4892, respectively.

In the second-from-the-right column (hurdle-Poisson) and the right column (hurdle-NB) of Table 7, we check the robustness of our hurdle model findings by using the GFG food-quality score, instead of the weighted combination of the Michelin and GFG food-quality scores. In the non-zero truncated dependent variable part, the coefficients of the GFG score coincide on both columns and equal to 0.3818. They are statistically significant, indicating anticorrelation.[11] Overall, the hurdle models in Table 7 exhibit statistically significant anticorrelation between hygiene and food quality only among the non-best hygiene (i.e., $Y_i > 0$) restaurants, but not among the best hygiene restaurants (i.e., $Y_i = 0$). We will discuss economic interpretations of this finding in Section 4.7.

[11]Focusing on the non-competitive area restaurants, and conditional on observing $Y_i > 0$, the partial-derivative-based marginal associations between hygiene and food quality (with 95% confidence intervals) of the hurdle model second parts, as listed in the order presented in Table 7, are 0.3859 (0.1070, 0.6647), 0.3653 (0.0892, 0.6414), 0.2970 (0.1103, 0.4836), and 0.2944 (0.1095, 0.4792), respectively.

4.4. **Competitive Location and Survival Selection.** In this section, we consider a restaurant's location and survival selection possibilities. Potentially endogenous restaurant location and survival, notably in the competitive city-center areas, arises as follows: A restaurant which is located and/or survives in a competitive area has specific characteristics, such as demand strong enough to survive fierce competition among high-end restaurants, and which may not be captured by observed variables.[12] If such location-selection and survival-related characteristics are omitted, reported estimates suffer from biases. To investigate such selection possibilities, we report the estimates, derived by using the two-step correction method, for a count dependent variable (Terza, 1998), which could be considered as a variant of the well-known two-step sample selection correction methods of Heckman (1976, 1978) for count data.

In the first step, the location and survival selection process are described by a selection equation

$$C_i = \boldsymbol{\delta}\mathbf{W}_i + u_i, \tag{3}$$

where C_i is a binary competitive area indicator variable (which is equal to one if a restaurant is located in a competitive area), $\mathbf{W_i}$ is the set of selection-related variables, including pre-determined reputation variables, and $\boldsymbol{\delta}$ is a vector of coefficients. Importantly, u_i is a residual term, which absorbs the selection-related omitted variables, as well as other idiosyncrasies. The second step characterizes an exponential mean equation

$$\lambda_i = \exp(\mathbf{b}\mathbf{X}_i + \rho u_i + \alpha\epsilon_i), \tag{4}$$

where λ_i is the notation for a mean, $\mathbf{X}_i$ is a vector of observables as listed in equation (2) (including C_i), $\boldsymbol{\delta}$ is a vector of coefficient parameters, ρ is the coefficient of u_i (1st stage residual term). Moreover, the term of ϵ_i is a random heterogeneity term, which captures factors not explained by $\mathbf{X_i}$ and u_i. Note that α is a gamma-distribution parameter, and we assume that the random variable $\exp(\alpha\varepsilon_i)$ has gamma distribution.

Practically, in the first step of Terza's method with equation (3), we run a linear regression with the binary dependent variable of a competitive area indicator.[13] In this first step, we use the number of Tripadvisor reviews and accumulated listing years in the Michelin guide and Good Food Guide explanatory variables (i.e., accumulated reputation), which is considered

[12]Focusing on restaurants in a non-competitive area, an alternative interpretation is that the restaurants which are located and survive in non-competitive (typically non-city center) areas have specific characteristics to compensate for being in a less populated neighborhood, which might not be observable.

[13]In the original study (Terza, 1998), the first step is further specified by the probit model, which is more efficient under the correct specification of the normally distributed error terms. In this study, we conservatively refrain from assuming this normality by using the linear regression in the first step, and the standard errors in the second step are calculated by bootstrap with the 400 repetitions of the same re-sample size (as original data) draws, as described by Cameron and Trivedi (2005, 2013).

to influence the location and survival in a competitive area. In addition, in the vector of $\mathbf{W_i}$ in equation (3), we also include the restaurant's other contemporary quality variables, food-quality and service-quality variables, which could be related to its current competitive location choice.[14] We then calculate a residual term in this linear regression. This residual term (in the first step) captures competitive-location- and survival-selection-related omitted variables, which are absent in baseline Poisson and NB regression specifications. Next, in the second step of Terza's correction method with equation (4), we run the count data dependent variable Poisson or NB regression, which includes the first-step residual term, described above. The first-step residual term is included in the second step to mitigate the potential bias in the estimates.

Table 8 reports the second-step results of two-step selection-corrected regressions, while Table 9 describes those of the first step.[15] Parentheses in Table 8 report bootstrapped standard errors, resulting from the two-step nature of Terza's method. Also, in Table 8, we use the Michelin-GFG weighted food quality variable for the left and second-from-the-left column regressions. In the left column, we report selection-corrected Poisson regression with the food quality coefficient estimate of 0.2241 (indicating anticorrelation) with statistical significance. The second-from-the-left column reports the selection-corrected NB regression outcomes with the positive and significant food quality coefficient estimate of 0.2218.

In the second-from-the-right and the right columns, we report the regressions based on the GFG food-quality score variable. These GFG food-quality-based regression outcomes report similar findings, the significant coefficient estimates of 0.2081 and 0.2178, respectively. These selection-correction regressions persistently suggest that, even after adjusting for the location and survival selection possibilities, anticorrelation between hygiene and food quality still remains.[16]

Lastly, we report that the estimated coefficient of u_i, reported in Table 8, are all positive and significant, statistically indicating the existence of competitive-location- and selection-related

[14]For example, a restaurant may choose to be located in a competitive area, while compromising its service quality due to a smaller space (for each table), involving a relatively high rent in a city-center area.

[15]The signs of the estimated coefficients in Table 9 are in line with our general perceptions (e.g., competitive area restaurants are conditionally associated with low service ratings). The exception is the sign of reputation variable (i.e., the variable of Michelin & GFG Past Listed Years). The potential reason for this negative conditional association could be a relatively high turnover rate, as the competitive area restaurants have a relatively high frequency of entries and exits.

[16]For each of these selection specification count-data regressions, focusing on non-competitive area restaurants, we can calculate a marginal association, which is the partial derivative of the exponential mean equation (4) with respect to the food-quality variable. The marginal associations between hygiene and food quality (with 95% confidence intervals), as listed in the order presented in Table 8, are 0.0847 (0.0064, 0.16305), 0.0845 (0.0049, 0.1641), 0.0786 (0.0226, 0.1347), and 0.0830 (0.0241, 0.1418), respectively.

omitted variable in the baseline specification of equation (2), providing a strong support for the usage of Terza's two-step bias-correction method.

4.5. **Information Criteria.** In this section, we report the goodness of fits, measured by the traditional information criteria, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC).[17] Regarding the information criteria, as each of the econometric specifications has a varying number of estimating parameters, we use the classical AIC, BIC, and HQIC to access the goodness of fit. Note that the BIC and HQIC have relatively heavier penalties on the number of parameters than the AIC.

The log pseudolikelihoods, number of parameters, AIC, BIC, and HQIC are reported in the second through sixth columns of Table 10. The evaluation of the goodness of fit involves the subtle tradeoff between the improvement of log pseudo-likelihood and increased number of parameters for insightful economic interpretations. In this vein, the hurdle-Poisson and hurdle-NB models, although having nearly doubled numbers of parameters, have AIC numbers 1370.3 and 1372.1, similar to the basic Poisson specification (1370.0), which provide a justification for hurdle specifications, motivated by the market segmentation interpretation. Moreover, regarding the selection-Poisson model (Terza, 1998), it has the improved log pseudo-likelihood numbers (1364.2 and 1349.6), adding further support for the competitive location and survival selection possibilities. Regarding the BIC and HQIC statistics, which further penalize the usage of a large number of parameters, they provide further support for the selection models. Overall, the advanced econometric specifications reported in Tables 7 and 8, which examine the various possibilities and provide insigutful economic interpretations of restaurants' behavior, have similar AIC statistics or improved information criteria.

4.6. **Policy Implications.** There are several takeaway policy implications from our empirical analysis. First and most importantly, there is statistically significant conditional anticorrelation between hygiene and food quality, notably among non-best hygiene quality restaurants and after adjusting location and survival selections. Second, important for policymaking, we obtain largely different regression results with and without considering market competition and selections. Thus, upon the consideration of hygiene policy, it could be suggested for researchers and policymakers to consider and include market background information, such as competitive location and survival selections. Third, we also report several observable restaurant characteristics, which are conditionally associated with hygiene, such as capacity, multiple chefs, and service rating. These observables could be used for inspection design, such as to determine which restaurants need to be prioritized for inspections.

[17]See Akaike (1998), Schwarz (1978), and Hannan and Quinn (1979) for details.

4.7. **Discussion.** Though our results show a generally negative association between food quality and hygiene, there are nuances across the econometric approaches that we employ. These provide some tentative explanation as to the causes of the observed association.

We conceptually distinguish two possibilities: one that restaurant's choice of food quality and hygiene are strategic choices, shaped by the market environment. This could happen, for example, if there are two types of consumers: those who are primarily taste conscious, and those who are primarily hygiene conscious. Restaurants would segment the market, and thus focus on improving one of the two dimensions while being somewhat indifferent with respect to the other. This would straightforwardly yield a negative correlation between taste quality and hygiene. We refer to this as the *strategic hypothesis*.

The second possibility is that the association could be due to survival selection under competition. Restaurants may have a fixed taste quality and a fixed hygiene level, i.e., they are endowed with them but do not choose them. Such fixed restaurant attributes could be chef skill and attitudes or routines about food handling. If demand is a function of both hygiene and taste quality, then restaurants that are low in both will go out of business. The ones that will survive will be either good in both, or good in one of the two. Therefore, there will be a negative correlation between taste quality and hygiene. Moreover, the selection effect would be stronger in more competitive markets. However, more competitive markets would not have a causal effect on taste quality and hygiene, it only has an effect on the distribution of these attributes in the surviving restaurants. We refer to this as the *selection hypothesis*.

Our results in Tables 6 and 7 (trucated and hurdle models) show that taste quality does not affect the likelihood that restaurants attain the highest level of hygiene, but it is associated with lower hygiene for restaurants that are not the most hygienic. This is inconsistent with the strategic hypothesis, which posits that restaurants with top food quality would not need to specialize in being very hygienic, and thus, top taste quality should make a restaurant less likely to in the most hygienic category. However, the results are consistent with the selection hypothesis. Taste quality should not have much of an impact on restaurant survival among the most hygienic restaurants, but restaurants that are not very hygienic would need high taste quality to survive.

Moreover, when we do not control for the endogeneity of competition, we find that competition is negatively associated with being of the top hygienic quality and also negatively associated with hygienic quality below the top hygienic category (the second finding is less pronounced for higher food quality restaurants). When we do account for the endogeneity of competition in Tables 8 and 9 (Terza's selection model), we find that these effects disappear: though higher competition reduces the anti-correlation and also increases hygienic quality, neither effect is statistically significant. These findings are again not consistent with the strategic hypothesis, which suggests that competition should have a causal effect of restaurant specialization (and strengthen it). The findings, though, could be consistent with the selection hypothesis: stronger competition may weaken the anti-correlation, as restaurants of poor hygiene may not survive in a competitive environment even if their food quality is high. Overall, our empirical investigation indicates that the selection hypothesis has more support.

5. Conclusion

In this study, we empirically analyze the relationship between hygiene and food quality, using high-end restaurant data from the UK. We show that, under several econometric specifications, there is a statistically significant negative conditional correlation between the food hygiene and the food quality rating, and the relationship holds for two different measurements of food quality, as evaluated by professional reviewers. We examine two hypotheses. The first is the *strategic hypothesis*, in which the segmented market with hygiene- and taste-aware consumer groups makes a restaurant focus on either one of them. Such a focusing choice straightforwardly results in a negative correlation between hygiene and food quality. The second is the *selection hypothesis*, in which restaurants are endowed with fixed quality characteristics rather than choosing them. Then, competition forces only specific combinations of quality and hygiene to survive. This process results in survival selection, and a negative correlation between hygiene and food quality emerges through the selection. Our analysis indicates the latter hypothesis has stronger empirical support.

The reported empirical result is often perceived as counterintuitive, as diners would expect professional reviews to provide reliable indication not only about the food quality, but also about the cleanliness of the kitchen and adherence of the personnel to food hygiene rules and procedures. In addition, appealing to hygiene regulators and policymakers, our study indicates factors that can be influenced in order to mitigate the negative association with hygiene standards. The relationship between hygiene scores and other variables (such as service quality rating and multiple chefs) can be informative for hygiene inspectors to decide which restaurants should be prioritized in their inspection schedule design.

References

Adalja, A., and E. Lichtenberg (2018): "Produce growers' cost of complying with the Food Safety Modernization Act," *Food Policy*, 74, 23–38.

Akaike, H. (1998): "Information theory and an extension of the maximum likelihood principle," in *Selected papers of hirotugu akaike*, pp. 199–213. Springer.

Anton, W. R. Q., G. Deltas, and M. Khanna (2004): "Incentives for environmental self-regulation and implications for environmental performance," *Journal of environmental economics and management*, 48, 632–654.

Bailey, A. P., and C. Garforth (2014): "An industry viewpoint on the role of farm assurance in delivering food safety to the consumer: The case of the dairy sector of England and Wales," *Food Policy*, 45, 14–24.

Baye, M. R., and J. Morgan (2001): "Information gatekeepers on the internet and the competitiveness of homogeneous product markets," *American Economic Review*, 91, 454–474.

Bederson, B. B., G. Z. Jin, P. Leslie, A. J. Quinn, and B. Zou (2018): "Incomplete Disclosure: Evidence of Signaling and Countersignaling," *American Economic Journal: Microeconomics*, 10, 41–66.

Belleflamme, P., and M. Peitz (2014): "Asymmetric information and overinvestment in quality," *European Economic Review*, 66, 127–143.

——— (2015): *Industrial organization: Markets and strategies*. Cambridge University Press, 2nd edn.

——— (2020): "Ratings, reviews, and recommendations," in *Handbook of Cultural Economics*. Edward Elgar Publishing, 3rd edn.

Cameron, A. C., and P. K. Trivedi (2005): *Microeconometrics: Methods and Applications*. Cambridge university press.

——— (2013): *Regression Analysis of Count Data*. Cambridge University Press, 2nd edn.

Cuevas, C., and D. Bernhardt (2023): "When financial advice rocks the market," *Emerging Markets Review*, 56, 101051.

Dai, W., and M. Luca (2020): "Digitizing disclosure: The case of restaurant hygiene scores," *American Economic Journal: Microeconomics*, 12, 41–59.

De Silva, D. G., D. Rietzke, A. R. Schiller, and H. Yoshimoto (2023): "Multitasking, complementarities, and information disclosure: Theory and evidence from the UK," *Available at SSRN 4375918*.

Dranove, D., and G. Z. Jin (2010): "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 48, 935–63.

Figueroa, N., and C. Guadalupi (2023): "Signaling through tests," *The Quarterly Review of Economics and Finance*, 92, 25–34.

Hannan, E. J., and B. G. Quinn (1979): "The determination of the order of an autoregression," *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2), 190–195.

Heckman, J. J. (1976): "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models," in

Annals of economic and social measurement, volume 5, number 4, pp. 475–492. NBER.

——— (1978): "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica: Journal of the Econometric Society*, pp. 931–959.

Henson, S., S. Majowicz, O. Masakure, P. Sockett, A. Jones, R. Hart, D. Carr, and L. Knowles (2006): "Consumer Assessment of the Safety of Restaurants: The Role of Inspection Notices and Other Information Cues," *Journal of Food Safety*, 26, 275–301.

Ho, D. E. (2012): "Fudging the Nudge: Information Disclosure and Restaurant Grading," *Yale Law Journal*, 122, 574–689.

Ho, D. E., Z. C. Ashwood, and C. Handan-Nader (2019): "New evidence on information disclosure through restaurant hygiene grading," *American Economic Journal: Economic Policy*, 11, 404–28.

Hong, S.-H. (2022): "Real estate agents' influence on housing search," *Journal of Applied Econometrics*, 37(3), 563–582.

Jin, G. Z., and J. Lee (2014): "Inspection Technology, Detection, and Compliance: Evidence from Florida Restaurant Inspections," *RAND Journal of Economics*, 45, 885–917.

——— (2018): "A Tale of Repetition: Lessons from Florida Restaurant Inspections," *The Journal of Law and Economics*, 61, 159–188.

Jin, G. Z., and P. Leslie (2003): "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, 118, 409–451.

——— (2005): "The Case in Support of Restaurant Hygiene Grade Cards," *Choices*, 20, 97–102.

——— (2009): "Reputational Incentives for Restaurant Hygiene," *American Economic Journal: Microeconomics*, 1, 237–267.

Kang, J. S., P. Kuznetsova, M. Luca, and Y. 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*, pp. 1443–1448.

Khanna, M., G. Deltas, and D. R. Harrington (2009): "Adoption of pollution prevention techniques: the role of management systems and regulatory pressures," *Environmental and Resource Economics*, 44, 85–106.

Kovács, B., D. W. Lehman, and G. R. Carroll (2020): "Grade inflation in restaurant hygiene inspections: Repeated interactions between inspectors and restaurateurs," *Food Policy*, 97, 101960.

Lehman, D. W., B. Kovács, and G. R. Carroll (2014): "Conflicting Social Codes and Organizations: Hygiene and Authenticity in Consumer Evaluations of Restaurants," *Management Science*, 60, 2602–2617.

Lemus, J., and E. Temnyalov (2023): "Diversification and information in contests," *Economic Theory*, pp. 1–32.

Lizzeri, A. (1999): "Information revelation and certification intermediaries," *RAND Journal of Economics*, 30, 214–231.

Loginova, O., and A. Mantovani (2019): "Price competition in the presence of a web aggregator," *Journal of Economics*, 126, 43–73.

Makofske, M. P. (2019): "Inspection regimes and regulatory compliance: How important is the element of surprise?," *Economics Letters*, 177, 30–34.

——— (2020a): "Disclosure policies in inspection programs: The role of specific deterrence," *Economics Letters*, 196, 109533.

——— (2020b): "The effect of information salience on product quality: Louisville restaurant hygiene and Yelp. com," *Journal of Industrial Economics*, 68, 52–92.

——— (2021): "Spoiled food and spoiled surprises: Inspection anticipation and regulatory compliance," *Journal of Economic Behavior & Organization*, 190, 348–365.

Salis, S., N. Jabin, and S. Morris (2015): "Evaluation of the impact of the Food Hygiene Rating Scheme and the Food Hygiene Information Scheme on food hygiene standards and food-borne illnesses," Food Standards Agency, available at https://e-space.mmu.ac.uk/618918/1/fhrs-fhis-eval2011-14foodborne.pdf.

Schwarz, G. (1978): "Estimating the dimension of a model," *The annals of statistics*, pp. 461–464.

Simon, P. A., P. Leslie, G. Run, G. Z. Jin, R. Reporter, A. Aguirre, and J. E. Fielding (2005): "Impact of Restaurant Hygiene Grade Cards on Foodborne-Disease Hospitalizations in Los Angeles County," *Journal of Environmental Health*, 67, 32–36.

Terza, J. V. (1998): "Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects," *Journal of econometrics*, 84(1), 129–154.

Wong, M. R., W. McKelvey, K. Ito, C. Schiff, J. B. Jacobson, and D. 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, e81–e87.

Zapechelnyuk, A. (2019): "Optimal Quality Certification," *American Economic Review: Insights*, 2, 161–176.

Zhu, J. (2020a): "Earning the Grade or Just Window Dressing? The Effects of Information Disclosure on Restaurant Hygiene Quality," mimeo.

——— (2020b): "Inspection in Disguise: A Study of Repeated Restaurant Inspections," mimeo.

Appendix A. Other Variables

This Appendix section illustrates the details of the right-hand-side variables, which are not explained in the Section 3. As a short summary, we collected the set of variables from two published guidebooks (Michelin and Good Food Guide), Food Standard Agency, and Tripadvisor websites of each (guidebook-lists) restaurant, which capture the various aspects of restaurant operation. Below, we illustrate the variable constructions, in the order of price, competition (or metropolitan area proxy), workload of hygiene inspector, and other variables.

I. **Competition.** To capture the presence of local competition, we include a competition dummy variable C_i that indicates if there are three or more rival restaurants within a 1/4-mile radius, which splits our sample into two approximately equal-sized subgroups. As the locations of guidebook-listed restaurants often tend to concentrate in urban areas, such as city centers, this variable could also indicate whether a restaurant is in a competitive metropolitan area.

To establish the presence of neighbor rivals, we identify each restaurant's geographic location by its postcode (see Figure 2 for the map of restaurants). In the UK, there are about 1.8 million postcodes, and each postcode consists of on average 15 properties. We use a centroid at each postcode as a proxy for each restaurant's latitude and longitude. Then, using ArcGIS, we calculate the number of guidebook-listed restaurants within 1/4-mile. While it would be ideal to take into account all existing restaurants, our competition proxy is sensible for two reasons. First, it is reasonable to assume that the strongest competition effect comes from similar high-end restaurants listed in the guidebooks, rather than medium and lower-end restaurants. Second, restaurants tend to be clustered, e.g., multiple restaurants in town centers and along high streets, and many such clusters contain several guidebook-listed restaurants.

The presence of competition is a demand-shifting factor that has an ambiguous economic association with hygiene standards. On the one hand, competition increases the marginal return of hygiene standards, thus stimulating the incentive to provide better hygiene. On the other hand, the revenues are competed away, so the cost substitution effect between hygiene and other inputs intensifies. Thus, our particular interest is in the coefficient of the interaction term between the food quality variable and the competition dummy. This coefficient indicates how the relationship between food quality and hygiene changes when competition is present, as compared to the baseline with fewer or no rivals in sight.

II. **Workload of Hygiene Inspectors.** Our model takes into account the workload of hygiene inspectors. For that purpose, we include the variable W_i, a logarithm of the number of food dealing premises per one hygiene inspector in the local authority district where restaurant i is located. This allows us to control for the effect of the variation in the excess or shortage of hygiene inspectors across local authority districts. Similarly to the hygiene rating, the data on the workload of hygiene inspectors comes from FSA.[18]

The workload of hygiene inspectors is a cost-shifting factor that, intuitively, has a negative economic association with hygiene standards. More overloaded hygiene inspectors are likely to backlog inspections and can be less thorough, so the restaurants can have more incentive to invest into food quality rather than hygiene.

III. **Price.** A major restaurant-specific characteristic is its price level. We include into our regression a price listed in guidebooks, P_i. This listed price is fixed in the short run: while the restaurant is free to adjust its prices any time, the guidebook listed price is only updated once a year, when an annual guidebook publication is out. Exploiting this unique nature of the printed guidebooks that the information in their issues remains constant for a substantial period of time, we use the guidebook listed price, as a control variable as it captures consumers' expectations and taste heterogeneity that could be otherwise difficult to observe.

For the regressions where the food quality score is the combined MG and GFG score, we define P_i as the price listed in GFG. This is the GFG evaluation of the price of a three-course meal at restaurant i.

IV. **Other Variables.** We use a number of other restaurant-level control variables obtained from GFG and Tripadvisor. Variables from GFG include the listed price, the seat capacity, the presence of multiple chefs, and the availability of a wine list. The variables from Tripadvisor include cuisine categories, the number of consumer reviews, and the Tripadvisor service rating. We use these variables to capture restaurant-level demand and cost heterogeneity. We will now provide detailed descriptions of these variables.

(a) The capacity variable is the logarithm of the GFG-published number of seats in a restaurant. The capacity is indirectly related to the restaurant's variable cost through the economy of scale.[19]

[18] https://webarchive.nationalarchives.gov.uk/20171207164658/https://www.food.gov.uk/ enforcement/monitoring/laems/mondatabyyear.

[19] A restaurant with a large capacity is frequently affiliated with a larger organization such as a hotel chain or an airport, so it can benefit from economy of scale regarding quality control and personnel training.

(b) The GFG's wine list availability dummy variable is added to capture the type of customers' dining experience and overall quality of the establishment.

(c) The multiple chef dummy variable MC_i from GFG specifies whether there are multiple chefs ($MC_i = 1$) or a single chef ($MC_i = 0$). It is included as an indication of both scale and the availability of extra hands to take care of multiple tasks, such as hygiene routines.

(d) We control for cuisine types as listed in Tripadvisor. For this purpose, we add 12 indicator variables D_i^{cui} for the most popular cuisine categories in Tripadvisor: $cui \in$ {Asian, British, Chinese, European, French, Indian, Italian, Japanese, Seafood, Steak, Bar/Pub, Cafe}, and all other cuisine categories are the baseline. Note that the list of categories combines regional cuisines (e.g., *British*) and types of establishments (e.g., *Cafe*). These are mutually exclusive, so each restaurant has only one cuisine category on Tripadvisor.

(e) We also include regional dummy variables to capture potential FSA administrative differences. The FSA divides the UK into 12 administrative regions: East Counties (EC), East Midlands (EM), London, North East (NE), North West (NW), South East (SE), South West (SW), West Midlands (WM), Yorkshire and Humberside (YH), Northern Ireland, Scotland, and Wales. We exclude Scotland from our study, because it has a different, 2-tier hygiene scoring system, as noted in Footnote 8. We merge Northern Ireland and Wales into a single region, abbreviated as *NIW*, because of very small samle sizes in each of these regions and because of their similar hygiene policies. We set London as the baseline for the regional dummies. We thus have 9 regional dummies D_i^{reg} with $reg \in \{EC, EM, NE, NW, SE, SW, WM, YH, NIW\}$.

(f) We include two additional variables from Tripadvisor: the logarithm of the number of consumer reviews and the service quality score. These factors can play a significant role in the tradeoff between food quality and hygiene. The number of consumer reviews is considered to be a reputation variable, as well as a proxy for whether the consumer base is recurrent (as in local restaurants) or constantly renewed (as in restaurants near tourist attractions). Similarly to Tripadvisor's bubble rating, the service quality score displays the aggregate consumer review rating of the service quality. The service quality is, on the one hand, a cost-shifting factor: a higher service score could potentially mean that there is a smaller amount of resources spend on hygiene. On the other hand, service and hygiene quality can be complementry, as, for example, both depend on staff training and qualifications. Thus, a priori, the relationship between service quality and hygiene is ambiguous.

(g) We use the MG and GFG time-stamp variables to parsimoniously include such potential dynamics, which capture the reputations, as well as food quality transitions from the preceding year. There are two types of time-stamp variables: past listed years and food-quality up-down transition variables, explained as follows.

First, as a pre-determined variable, we can use the accumulated years of the restaurant's Michelin-stared and GFG-listed status, which could proxy for its historical reputation. Specifically, we use the years in which a restaurant obtained Michelin-star status (1-star, 2-stars, and 3-stars), as well as the GFG listed status, since 2008. For example, during 2008-2018 (for 11 years), if a restaurant is awarded MG stars for 7 years/editions, regardless of the continuous years or discontinuous years, we construct the accumulated MG year variable of 7 for this restaurant. If this restaurant is also listed in the GFG for every year during 2008-2018, we construct the accumulated GFG year variable of 11. We then construct the MG-and-GFG past-listed-year variable by taking the equal weight of the accumulated MG and GFG years. For the previously exemplified restaurant, this variable is $9 = \frac{1}{2}(7+11)$. This variable aims to capture the long-term (and pre-determined) reputation in a minimalistic manner.

Second, as a contemporaneous variable, given the construction of our food quality variable (i.e., F_i^G in equation (1)), and by comparing the food quality scores between 2018 (previous year) and 2019 (relevant year), we can construct three mutually exclusive categories for capturing food quality transition: (1) Remain, which works as a baseline, (2) Up, and (3) Down. The first category, (1) Remain, describes the transition state that a relevant restaurant food-quality score remains the same since the previous year. The second category, (2) Up, applies to the situation that a relevant restaurant food-quality score experiences improvement from the last year. This category also includes brand-new restaurants. Lastly, the third category, (3) Down, indicates that a relevant restaurant food-quality score has deteriorated from the previous year. The advantage of these simple categorizations is that, beyond the present level of food quality, we can parsimoniously capture the conditional correlation related to the dynamics of the food quality score, which is testable. For example, conditional on other observable variables, if a restaurant gives up its ambition by moving down the ladder of food quality evaluation score in the present year (e.g., losing MG star[s] or a lower GFG score), it may be associated with an improved hygiene condition.

Appendix B. Tables

Table 5. Basic Poisson and Negative Binomial

	(1)		(2)		(3)	
Dependent Variable: Y_i	Poisson		NB		Generalized Poisson	
Food Quality (MG&GFG)	0.1192	(0.0876)	0.1198	(0.0905)	0.0762	(0.0893)
Food Quality × Competition	-0.1545	(0.1075)	-0.1629	(0.1036)	-0.1552	(0.1095)
Competition	0.6493***	(0.2495)	0.6876***	(0.2452)	0.6001**	(0.2527)
Wine List	0.0687	(0.1949)	0.0814	(0.1963)	0.1012	(0.1972)
Capacity	-0.3240**	(0.1414)	-0.3473**	(0.1423)	-0.3599***	(0.1367)
Multiple Chefs	-0.1900	(0.1676)	-0.1686	(0.1664)	-0.1356	(0.1654)
Listed Price	-0.0128**	(0.0064)	-0.0133**	(0.0062)	-0.0114*	(0.0062)
Tripadvisor Online Service Rating	-0.5652***	(0.2018)	-0.6248***	(0.2084)	-0.5918***	(0.2003)
Food Quality Up or New (Binary)	-0.1050	(0.1833)	-0.1202	(0.1827)	-0.0523	(0.1846)
Food Quality Down (Binary)	-0.3674*	(0.2225)	-0.3847*	(0.2134)	-0.3255	(0.2198)
Inspector Workload (in Natural Log)	0.2984*	(0.1791)	0.2722	(0.1960)	0.3372**	(0.1713)
Cuisine Categories	Yes		Yes		Yes	
Regions	Yes		Yes		Yes	
Log Pseudolikelihood	-652.0225		-642.8283		-640.5505	
α (p-value)			0.6296 (0.0000)			
δ (p-value)					0.1276 (0.0000)	
N	852		852		852	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The cuisine categories are: *Asian*, *Bar & Pub*, *British*, *Cafe*, *Chinese*, *European*, *French*, *Indian*, *Italian*, *Japanese*, *Seafood*, *Steak*. The administrative regions include: *Northern Ireland and Wales*, *East Counties*, *East Midlands*, *North East*, *North West*, *South East*, *South West*, *West Midlands*, *Yorkshire and Humberside*. We use the food quality evaluation variable based on Michelin and Good Food Guide (GFG) scores (weighted half for the Michelin score and half for the GFG score; see the data description section for details). α is a gamma-distribution parameter, while δ is a generalized Poisson distribution parameter.

Table 6. Truncated Poisson and Negative Binomial

	(1)		(2)	
Dependent Variable: $Y_i > 0$	Poisson		NB	
Food Quality (MG&GFG)	0.5041***	(0.1823)	0.4958***	(0.1869)
Food Quality × Competition	-0.1568	(0.1908)	-0.1550	(0.1964)
Competition	0.7827**	(0.3718)	0.7738**	(0.3826)
Wine List	-0.3783	(0.3804)	-0.3653	(0.3900)
Capacity	0.0495	(0.2412)	0.0520	(0.2502)
Multiple Chefs	-0.7817**	(0.3640)	-0.7665**	(0.3717)
Listed Price	-0.0161	(0.0107)	-0.0159	(0.0111)
Tripadvisor Online Service Rating	-0.1790	(0.2867)	-0.1811	(0.2987)
Food Quality Up or New (Binary)	-0.3101	(0.2619)	-0.3060	(0.2717)
Food Quality Down (Binary)	-0.6784	(0.4303)	-0.6636	(0.4406)
Inspector Workload (in Natural Log)	-0.0276	(0.2258)	-0.0227	(0.2360)
Cuisine Categories	Yes		Yes	
Regions	Yes		Yes	
Log Pseudolikelihood	-174.7422		-174.3860	
α (p-value)			0.0085 (0.3986)	
N	224		224	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The cuisine categories are: *Asian*, *Bar & Pub*, *British*, *Cafe*, *Chinese*, *European*, *French*, *Indian*, *Italian*, *Japanese*, *Seafood*, *Steak*. The administrative regions include: *Northern Ireland and Wales*, *East Counties*, *East Midlands*, *North East*, *North West*, *South East*, *South West*, *West Midlands*, *Yorkshire and Humberside*. We use the food quality evaluation variable based on Michelin and Good Food Guide (GFG) scores (weighted half for the Michelin score and half for the GFG score; see the data description section for details). α is a gamma-distribution parameter.

TABLE 7. Hurdle Poisson and Negative Binomial

	(1) Poisson		(2) NB		(3) Poisson		(4) NB	
Dependent Variable: $\mathbf{1}(Y_i = 0)$								
Food Quality (MG&GFG)	-0.0063	(0.1154)	-0.0073	(0.1164)				
Food Quality × Competition	0.1851	(0.1284)	0.1863	(0.1302)				
Food Quality (GFG Only)					-0.0570	(0.0857)	-0.0560	(0.0855)
Food Quality (GFG Only) × Competition					0.1042	(0.1111)	0.1045	(0.1107)
Competition	-0.6628**	(0.3345)	-0.6671**	(0.3372)	-0.6602	(0.4354)	-0.6601	(0.4343)
Wine List	-0.2551	(0.2684)	-0.2571	(0.2703)	-0.2378	(0.2686)	-0.2372	(0.2680)
Capacity	0.5642***	(0.1749)	0.5650***	(0.1759)	0.5491***	(0.1767)	0.5491***	(0.1763)
Multiple Chefs	-0.0501	(0.2175)	-0.0483	(0.2189)	-0.0304	(0.2191)	-0.0316	(0.2186)
Listed Price	0.0111	(0.0070)	0.0115	(0.0071)	0.0161**	(0.0067)	0.0159**	(0.0066)
Tripadvisor Online Service Rating	0.8862***	(0.2853)	0.8875***	(0.2867)	0.9135***	(0.2885)	0.9127***	(0.2880)
Food Quality Up or New (Binary)	0.0215	(0.2393)	0.0192	(0.2405)	0.0707	(0.2417)	0.0706	(0.2412)
Food Quality Down (Binary)	0.3287	(0.2558)	0.3319	(0.2581)	0.4158	(0.2531)	0.4139	(0.2522)
Inspector Workload (in Natural Log)	-0.4192	(0.2599)	-0.4262	(0.2610)	-0.4323*	(0.2613)	-0.4298*	(0.2609)
Dependent Variable: $Y_i > 0$								
Food Quality (MG&GFG)	0.4974***	(0.1820)	0.4892***	(0.1702)				
Food Quality × Competition	-0.1568	(0.1908)	-0.1545	(0.1879)				
Food Quality (GFG Only)					0.3818***	(0.1229)	0.3818***	(0.1059)
Food Quality (GFG Only) × Competition					-0.2686*	(0.1417)	-0.2688**	(0.1319)
Competition	0.7844**	(0.3722)	0.7709**	(0.3515)	1.3723***	(0.5096)	1.3756***	(0.5012)
Wine List	-0.3668	(0.3798)	-0.3696	(0.2888)	-0.3931	(0.3824)	-0.3918	(0.3708)
Capacity	0.0539	(0.2427)	0.0478	(0.2364)	0.0609	(0.2503)	0.0621	(0.2330)
Multiple Chefs	-0.7604**	(0.3614)	-0.7708*	(0.4389)	-0.7647**	(0.3708)	-0.7622***	(0.2639)
Listed Price	-0.0160	(0.0108)	-0.0144	(0.0115)	-0.0141	(0.0096)	-0.0141*	(0.0073)
Tripadvisor Online Service Rating	-0.1783	(0.2902)	-0.1914	(0.3169)	-0.1547	(0.2796)	-0.1537	(0.1864)
Food Quality Up or New (Binary)	-0.3038	(0.2648)	-0.3035	(0.1885)	-0.2367	(0.3035)	-0.2363	(0.2819)
Food Quality Down (Binary)	-0.6554	(0.4270)	-0.6760	(0.4322)	-0.3471	(0.3655)	-0.3444	(0.3516)
Inspector Workload (in Natural Log)	-0.0167	(0.2283)	0.0039	(0.2434)	0.0157	(0.2210)	0.0174	(0.1827)
Cuisine Categories	Yes		Yes		Yes		Yes	
Regions	Yes		Yes		Yes		Yes	
Log Pseudolikelihood	-619.1679		-619.0749		-619.7163		-619.6777	
α (p-value)			0.0055 (0.6663)				3.9825e-07 (0.7811)	
N	852		852		852		852	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The cuisine categories are: *Asian*, *Bar & Pub*, *British*, *Cafe*, *Chinese*, *European*, *French*, *Indian*, *Italian*, *Japanese*, *Seafood*, *Steak*. The administrative regions include: *Northern Ireland and Wales*, *East Counties*, *East Midlands*, *North East*, *North West*, *South East*, *South West*, *West Midlands*, *Yorkshire and Humberside*. The cuisine category and region variables are included for each part of the hurdle model. For (1) and (2), we use the food quality evaluation variable based on Michelin and Good Food Guide (GFG) scores (weighted half for the Michelin score and half for the GFG score; see the data description section for details). For (3) and (4), we use the food quality scores from the Good Food Guide (GFG). α is a gamma-distribution parameter.

TABLE 8. Selection Poisson and Negative Binomial (NB): Second Stage

	(1)		(2)		(3)		(4)	
Dependent Variable: Y_i	Poisson		NB		Poission		NB	
Food Quality (MG&GFG)	0.2241**	(0.1131)	0.2218*	(0.1146)				
Food Quality (MG&GFG) × Competition	-0.1381	(0.1150)	-0.1478	(0.1134)				
Food Quality (GFG Only)					0.2081***	(0.0793)	0.2178***	(0.0816)
Food Quality (GFG Only) × Competition					-0.1039	(0.0913)	-0.1101	(0.0916)
Competition	-1.1553	(0.8488)	-1.1340	(0.8655)	-1.3391	(0.9876)	-1.3333	(1.0144)
Wine List	-0.0005	(0.2076)	0.0188	(0.2100)	-0.0058	(0.2075)	0.0070	(0.2103)
Capacity	-0.2587*	(0.1493)	-0.2814*	(0.1511)	-0.2197	(0.1510)	-0.2381	(0.1535)
Multiple Chefs	-0.2008	(0.1717)	-0.1795	(0.1716)	-0.2114	(0.1714)	-0.1921	(0.1712)
Listed Price	-0.0146**	(0.0068)	-0.0148**	(0.0067)	-0.0171***	(0.0055)	-0.0180***	(0.0056)
Tripadvisor Online Service Rating	-0.7693***	(0.2674)	-0.8372***	(0.2762)	-0.7748***	(0.2693)	-0.8542***	(0.2789)
Food Quality Up or New (Binary)	-0.0601	(0.2041)	-0.0745	(0.2023)	-0.1344	(0.2103)	-0.1558	(0.2077)
Food Quality Up or New (Binary)	-0.4571*	(0.2444)	-0.4735**	(0.2398)	-0.4419*	(0.2337)	-0.4664**	(0.2296)
Inspector Workload (in Natural Log)	0.2846	(0.1970)	0.2624	(0.2086)	0.2929	(0.1950)	0.2735	(0.2057)
u	1.8399**	(0.8138)	1.8626**	(0.8342)	2.1145**	(0.9320)	2.1580**	(0.9631)
Cuisine Categories	Yes		Yes		Yes		Yes	
Regions	Yes		Yes		Yes		Yes	
Log Pseudolikelihood	-648.1218		-639.8139		-647.3854		-639.1044	
α (p-value)			0.5883 (0.0000)				0.5842 (0.0000)	
N	852		852		852		852	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors in parentheses. The cuisine categories are: *Asian*, *Bar & Pub*, *British*, *Cafe*, *Chinese*, *European*, *French*, *Indian*, *Italian*, *Japanese*, *Seafood*, *Steak*. The administrative regions include: *Northern Ireland and Wales*, *East Counties*, *East Midlands*, *North East*, *North West*, *South East*, *South West*, *West Midlands*, *Yorkshire and Humberside*. For (1) and (2), we use the food quality evaluation variable based on Michelin and Good Food Guide (GFG) scores (weighted half for the Michelin score and half for the GFG score; see the data description section for details). For (3) and (4), we use the food quality scores from the Good Food Guide (GFG). u is the residual term calculated in the first stage. α is a gamma-distribution parameter.

TABLE 9. Selection Poisson and Negative Binomial: First Stage

Dependent Variable: Competition	(1) & (2)		(3) & (4)	
Food Quality (MG&GFG)	0.0675***	(0.0135)		
Food Quality (GFG Only)			0.0360***	(0.0111)
Tripadvisor Online Service Rating	-0.1566***	(0.0516)	-0.1333**	(0.0519)
# of Tripadvisor Reviews (in Natural Log)	0.0792***	(0.0205)	0.0823***	(0.0206)
Michelin & GFG Past Listed Years	-0.0369***	(0.0065)	-0.0284***	(0.0062)
N	852		852	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. A constant term is included. For (1) and (2), we use the food quality evaluation variable based on Michelin and Good Food Guide (GFG) scores (weighted half for the Michelin score and half for the GFG score; see the data description section for details). For (3) and (4), we use the food quality scores from the Good Food Guide (GFG).

Table 10. Information Criteria

Specification	Log Pseudolikelihood	# of Parameters	AIC	BIC	HQIC
Poisson	-652.0225	33	1370.0	1526.7	1430.1
Negative Binomial (NB)	-642.8283	34	1353.7	1515.1	1415.5
Hurdle-Poisson	-619.1679	66	1370.3	1683.7	1490.3
Hurdle-NB	-619.0749	67	1372.1	1690.2	1494.0
Selection-Poisson	-648.1218	34	1364.2	1525.7	1426.1
Selection-NB	-639.8139	35	1349.6	1515.8	1413.3

Notes: The sample size is 852. Akaike Information Criterion (AIC) $= -2\ln L + 2k$, Bayesian Information Criterion (BIC) $= -2\ln L + k\ln(n)$, and Hannan-Quinn Information Criterion (HQIC) $= -2\ln L + 2k\ln(\ln(n))$, where $\ln L$ is the log (pseudo) likelihood, and k is the number of parameters, and n is the number of samples. For the reported statistics in this table, we use the food quality evaluation variable based on Michelin and Good Food Guide (GFG) scores (weighted half for the Michelin score and half for the GFG score; see the data description section for details).