The Effect of Quality Disclosure on Firm Entry and Exit Dynamics: Evidence from Online Review Platforms

Ying Bao* Limin Fang† Matthew Osborne‡§

First draft: November 23, 2023
This draft: September 26, 2024

Click here for the most recent version

Abstract

This paper examines the impact of quality disclosure on the entry and exit dynamics of firms in an industry with many firms. We develop a theoretical model highlighting two key forces: (1) direct effect, where quality disclosure influences consumers' preferences for a firm by altering perceptions of its quality, and (2) competition effect, where quality disclosure alters a firm's competitive environment by affecting perceptions of competitors' qualities. Depending on which force dominates, different scenarios emerge. For example, quality disclosure can either intensify competition so much that high-quality firms are discouraged from entry, or reduce competition such that even low-quality firms are encouraged to enter. We test the model predictions using a unique dataset on restaurant entries, exits, and online reviews in Texas from 1995 to 2015. The findings reveal varied outcomes across market types: in college towns, all firms are discouraged from entry, whereas in highway-exit markets, all firms are encouraged to enter. In other markets, high-quality firms are encouraged to enter while low-quality ones are deterred. Across all market types, young high-quality independent firms tend to stay longer, while low-quality ones exit sooner, with no significant impact on chain or established firms.

*University of Illinois Urbana-Champaign, Gies College of Business. Email: ybao@illinois.edu.

†University of British Columbia, Sauder School of Business. Email: limin.fang@sauder.ubc.ca.

‡University of Toronto, Rotman School of Management. Email: matthew.osborne@rotman.utoronto.ca.

§We thank Chiara Farronato and Xinyao Kong for their helpful comments on earlier drafts of the paper. We also appreciate the invaluable feedback from conference participants at the Asia-Pacific Industrial Organisation Conference (APIOC), Digital Economics Conference, NBER Digital Economics and AI Meeting, Marketing Strategy Consortium, Marketing Dynamics Conference, and Summer Institute in Competitive Strategy (SICS) Conference, as well as from seminar participant at the University of British Columbia, Shanghai University of Finance and Economics, Chinese University of Hong Kong (Shenzhen), and University of Toronto. We greatly benefited from discussions with Avi Goldfarb, Przemyslaw Jeziorski, Miguel Villas-Boas, Heski Bar-Isaac, Yi Qian, and Aric Rindfleisch. All errors are our own.

1 Introduction

Information on product quality plays a critical role in consumer purchase decisions. In the digital age, an abundance of resources is at consumers' fingertips to search for product information, such as online reviews, social media influencers, images or videos, consumer reports, etc. Searching for quality information has become an integral part of the consumer shopping journey. According to statistica.com, over 90% of consumers read reviews before buying a product (van Gelder, 2023). Given the crucial role of quality information in consumer decisions, policymakers are mandating quality disclosure in various domains, such as restaurant hygiene grade, automobile manufacturer fuel efficiency measure, and mortality rates for hospitals (Dranove and Jin, 2010). These phenomena have intrigued academic researchers to study the impact of quality disclosure on a number of market outcomes, including consumer learning (e.g. Fang, 2022; Luca, 2016; Wu et al., 2015), product sales (e.g. Chevalier and Mayzlin, 2006; Zhu and Zhang, 2010; Hollenbeck, 2018), firms' strategic voluntary disclosure decisions (e.g. Guo and Zhao, 2009; Hotz and Xiao, 2013) and quality improvement (e.g. Jin and Leslie, 2003).

However, few studies have systematically examined the effect of quality disclosure on the entry and exit dynamics of firms in an entire industry. Understanding this effect is important because it helps policymakers assess the full socioeconomic impact of policies related to quality information (e.g. disclosure mandates, bans on fake reviews). Firm entry and exit dynamics directly affect the number and variety of products available to consumers. They also impact the market environment in which firms operate. Extending to the broader economy and society, these dynamics influence employment and business turnover in local communities and affect the quality of life.[1]

To shed light on this understudied area, this paper investigates the effect of quality disclosure on firm entry and exit, both theoretically and empirically. On the theory side, we develop a model of firm entry and exit in an industry with many firms, where consumers are unsure of firm quality and rely on quality information to learn. The model differentiates the effect across firms' quality levels, age, and chain affiliation. It yields a number of insightful predictions about how market characteristics relate to the types of firms that enter or exit. We then test these predictions using data from the restaurant industry in Texas. In the empirical application, quality information is provided through online reviews. We investigate the effect of online reviews on restaurant entry and exit in Texas.

Our theory highlights the interplay between two mechanisms through which quality information drives firms' profits and in turn, market dynamics. One is the *direct effect*, which is the effect on firm entry and exit through the change in consumers' preferences for a firm upon learning its quality. The other is the *competition effect*, which is the effect on entry and exit through the

[1]For small businesses, in particular, turnover can have a substantial impact on the cultural and social life of a local community as small restaurants and shops often serve as the locale where local residents socialize with each other (Porter, 2015; Waxman, 2006; Jacobs, 1962).

change in a firm's competitive environment once consumers learn its competitors' qualities. The direct effect is very intuitive: if a firm is of high quality, then quality information allows consumers to recognize the true quality, and consumers will like that firm more than before. However, whether the firm's profits will increase or not also depends on its competitors' revealed qualities. If most competitors are high-quality, then quality disclosure reveals their true qualities to consumers too, and competition intensifies. If the increase in competitive intensity is sufficiently large, then the focal firm's profit could actually decrease. On the other hand, if most competitors are low-quality, the competition faced by the focal firm may diminish, and its profit could increase as a result. By affecting profits, this change in the composition of firms in the broader market can influence firms' entry and exit decisions. We refer to this mechanism as the competition effect.

The overall impact of quality disclosure on firm entry and exit is the sum of the direct effect and the competition effect. These effects may have different signs. Thus, depending on which effect dominates, three scenarios can arise: one being intuitive, and the other two counterintuitive. If the direct effect dominates the competition effect, then the intuitive scenario occurs, where quality disclosure encourages the entry of high-quality firms but discourages the low-quality ones, and it speeds up the exit of young low-quality independent firms but keeps young high-quality independent firms in the market for longer. If the competition effect dominates the direct effect, then counterintuitive scenarios can occur. If quality disclosure significantly increases competition, then it could discourage the entry of all firms, including high-quality firms. However, if quality disclosure substantially decreases competition, then it could encourage the entry of all firms, even low-quality firms. Regarding exit, the effects in the counterintuitive scenarios are the same as those in the intuitive scenario. To better understand under what circumstances these different scenarios may arise, we conduct numerical simulations where we compute equilibrium entry and exit probabilities from the model for a wide range of parameters (for example, entry costs or the quality differentials across firms). Graphs of the differences in entry probabilities with and without quality disclosure demonstrate for what values of underlying model parameters, such as entry costs, we would expect to see these different scenarios arise. The simulation exercises thus provide guidance for our empirical analysis. For example, the numerical simulations suggest that one of the counterintuitive scenarios (discouraged entry of high-quality firms) can occur when the entry cost is low and the quality difference between high- and low-quality firms is large; the other counterintuitive scenario (encouraged entry of all firm types) occurs when there are many low quality potential entrants, and quality differentials are low.

We formalize the predictions described above into four hypotheses, which we test using data covering the restaurant industry in Texas during a time period when online review platforms were becoming more and more widespread. Specifically, we examine the effect of online review platform penetration on restaurant entry and exit. The restaurant industry is an excellent context for studying the role of quality disclosure. This industry features many small differentiated firms, and online reviews have large impacts on consumers' choices of restaurants.[2] For this study, we constructed a unique data set with a complete record of the entry and exit timing of full-service restaurants in Texas from 1995 to 2015. To gauge consumers' usage of online reviews, we also gathered review information from three major platforms: Google, Yelp and TripAdvisor. These online review platforms penetrated different regions in Texas at different times. This variation allows us to tease out the effect of online review platform penetration on the entry and exit of restaurants.

Following the insights from the theoretical analyses on the economic conditions that give rise to different scenarios, we divide the restaurant markets in Texas into three types: college-town markets, highway-exit markets, and the other markets. In college-town markets, where entry costs are low, but the quality range among restaurants is high, we discover that the penetration of online review platforms discourages the entries of all restaurants, including high-quality ones. However, in highway-exit markets, which are dominated by low-quality restaurants, the reverse holds true: the penetration of online review platforms increases the entry of all restaurants, including low-quality ones. These findings are consistent with the counterintuitive scenarios that are predicted by our theoretical model. In other markets, the empirical evidence is consistent with the intuitive scenario: the penetration of online review platforms encourages (discourages) the entry of high (low)-quality independent restaurants. There is very little effect on chain restaurants' entry. On the exit side, our findings support the model predictions that young high-quality independent restaurants stay in the market longer, while young low-quality independent restaurants exit more quickly. No statistically significant effect is detected for either chain or established independent restaurants. These empirical findings have implications about the improvement of the overall quality of restaurants in the market. We show using additional analysis that doubling consumers' review activity on online platforms increases the share of high-quality restaurants by about two percentage points, leading to an overall rise in the average rating of restaurants.

The findings in our paper have important policy and managerial implications. For policymakers, our results suggest that increasing the availability of quality information will enhance the average quality of firms in the market by weeding out low-quality firms. This is a positive outcome. However, policymakers need to be wary that in certain markets, such as highway-exit markets, heightened business turnover can be harmful for general business health and employment stability. For managers, our findings illuminate the effect of quality information on the competitive environment, which is a crucial factor for firms to consider when choosing markets. For example, managers of high-quality firms may want to reconsider entry into markets with abundant quality information, as such markets may have fiercer than expected competition.

In this context, our paper's contribution is two-fold. First, we develop a novel theoretical model that elucidates the direct and competition effects of quality disclosure. To the best of

[2]As shown in a survey by Tripadvisor, a staggering 94% of American diners acknowledge the significant influence of online reviews (Tripadvisor.com, 2018).

our knowledge, this paper is the first that studies these effects. Second, our study is the first empirical study that provides insights on the effect of online reviews on firms' entry behaviors. Prior literature (e.g. Donati (2022) and Anenberg et al. (2019)) has examined the effect on exit, but none has investigated the impact on entry, likely due to data limitations. Using a unique dataset that includes a complete record of restaurants in Texas over a 21-year horizon, our paper is well-positioned to investigate this important topic. We quantify the effects of online reviews on firm dynamics in the restaurant industry and document a number of novel empirical findings that can guide policymakers and firm managers in making informed public policy and strategic decisions.

The rest of the paper is organized as follows. In Section 2, we review the related literature. In Section 3, we set up the model and provide characterizations of the equilibrium solutions. In Section 4, we develop and prove comparative statics and provide numerical examples. In Section 5, we empirically test our model predictions using data on the restaurant industry in Texas. In Section 6, we conclude and discuss policy and managerial implications.

2 Literature Review

Our paper contributes to several strands of literature. First, it expands the body of research on quality disclosure and its impact on market outcomes. Previous studies have predominantly focused on the effects of mandatory disclosure on firms' quality improvement, sales, market shares, and voluntary disclosure decisions. Theoretically, Hotz and Xiao (2013) demonstrate that disclosure leads to more elastic demand and more intense price competition. Guo and Zhao (2009) and Hopenhayn and Saeedi (2023) show that competition and dispersed price distribution may incentivize strategic concealment, reducing the extent of information disclosure. Empirically, Jin and Leslie (2003) find that the disclosure of restaurant hygiene leads to significant improvements in hygiene quality in Los Angeles. In the healthcare industry, Jin and Sorensen (2006) and Dafny and Dranove (2008) provide evidence that public ratings of health plans increased the market share of higher-rated plans in HMO markets. Additionally, Dranove and Sfekas (2008) show that low-rated hospitals experienced a decline in market share due to quality report cards. So far, few studies have explored the effects of quality disclosure on market dynamics. A particularly relevant study is Hui et al. (2023), which examines the impact of eBay's stricter certification thresholds on the distribution of quality and incumbent behavior. In contrast, our paper contributes to the literature by systematically analyzing the effects of quality disclosure on firm entry and exit behaviors in an entire industry, and by providing insights into the interplay between the direct and competitive effects of quality disclosure.

Second, our paper contributes to the research on online reviews and their impact on market outcomes. Unlike government-mandated quality disclosure policies, online reviews represent a user-initiated alternative approach to quality disclosure. Users on online platforms rate firms' quality and share their experiences through written reviews. Much of the existing literature has focused on the empirical impact of online reviews on business demand and sales (Godes and Mayzlin, 2004; Moe and Trusov, 2011; Li and Hitt, 2008). For example, Chevalier and Mayzlin (2006) examine the effect of reviews on book sales and find a significant positive impact. Anderson and Magruder (2012) and Luca (2016) find that an additional half-star rating can lead to a substantial increase in restaurant revenue. Researchers have also explored this topic from the perspective of consumer learning. For instance, Zhao et al. (2013) investigate how consumers learn from book reviews as well as from their own experiences. Wu et al. (2015) and Fang (2022) study consumers' learning behaviors from online review platforms in the restaurant industry and estimate the economic value and consumer welfare generated by these platforms. Numerous studies have highlighted the heterogeneous effects of online reviews across various industries, including entertainment industries such as video games, movies, and music (Zhu and Zhang, 2010; Chintagunta et al., 2010; Dhar and Chang, 2009), online stores (Newberry and Zhou, 2019; Forman et al., 2008), and hospitality sectors such as hotels and restaurants (Luca, 2016; Hollenbeck, 2018; Fang, 2022). For example, Luca (2016) finds that the positive effect of Yelp reviews is primarily driven by independent restaurants rather than chains. Fang (2022) shows that online reviews have a greater impact on the revenues of restaurants in tourist areas compared to those in local areas dominated by repeat customers. A few studies have examined firms' responses to online reviews and the challenges with platforms' rating aggregation algorithms. For example, Farronato and Zervas (2022) find that restaurants are more likely to improve their hygiene if they are more exposed to online reviews. Donati (2022) explores how easier access to online review platforms influences consumer behavior and restaurants' incentives to upgrade quality.

To date, only a few studies have examined the effect of online reviews on the entry and exit dynamics of firms. Theoretically, a notable paper is Vellodi (2018), which investigates the impact of platforms' rating designs on firms' incentives to participate in the market. The study shows that consumers' queuing at capacity-constrained, highly-rated firms can raise entry barriers and accelerate the exit of firms that initially receive low ratings. The paper suggests suppressing the reviews of highly-rated firms to encourage market entry. Empirically, Dendorfer and Seibel (2024) examine the welfare loss caused by the "cold-start" problem on Airbnb by considering hosts' entry and exit, and they recommend lower rental rates for listings with few reviews to mitigate the issue.[3] Although this study accounts for entry and exit dynamics in the welfare calculation, it does not specifically test the effect of online reviews on firm entry and exit. Two other relevant studies, Donati (2022) and Anenberg et al. (2019), have explored the impact of online reviews on firm exits, both finding that online reviews accelerate the exit of low-quality firms. Compared to these studies, our paper is the first to *empirically* examine the effect of online reviews on firm *entry* behavior. This is primarily due to the advantage of our dataset, which covers the universe of full-service restaurants in Texas over a long horizon (21 years). Moreover, because Texas is geographically

[3]The "cold-start" problem refers to the inefficiency in review platforms where less frequently purchased products have fewer reviews, leading to even fewer purchases.

large, we have enough data to identify the way the impact of review platform penetration varies across different markets.

Finally, and more broadly, our research contributes to the literature on dynamic oligopoly games of entry and exit, particularly through the application of the OE concept to simplify dynamic games. The OE concept was proposed by Weintraub et al. (2008), who demonstrate that in industries with many small firms applying OE allows researchers to bypass the curse of dimensionality in analyzing Ericson and Pakes (1995)-style dynamic models. Moreover, they show that the OE approximates the Markov perfect equilibrium (MPE) well. Due to the tractability advantage of OE, several empirical papers have applied OE in various contexts. For example, Saeedi (2019) examines the effect of online reputation on adverse selection on eBay and estimates the value of reputation by applying the OE concept. Chen and Xu (2023) investigate R&D investment and knowledge spillover in the Korean electric motor industry, using OE to link individual plant R&D decisions to aggregate industry productivity and output. In our study, we use OE to significantly simplify our theoretical model, making analytically tractable comparative statics possible. We contribute to this literature by showcasing the broad applicability of the OE concept.

3 Model

Our model is based on the dynamic game framework introduced in the seminal work of Ericson and Pakes (1995) (EP), where firms are forward-looking and make entry and exit decisions over time. We develop our model in the context of an industry with many firms, chosen for two key reasons: first, industries significantly impacted by quality disclosure often consist of many small firms, such as restaurants and online apps. Second, modeling in this type of industry context allows for substantial simplification using the oblivious equilibrium (OE) concept proposed by Weintraub et al. (2008). The OE concept has been shown to closely approximate Markov perfect equilibria in industries with many firms (Weintraub et al., 2008). The use of this concept greatly simplifies our model and facilitates the development of analytical comparative statics. To provide a concrete setting, we focus on the restaurant industry as the specific context. However, the insights developed here are broadly applicable to other industries with many firms. Below, we present the detailed setup of our model.

3.1 Firms, Actions, Profits and States

Our model includes high- and low-quality restaurants as well as chain and independent restaurants. Every period potential entrants decide whether to enter the market or not. The types of entrants in terms of their quality levels $T \in \{H, L\}$ and chain affiliation $D \in \{I, c\}$ are determined before entry. The quality q for high-quality restaurants (either independent or chain) is $\bar{q}$, and the quality for low-quality restaurants is $\underline{q}$. The quality q is set before entry and does not change after entry. The number of potential entrants from each type is N_H for high-quality independent restaurants,

N_L for low-quality independent restaurants, N_{Hc} for high-quality chain restaurants, and N_{Lc} for low-quality chain restaurants. The number of potential entrants from each type is assumed to be exogenously given. Once potential entrants enter the market, they become incumbents. Every period, an incumbent decides whether to exit the market or not. Once they exit the market, they cannot reenter.

Perceived Quality of Firms Incumbents' profits depend on consumers' perception of their quality. For chain restaurants, the quality is assumed to be known to consumers as soon as they open because they benefit from a chain's reputation. For independent restaurants, however, the quality is unknown initially, and consumers need to learn about their quality over time. How much consumers learn depends on a parameter, $\gamma \in [0, 1]$, which measures the amount of information on quality. This information comes from word of mouth (WOM), such as reviews on online platforms or recommendations from friends and family. At extremes, 0 means that there is no information and therefore no quality disclosure at all; 1 means abundance of information and full disclosure.

Depending on how long independent restaurants have been in the market, they experience three age-related stages: new, young and established. At each stage, the perceived quality by consumers, denoted as $\hat{q}$, varies. At the new stage, consumers assign an exogenous prior $\hat{q}_0$ to the restaurant's quality, which is assumed to be $\hat{q}_0 \equiv \frac{1}{2}(\bar{q} + \underline{q})$.[4] After being in the market for one period, a new restaurant reaches the young stage, where consumers have partially learned its quality. The perceived quality at this stage is denoted by $\hat{q}_1$, and how close it is to the true quality depends on the level of information γ. We assume that it is somewhere between the prior $\hat{q}_0$ and the true quality q: $\hat{q}_1 = \hat{q}_0 + (q - \hat{q}_0)\gamma, \forall q \in \{\underline{q}, \bar{q}\}$. If $\gamma = 0$, then $\hat{q}_1 = \hat{q}_0$. If $\gamma = 1$, then $\hat{q}_1 = q$. If $0 < \gamma < 1$, then $\hat{q}_1$ is in between, and the greater the γ, the closer $\hat{q}_1$ is to the true quality q. After being at the young stage for one period (i.e., in the market for two periods in total), the young restaurant enters the established stage, where consumers learn the true quality of the restaurant; in other words, the perceived quality of an established restaurant is its true quality q. We assume that once a restaurant's true quality is revealed, consumers' beliefs of its quality will stay at the true quality forever.

Our formula for $\hat{q}_1$ represents the population average of beliefs across all consumers who may experience different belief updates; for example, some consumers may have received consistently low-quality signals for a restaurant, and other consumers may have received consistently high-quality signals. The formula for $\hat{q}_1$ reflects that in aggregate, more consumers received the correct signals than those who received false ones. Therefore, $\hat{q}_1$ is always closer to the true quality than the

[4] The priors can be set anywhere between $\bar{q}$ and $\underline{q}$. The key assumption here is that it is exogenously determined. We assume the exogenous prior for tractability. Introducing a new stage in an independent restaurant's life cycle is to reflect the reality that it usually takes time for information to accumulate, and earlier signals during this stage may be too sparse and noisy for consumers to make any meaningful update to their beliefs. As a result, consumers are initially very uncertain about a new restaurant's quality, and their beliefs are assumed to be their priors.

initial prior $\hat{q}_0$ and becomes even closer to the true quality as γ increases.[5] The micro foundation for this formulation based on the Bayesian learning process is provided in Fang (2022).[6] Our demarcation of stages into new, young and old is an abstract way to capture the consumer learning process over time. In reality, it may take consumers many weeks or months to fully learn the quality of a restaurant. One could interpret the length of two periods in our theoretical model as capturing the total amount of calendar time it takes a consumer to fully (or almost fully) learn quality.

Firm Profits Every period, firms take the market structure as given and engage in static price competition with differentiated products.[7] We define firms' profit function from static competition as follows:

$$\pi(\cdot) = \frac{M \exp(\hat{q})}{\sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}') + 1} - C(q, D), \tag{1}$$

where $\hat{q} \in \{\hat{q}_0, \hat{q}_1, q\}$ is the restaurant's own perceived quality. $\hat{q}'$ denotes the perceived quality of any given restaurant in the market. $n(\hat{q}')$ represents the number of restaurants with perceived quality $\hat{q}'$ in the market. M represents the market size, measured in dollar value (it is the product of the number of potential consumers and a constant markup). $D \in \{I, c\}$ is chain affiliation indicator; I means independent and c, chain. C is the per-period fixed cost of the restaurant that depends on the restaurant's true quality and its chain affiliation, such as rent and the chef's salary.

This profit function is derived from the logit demand formula based on a micro demand model. The details are provided in Online Appendix G.[8] There are some important aspects of this profit function to highlight. First, $\hat{q}$ measures a consumer's indirect utility from visiting restaurant j, net of the logit error term: in other words, it is underlying quality minus price, under the normalization the price coefficient is one. This formulation aligns with what consumers ultimately care about, which is quality net of price, and reflects that fact that platforms often present quality adjusted for price. Second, firms have constant markups. Constant markups arise when static price competition is monopolistic competition (Benassy, 1996): there are many small and differentiated firms in the market, where firms have limited pricing power, but due to differentiation, can receive positive markups. In Online Appendix G, we show constant markups arise in our setting. Third, firms are assumed to have the same marginal costs in the model, such that consumers cannot infer their

[5] Another simplification made here is that we do not take into account the effect of a restaurant's sales on consumers' perception of its quality. For example, greater sales may generate more quality signals. Here, we assume that all restaurants share the same γ.

[6] Please see Section 4.2 of Fang (2022).

[7] Static competition is a common assumption in the EP framework. That is, we abstract away from firms' dynamic considerations when choosing prices, such as predatory pricing or pricing low at the beginning in order to encourage more consumers to learn about firms' quality.

[8] Note that this profit formula models consumers' demand for new and established independent restaurants exactly, but for young independent restaurants, it is an approximation. The profit function in Equation (1) approximates the aggregation of demand from consumers with heterogeneous beliefs of young independent restaurants' quality by using the demand from a representative consumer who has the belief at the population average.

expected utility from a restaurant based on its price. This assumption is not restrictive.[9]

Assumptions on Timing of Events Within Each Period Firms choose whether to be active or inactive in the market each period. Let $a \in \{0, 1\}$ denote a firm's action, with 0 representing inactive and 1 active. In each period, events in this dynamic game are assumed to occur as follows:

1. At the beginning of each period, potential entrants and incumbents simultaneously make entry and exit decisions. If a potential entrant decides not to enter, it can choose to enter in the next period.
   - If a potential entrant decides to enter ($a = 1$), it incurs an entry cost of $\kappa(q, D) + \varepsilon$, where $\kappa(q, D) > 0$ is the average entry cost incurred by a restaurant with quality q and chain affiliation type D. $\varepsilon \sim U(-b_0, b_0)$ is a private random shock, and is independent and identically distributed (i.i.d.) across time and firms. The potential entrant observes the shock first before making the entry decision.
   - If an incumbent decides to exit, it receives a scrap value of $\phi \sim U[-b_1, b_1]$, which is a private shock only known to the firm itself and is observed before the decision of exit. It is i.i.d. across time and firms. An incumbent that exits stays out of the market forever.
2. Firms' decisions take effect right away, and the market structure changes.[10] The new entrants and the remaining incumbents engage in static competition and make profits according to Equation (1). A firm that is not in the market earns 0 profit.

States Given this setup, the payoff relevant states include the private shocks to the entry cost and scrap value (ε, ϕ), a firm's own quality q, the firm's chain affiliation D, and the firm's stage, denoted by $g \in \{0, 1, 2\}$, where $\{0, 1, 2\}$ represent a firm's stages of new, young, and established, respectively. In addition, the distribution of firms in the market over the various perceived quality levels, $\boldsymbol{n} = \{n(\hat{q}')\}$, is particularly important because it summarizes the competitive environment faced by a firm. The parameters M, $\kappa(\cdot)$ and $C(\cdot)$ do not vary over time in our setting, and thereby omitted from the payoff relevant states. The payoff relevant state can thus be summarized by the tuple $(q, g, D, \boldsymbol{n}, \varepsilon, \phi)$.

[9]The model is equivalent to one where firms are assigned marginal costs that may be heterogeneous, and their average qualities are linear functions of marginal costs, but where each firm's deviation from its average quality is uncorrelated with its marginal cost. Consumers can infer a firm's average quality from its price due to constant markups, but they need to learn about the firm's deviation from the average quality. Therefore, even with heterogeneous marginal costs, consumers' indirectly utility can still be represented by $\hat{q}$ because we use a linear indirect utility function.

[10]The absence of the time-to-build assumption does not affect the predictions of the model.

3.2 Firm's Problem

A firm chooses a strategy (denoted by σ) to maximize the expected net present value of profit in every period. It takes the competitors' strategies (denoted by σ_-) as given. Following the literature, we focus on symmetric Markov strategies of firms. In an abuse of notation, we will use the shorthand $\widetilde{V}(s, \boldsymbol{n}, \varepsilon, \phi, a|\sigma) \equiv \widetilde{V}(s, \boldsymbol{n}, \varepsilon, \phi, a|\sigma, \sigma_-)$ to represent the action-specific value function that accounts for the expected present value of profits when a firm follows the same strategy σ as its competitors. A firm solves the following maximization problem:

$$\max_a \widetilde{V}(s, \boldsymbol{n}, \varepsilon, \phi, a|\sigma) = \Pi(s, \boldsymbol{n}, \varepsilon, \phi, a) + (\mathbb{1}\{g = 0\} + a\mathbb{1}\{g > 0\})\, \beta\mathbb{E}\left(\widetilde{V}(s', \boldsymbol{n}', \varepsilon', \phi', a'|\sigma, s, \boldsymbol{n}, a)\right), \tag{2}$$

where $s = (q, g, D)$, and $\Pi(s, \boldsymbol{n}, \varepsilon, \phi, a) \equiv a\pi(s, a) - a\mathbb{1}\{g = 0\}(\kappa(q, D) + \varepsilon) + (1 - a)\mathbb{1}\{g > 0\}\phi$. $\beta \in (0, 1)$ is a discount factor. The apostrophe ' denotes the next period. The law of motion for g is $g' = (g + a)\mathbb{1}\{g < 2\} + g\mathbb{1}\{g = 2\}$. The next period action $a' = \sigma(s', \boldsymbol{n}', \varepsilon', \phi')$. The next period $\boldsymbol{n}'$ incorporates the entry and exit behaviors of firms from different types and stages. The law of motion of $\boldsymbol{n}$ will be discussed in detail when we discuss equilibrium conditions of the model.

3.3 Oblivious Equilibrium

Given that the industry has many small firms, we use the concept of oblivious equilibrium (OE) instead of Markov perfect equilibrium (MPE). In an OE, firms take, as given, the time-invariant long-run average of the industry state $\boldsymbol{n}$ when making entry and exit decisions. In an abuse of notation, we still use $\boldsymbol{n}$ to represent this long-run average of the industry state. The OE in our context is defined as

$$\max_\sigma \widetilde{V}(s, \varepsilon, \phi|\sigma, \sigma_-^*, \boldsymbol{n}) = \widetilde{V}(s, \varepsilon, \phi|\sigma^*, \boldsymbol{n}), \tag{3}$$

where σ^* represents the equilibrium strategy of an OE.

This equation demonstrates that when competitors are playing the equilibrium strategy, a firm's optimal strategy is the equilibrium strategy. Weintraub et al. (2008) show that when the market size M is very large and when there are many small firms in the market, the chance of a single firm's action having a large impact on another firm's profit is very small. Therefore, firms only need to keep track of this time-invariant long-run average of the industry state $\boldsymbol{n}$ rather than individual competitors' states. Weintraub et al. (2008) label this as the "light-tailed" condition, and a number of assumptions of the model have to be satisfied in order to generate this condition. In Weintraub et al. (2008), logit demand is used as a special case to demonstrate compliance with all necessary assumptions for OE. Our model's microfoundation is based on logit demand, and we have verified that it adheres to these assumptions, including the "light-tailed" condition, ensuring its robustness and applicability in OE scenarios. Adopting the OE concept significantly simplifies the model solutions by avoiding the curse of dimensionality. In essence, the OE concept translates a dynamic game into a single agent problem. Nonetheless, in an OE, competitors' strategies are still relevant for a firm's optimal strategy; it is only that competitors' strategies affect a firm's decisions through impacting the long-run industry state $\boldsymbol{n}$, instead of competitors' individual states.

Firms' equilibrium strategies can be represented by entry and exit probabilities, because ε and ϕ are private shocks. Let $\mathbf{P}^* = (\mathbf{P^E}, \mathbf{P^X})$ denote the equilibrium probabilities associated with the optimal strategy σ^*, where the superscript E indicates entry and X represents exit. We present the equilibrium conditions of our model in Appendix A. Note that there are no closed-form analytical solutions for equilibrium entry and exit probabilities, although they can be solved numerically. Despite the lack of closed-form solutions, we can still sign the directions of γ's effects on entry and exit probabilities analytically for a wide range of parameters. In the following section, we discuss the analytical comparative statics of the OE solutions. For notation simplicity, in the remainder of the text, we express the (q, g, D) as subscripts and also omit the subscript I for independent restaurants, but still use c to indicate chain. For example, the exit probability of young high-quality independent restaurant, $P^X(\bar{q}, 1, I)$, will be denoted by P^X_{H1}, and that for chain, $P^X(\bar{q}, g, c)$, by P^X_{Hc}. We provide a detailed definition of each notation in Table 1.

Table 1: Table of Notation

Variable	Meaning
$q \in \{\underline{q}, \bar{q}\}$	true quality of restaurants, $\bar{q}$: quality of high-quality (H) restaurants; $\underline{q}$: quality of low-quality (L) restaurants
$g \in \{0, 1, 2\}$	stage of an independent restaurant, 0: new; 1: young; 2: established
$\hat{q} \in \{\hat{q}_0, \hat{q}_1, q\}$	perceived quality of a restaurant, $\hat{q}_0$: perceived quality of new entrants; $\hat{q}_1$: perceived quality of young restaurants; q: perceived (also true) quality of established restaurants
$D \in \{I, c\}$	chain affiliation c: chain; I: independent [1]
γ	level of information on quality
β	discount factor
M	market size in terms of dollar value
ϕ	scrape value, $\phi \sim U[-b_1, b_1]$
ε	random shock to the entry cost, $\varepsilon \sim U[-b_0, b_0]$
$T \in \{H, L\}$	type of the restaurant, H: high quality with $q = \bar{q}$; L: low quality with $q = \underline{q}$
N_T	number of potential independent entrants with quality type $T \in \{H, L\}$
N_{Tc}	number of potential chain entrants with quality type $T \in \{H, L\}$
κ_T	average entry cost for independent restaurants with quality type $T \in \{H, L\}$
κ_{Tc}	average entry cost for chain restaurants with quality type $T \in \{H, L\}$
π_{Tg}	flow profit for independent restaurants with quality type $T \in \{H, L\}$ at stage $g \in \{0, 1, 2\}$

[1] For notation simplicity, we dropped the subscription I for independent restaurants

4 Comparative Statics

The equilibrium outcomes of this model are the result of the interplay between two key forces: a direct effect and a competition effect. The sum of the two effects constitutes the total effect of γ on equilibrium probabilities. The direct effect is the impact of quality disclosure (γ) on a firm's entry and exit probabilities by influencing the firm's own perceived quality, $\hat{q}$. The competition effect is the impact of quality disclosure on equilibrium probabilities by altering the competitive environment of a firm. Quality disclosure affects competition through two channels: by changing the perceived quality of competitors and by affecting the number of firms in the market. Precise definitions of the direct and competition effects are provided in Definitions 1 and 2. These two effects work in the same direction for some types of restaurants but oppose each other for others. Depending on which effect dominates,[11] a number of scenarios can occur — some intuitive, others counterintuitive. To provide an intuitive illustration of the various scenarios, we conduct numerical simulations and discuss results in Section 4.3.

4.1 Definitions of Direct Effect and Competition Effect

Definition 1. The direct effect (DE) of γ on $\mathbf{P}^*_\mathbf{T}, \forall T \in \{H, L\}$ is $\frac{\partial \mathbf{P}^*_\mathbf{T}}{\partial \pi_{T1}} \frac{\partial \pi_{T1}}{\partial M \exp(\hat{q}_{T1})} \frac{\partial M \exp(\hat{q}_{T1})}{\partial \hat{q}_{T1}} \frac{d\hat{q}_{T1}}{d\gamma}$, and 0 on $\mathbf{P}^*_\mathbf{Tc}$.

Definition 2. The competition effect (CE) of γ on $\mathbf{P}^*_\mathbf{T}, \forall T \in \{H, L\}$ is $\left(\frac{\partial \mathbf{P}^*_\mathbf{T}}{\partial \pi_{T0}} \frac{\partial \pi_{T0}}{\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')} + \frac{\partial \mathbf{P}^*_\mathbf{T}}{\partial \pi_{T1}} \frac{\partial \pi_{T1}}{\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')} + \frac{\partial \mathbf{P}^*_\mathbf{T}}{\partial \pi_{T2}} \frac{\partial \pi_{T2}}{\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')} \right) \frac{d \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')}{d\gamma}$, and the CE on $\mathbf{P}^*_\mathbf{Tc}$ is $\frac{\partial \mathbf{P}^*_\mathbf{T}}{\partial \pi_{Tc}} \frac{\partial \pi_{Tc}}{\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')} \frac{d \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')}{d\gamma}$

In these definitions, $\mathbf{P}^*_\mathbf{T}$ is the vector of equilibrium probabilities of independent restaurants with quality level $T \in \{H, L\}$; for high quality, $\mathbf{P}^*_\mathbf{H} = (P^E_H, P^X_{H1}, P^X_{H2})$ and for low-quality, $\mathbf{P}^*_\mathbf{L} = (P^E_L, P^X_{L1}, P^X_{L2})$. Similarly defined, $\mathbf{P}^*_\mathbf{Tc}$ is the vector of equilibrium probabilities of chain restaurants.

As can be seen, the direct effect (DE) works through the numerator of the flow profit function, i.e., $M \exp(\hat{q})$. Because the numerators of π_{T0} (profit at the *new* stage) and π_{T2} (profit at the *established*-stage) are not functions of γ, their derivatives with respect to γ are zero. Only the numerator of π_{T1} (profit at the *young*-stage) yields a nonzero derivative because the perceived quality, $\hat{q}_{T1}$, is a function of γ. For chain restaurants, the numerators of the flow profits do not change with γ; thus, the direct effect for chain is 0.

Definition 2 of the competition effect (CE) states that CE operates through the effect of γ on the denominator of the profit function, i.e. $\frac{\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')}{\partial \gamma}$. For ease of interpretation, we refer to $\sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')$ as *competition* in the remaining text. Competition is affected by γ through two channels: one is that it discloses the quality of the competitors (i.e. $\hat{q}'$), and the other is that it changes the distribution of firms over the perceived quality levels in the market (i.e. $n(\hat{q}')$).

[11] As we describe below, when we characterize one effect as dominating another, we mean in absolute terms.

The latter channel is realized through the total effect on entry and exit, which is an equilibrium outcome. The sum of DE and CE constitutes the total effect of γ on equilibrium probabilities.[12]

4.2 The Impact of Quality Disclosure on Equilibrium Outcomes: Possible Scenarios

To determine the total effect of γ, we need to sign the DE and CE. For ease of presentation, let $DE(P)$ and $CE(P)$ denote the direct and competition effects, respectively, on a given equilibrium probability P; for example, $DE(P_H^E)$ denotes the direct effect of γ on P_H^E.[13] The signs for the direct effect are intuitive, as shown in Proposition 1:

Proposition 1. The DEs on chain restaurants are 0, and the DEs on independent restaurants have the following signs:

$DE(P_H^E) > 0$, $DE(P_{H1}^X) < 0$, $DE(P_{H2}^X) = 0$,

$DE(P_L^E) < 0$, $DE(P_{L1}^X) > 0$, $DE(P_{L2}^X) = 0$.

The proof of Proposition 1 is straightforward, and is shown in Online Appendix C.1.

We sign the impact of γ on the change in competition (the denominator of the profit function) in Proposition 2, which is presented in Appendix B.2, with proof in Online Appendix C.2. For simplicity of presentation, we omit the technical details of Proposition 2 in the main text and describe its implications. Proposition 2 shows that a higher γ can potentially increase or decrease competition faced by firms (i.e. $\sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}')$). The effect of γ has an effect on competition through two channels: by changing the perceived quality of firms ($\hat{q}'$) and by affecting the distribution of firms over perceived quality levels ($n(\hat{q}')$). The total effect of γ on this perceived quality distribution is the sum of the separate effects on low- and high-quality firms in the market, which we define as $\bar{F}_L + \bar{F}_H$ (formulas for these two terms are shown in Equations (C.9) and (C.10) in Online Appendix C.2.1). The sign of the impact of γ on $\bar{F}_L + \bar{F}_H$ mirrors its impact on competition. In particular, when a market is dominated by low-quality restaurants, competition is likely to decline with γ because γ has a negative impact on the perceived quality of low-quality firms. Conversely, if a market has many high-quality restaurants, competition is likely to increase with γ because more high-quality firms' true quality is revealed. We use this insight in choosing settings for the numerical examples in Section 4.3.

[12]This is because the equilibrium probabilities, $\mathbf{P}_\mathbf{T}^*$, can be expressed as a function of two elements: $\hat{q}_{T1}$ and $\sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}')$. The total derivative of the equilibrium probabilities, $\mathbf{P}_\mathbf{T}^*$, with respect to γ is then

$$\frac{d\mathbf{P}_\mathbf{T}^*(\hat{q}_{T1}, \sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}'))}{d\gamma} = \frac{\partial \mathbf{P}_\mathbf{T}^*}{\partial \hat{q}_{T1}}\frac{d\hat{q}_{T1}}{d\gamma} + \frac{\partial \mathbf{P}_\mathbf{T}^*}{\partial \sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}'))}\frac{d\sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}'))}{d\gamma}$$

The first element on the right hand side of this equation can be expanded into the same expression as the DE in Definition 1, and the second element can also be written as the CE in Definition 2.

[13]The functional form of $DE(P_H^E)$ is $DE(P_H^E) = \frac{\partial P_H^E}{\partial \pi_{H1}} \frac{\partial \pi_{H1}}{\partial M\exp(\hat{q}_{H1})} \frac{\partial M\exp(\hat{q}_{H1})}{\partial \hat{q}_{H1}} \frac{\partial \hat{q}_{H1}}{\partial \gamma}$.

An obvious corollary of Proposition 2 is that once the sign of the impact of γ on the change in competition is established, the competition effect is thus determined. We present this result in Corollary 1, in Appendix B.2.

Proposition 1 and Corollary 1 together demonstrate that under certain conditions, DE and CE work in the same direction for some equilibrium probabilities, while opposing each other for others. As a result, it is possible to sign the total effect of γ. We develop two additional propositions, 3 and 4, which show the set of all possible total effects given $\bar{F}_L + \bar{F}_H > 0$ and $\bar{F}_L + \bar{F}_H < 0$, respectively. We present the formal statements of these propositions in Appendix B.2 and proofs in Online Appendices C.3 and C.4, respectively. We note that it is easy to see that when $\bar{F}_L + \bar{F}_H = 0$, the effects of γ on all equilibrium probabilities are essentially the DEs, and the signs are as outlined in Proposition 1.

Proposition 3 and Proposition 4 cover all the possible scenarios that could arise as the result of quality disclosure. We summarize these scenarios in terms of changes in equilibrium probabilities in Table 2 for independent restaurants and Online Appendix Table C.1 for chain restaurants. As shown in these tables, for chain and established independent restaurants, the entry and exit probabilities are determined solely by the CE, meaning their changes align with the signs of $F_L + F_H$. For independent restaurants, the key difference between the scenarios lies primarily in the effects on the entry of independent restaurants. As indicated in Table 2, the cases in Proposition 3 (1), Proposition 4 (2) and Proposition 1 show the most intuitive effects: quality disclosure encourages the entry of high-quality independent restaurants and discourages that of low-quality restaurants. The other cases demonstrate the counterintuitive effects: Proposition 3 (2) suggests that high-quality independent restaurants can be discouraged from entry if quality disclosure makes the market too competitive, leading to a situation where all types of firms are discouraged from entry. Proposition 4 (1) illustrates that low-quality independent restaurants can be encouraged to enter the market if quality disclosure significantly reduces competition, resulting in a scenario where all types of firms are encouraged to enter.

Table 2: Summary of Effect of Quality Disclosure on Entry and Exit Probabilities of Independent Restaurants

Condition	**Case**	**High-quality**			**Low-quality**		
		P_H^E	P_{H1}^X	P_{H2}^X	P_L^E	P_{L1}^X	P_{L2}^X
$\bar{F}_L + \bar{F}_H > 0$	Proposition 3 (1)	↑	↓	↑	↓	↑	↑
	Proposition 3 (2)	↓	↓	↑	↓	↑	↑
$\bar{F}_L + \bar{F}_H = 0$	Proposition 1	↑	↓	-	↓	↑	-
$\bar{F}_L + \bar{F}_H < 0$	Proposition 4 (1)	↑	↓	↓	↑	↑	↓
	Proposition4 (2)	↑	↓	↓	↓	↑	↓

Note: The red arrows are the counterintuitive results. "-" means no change.

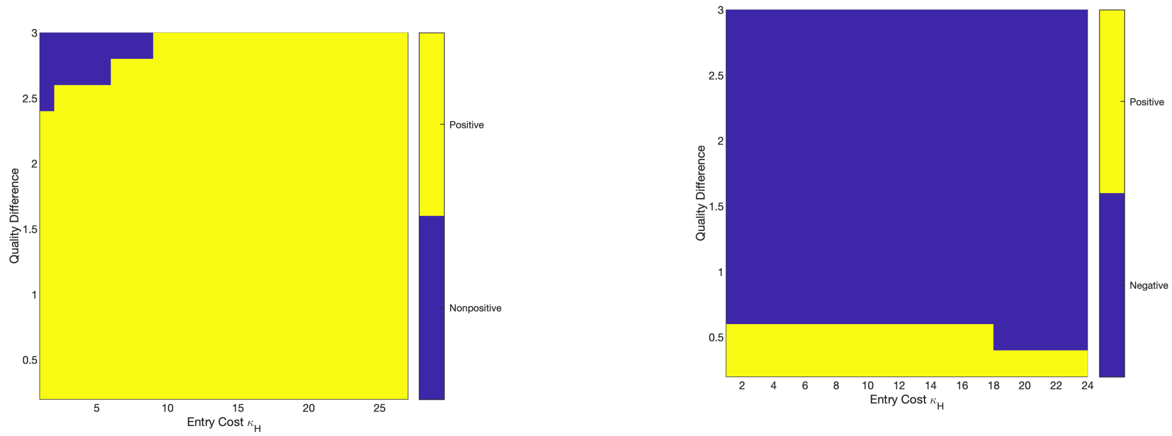
4.3 Numerical Examples

To illustrate the scenarios presented in Table 2 intuitively, we conduct numerical simulations. Based on the insights provided in Proposition 2, we choose two settings: one with $N_L = N_H$ and the other with $N_L >> N_H$. In the first setting, quality disclosure increases competition (i.e. $\sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')$) in most of the parameter space we explore. In the second setting, however, quality disclosure reduces competition in the entire parameter space. In Table D.1 of the Online Appendix, we specify the values and ranges of all model parameters used in the numerical examples presented below.

We focus mainly on the dimensions of the average entry costs and the difference between $\bar{q}$ and $\underline{q}$ for a few reasons: first, entry cost directly affects firms' entry decisions, which are the key differences between the various scenarios predicted by the model. Second, quality difference is equally crucial because it relates directly to the effect of quality disclosure — the bigger the quality difference, the larger the impact of quality disclosure. For quality difference, we normalize $\underline{q}$ to 0, and vary only $\bar{q}$. For the average entry costs, we vary the entry cost of high-quality independent restaurants (κ_H) and assume that the entry costs of low-quality independent and chain restaurants are of fixed proportions (0.7 and 0.8 respectively) to κ_H. This assumption is to reflect that in reality both low-quality independent and chain restaurants usually incur lower entry costs.

For all the numerical examples in this section, we compare the equilibrium probabilities when there is no disclosure (i.e. $\gamma = 0$) to those when there is full disclosure (i.e. $\gamma = 1$). In doing so, we illustrate the effect of γ on equilibrium probabilities. For both settings of $N_L = N_H$ and $N_L >> N_H$, we focus on the change in firms' entry probabilities because all other probabilities are perfectly predictable and the numerical results for them are consistent with the model.

Numerical Examples with $N_L = N_H$ In this case, quality disclosure increases competition (i.e. $\sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')$) for most of the parameter space under examination (The change in competition is shown in Figure D.D.1a in the Online Appendix). This case corresponds to the setting in rows 1 and 2 of Table 2. We display the change in the entry probabilities for high-quality independent restaurants (i.e., P_H^E) in Figure 1a . The vertical axis of the figure is quality difference, and the horizontal axis represents entry cost. As shown, for the majority of the parameter space in the figure (colored in yellow), high-quality independent restaurants are more likely to enter as a result of quality disclosure. This increase in the entry probability is consistent with the intuitive scenario in Proposition 3 (1). However, at the upper left corner of the graph (colored in blue), where the entry cost is very low but the quality difference is very high, high-quality independent restaurants are less likely to enter due to quality disclosure. This decrease in the entry probability is consistent with the counterintuitive scenario in Proposition 3 (2). The reason why this counterintuitive scenario occurs in this blue region is that the change in competition is the highest here. When the entry cost is lower, there are more firms in the market, resulting in fiercer competition with or without quality disclosure. When quality difference is higher, the increase in competition caused by quality



(a) $N_L = N_H$

(b) $N_L >> N_H$

Figure 1: Change in Entry Probabilities of Low-Quality Independent Restaurants

disclosure will be larger. The interaction of a small entry cost and a high quality difference results in a large increase in competition when γ changes from 0 to 1. Therefore, the CE dominates DE in the blue region, discouraging all firms from entry, including high-quality independent restaurants.

Numerical Examples with $N_L >> N_H$ In this case, quality disclosure reduces competition (i.e. $\sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')$) for all of the parameter space under examination, corresponding to the setting in rows 4 and 5 of Table 2 (The change in competition is shown in Figure D.D.1b in the Online Appendix). Competition decreases in this setting because the market is dominated by low-quality restaurants due to $N_L >> N_H$; once the true quality of low-quality restaurants is disclosed, competition in the market declines substantially. We display the change in the entry probabilities of low-quality independent restaurants (i.e., P_L^E) in Figure 1b. As shown, for the majority of the region of the parameter space (colored in blue), low-quality independent restaurants are less likely to enter as a result of quality disclosure. The outcomes in this region are consistent with the intuitive scenario in Proposition 4 (2). However, at the bottom of the graph (colored in yellow), where the quality difference is very low, low-quality independent restaurants are more likely to enter due to quality disclosure. The outcomes in this yellow region are consistent with the counterintuitive scenario in Proposition 4 (1). The rationale behind why this scenario occurs here is that the decline in the perceived quality of low-quality independent restaurants is the smallest in this yellow region, and its effect on entry can be easily outweighed by the impact of decreased competition on entry. As a result, in this yellow region, all firms are encouraged to enter the market, including low-quality independent restaurants.

Having demonstrated all the scenarios that could arise due to quality disclosure, we develop four testable hypotheses. Given that the various scenarios predicted by the model differ only in the entry probabilities of independent restaurants, in the first three hypotheses, we specifically test for the heterogeneous effects of quality disclosure on entry across different market types. For exits, because the predicted effects are common across all scenarios, we test whether the predicted effects on exit align with the model's predictions. Our hypotheses are as follows:

Hypothesis 1. In markets with low entry costs and a large difference between high and low qualities, the penetration of online review platforms can reduce the entry of all restaurants, including high-quality independent restaurants.

Hypothesis 2. In markets dominated by low-quality restaurants and a small difference between high and low qualities, the penetration of online review platforms can increase the entry of all restaurants, including low-quality independent restaurants.

Hypothesis 3. For markets that do not have the characteristics of those described in Hypotheses 1 and 2, the penetration of online review platforms can increase the entry of high-quality independent restaurants and discourage that of low-quality independent restaurants.

Hypothesis 4. The penetration of online review platforms discourages the exit of young high-quality independent restaurants and speeds up the exit of young low-quality independent restaurants.

5 Empirical Application

Our empirical context is the effect of online reviews platforms' penetration on firm entry and exit dynamics in the restaurant industry in Texas. The restaurant industry provides an ideal setting for testing our model predictions as it is often characterized by a large number of firms with small market shares. Online reviews disclose the quality of restaurants and allow consumers to learn from each other's experiences. In areas where consumers frequently use online reviews to share experiences, the quality disclosure of restaurants is likely fast, and the reverse would be true in areas where consumers do not share reviews often. Online review platforms, such as Yelp and Google, penetrated different regions in Texas during various periods of time. This variation in penetration rates across both regions and time allows us to tease out the effect of quality disclosure on industry dynamics.

To test Hypotheses 1 to 3 in Section 4 regarding the impact of quality disclosure on entry, we identify various market types that align with the settings in those hypotheses. We discuss these market types in detail in Subsection 5.1.2. To measure entry, we specify our dependent variable as the number of new entries instead of entry probabilities because we do not observe the number of potential entrants. We use the Poisson model to estimate entry and include regional fixed effects and time trends as controls for changes in the number of potential entrants.[14] For exits, recall that

[14] In Weintraub et al. (2008), firm entry is modeled using a Poisson distribution. In their model, the number of

Hypothesis 4 makes predictions about the impact of quality disclosure on exit, conditional on a firm's age. When we test this hypothesis, we model the exit probability directly because we observe the exit decisions of incumbent firms in each period. We split the restaurants based on their ages to test Hypothesis 4. We begin with a description of our data, the construction of various measures, and selections of market types in Subsection 5.1. Then we discuss our identification strategy and potential endogeneity in Subsection 5.2. Finally, we present and describe our empirical findings in Subsections 5.3 to 5.5.

5.1 Data, Measures and Market Types

The data are the same as those used in Fang (2022), which studies the effect of online review platforms on restaurant revenue. We use this data to investigate the effect of online review platforms' penetration on firm entry and exit dynamics. The data are drawn from a variety of sources: (1) Texas restaurant mixed beverage gross receipts data from the Texas Comptroller Office of Public Accounts. In addition to monthly revenue, this dataset tracks the entry and exit dates of each restaurant in Texas from January, 1995 to December, 2015. (2) Restaurant characteristics and review data collected from Yelp, TripAdvisor and Google, including snapshots of overall average ratings on each platform, prices in $ ranges and restaurant cuisine types. For restaurants listed on Yelp and TripAdvisor, we have the rating history data that show the time stamp and star rating for each review of each restaurant. (3) Consumer demographics data gathered from sources including the decennial census and the American Community Survey. Although the Texas mixed beverage gross receipts dataset contains all establishments that hold a liquor license, we restricted the sample to only full-service restaurants. In total, there are 15,417 restaurants in the dataset.[15]

5.1.1 Measures

We use these data to construct four measures that are relevant for our analysis: (1) online review platforms' penetration, (2) the quality of a restaurant, (3) the number of new entries, and (4) the action of exit. For market definition, we choose county as a geographic market because restaurant competition is relatively local. In total, there are 113 counties in our dataset.

Following Fang (2022), we use the average number of new reviews received by each restaurant per month on Yelp at the county level as the measure of online review platforms' penetration. This measure is constructed from the rating history data on Yelp. It represents consumers' review activity in each county each month.[16] Although only the review history data from Yelp is used

entries represents the entry probabilities. In our setting, by using the number of entries, we implicitly assume that online review platforms' penetration does not affect the number of potential entrants. In this regard, our Poisson regression model is consistent with Weintraub et al. (2008).

[15] For additional details of the dataset, please see Section 2 of Fang (2022).

[16] As described in Fang (2022), averaging the number of new reviews across restaurants makes the measure less sensitive to the size of a county, so that larger counties do not always have larger penetration measures. Furthermore, Berry and Waldfogel (2010) show that the number of restaurants in a region is generally proportional to the region's

in the construction of the measure, it can be seen as a measure capturing the penetration of all online review platforms because consumers' review activity on each platform is highly correlated. The penetration of online review platforms varies widely across different counties in Texas (see Figure 2). This variation reflects the availability of broadband internet in each county.[17] The rating history on Yelp shows that the first review on Yelp was written in March 2005. Because Yelp is the pioneer of online review platforms and has the longest time series, we use March 2005 as the start time of all online review platforms' penetration in our analysis.

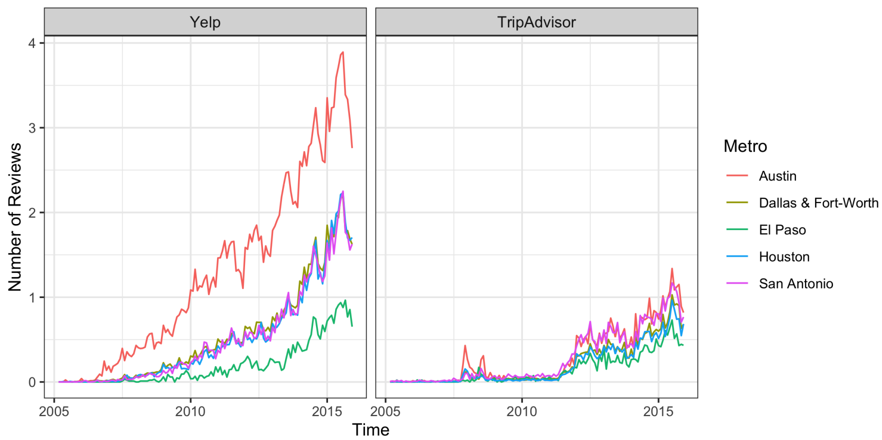


Figure 2: Yelp Penetration for Selected Metro Areas, 2005-2015 (Source: Fang (2022))

For the second measure, the quality of a restaurant, we leverage mainly the overall average rating on Google as of November 2016, the time of data collection. As demonstrated by Fang (2022), the Google rating covers the largest number of restaurants in the sample (9,0240 or about 60%) and is a good measure of quality. Among all restaurants operated after March 2005, the start of the online review platforms' penetration, over 68% of restaurants have Google ratings as of November 2016. By the end of the sample period (December 2015), almost all restaurants that were active had online presence (over 96%). For restaurants that do not have Google ratings, their quality is not available. The fact that some ratings are missing could be problematic if the selection of reviews is endogenous. For example, restaurants that exited long before the date of the data collection would not have ratings available.[18] Low-quality restaurants are more likely to exit than high-quality restaurants, and therefore, low-quality restaurants' ratings are more likely to be missing than those of the high-quality ones. In addition, the penetration of online review platforms might have sped up this process, leading to an endogenous data generating process.

population size. The average number of new reviews per restaurant therefore reflects the average review activity from each person.

[17]Based on the broadband internet coverage information provided by Connected Texas in 2014 (Texas, 2014), the correlation between the coverage of the fastest speed internet (6+ Mbps Download/1.5 Mbps Upload) and our measure of penetration is about 0.4.

[18]Google keeps the profiles of businesses that are recently closed, but if they are out of business for too long, their profiles are no longer accessible.

To address this missing data problem, we employ a random-forest machine-learning (ML) algorithm to predict the ratings for those restaurants with missing ratings. The ML prediction leverages the rich information contained in the dataset, particularly the revenue information. The rationale behind the prediction approach is that controlling for all observable restaurant characteristics and market trends, the change in restaurant revenue over time can inform us of the restaurant's quality. We expect there to be a monotonic relationship between revenue and ratings. For restaurants with ratings, we can estimate the relationship between revenue and ratings and use the revenues of restaurants without ratings to predict their ratings. The detailed implementation of the ML algorithm is explained in Online Appendix E. Our ML has a high out-of-sample fitting rate, with an R-squared of 0.924. In the main empirical analysis of the paper, we use the imputed rating data. As discussed in Online Appendix E, the imputed ratings and the original Google ratings have very similar distributions. To demonstrate that our results also hold using the original data with missing ratings, we conduct the same set of analyses on entry and exit by accounting for potential endogeneity. These analyses, presented in Online Appendix F.1, show that the results remain qualitatively the same.

For the third measure, the number of new entries, we count the number of newly entered restaurants by their chain affiliation and quality level in each county per month. An entry is defined as the first month that a restaurant makes positive revenue. In total, during our sample period, there were 10,309 new entries of independent restaurants and 2,500 of chain restaurants. To tally the number of new entries by quality level, the procedure is more involved. Recall that our hypotheses predict different effects for restaurants at different quality levels. Therefore, we need to model the number of restaurant entries across various quality levels. Given that Google ratings are continuous, we discretize quality into intervals to count the number of entries at each quality level. The discretization of Google ratings is based on the quantile of the rating distribution of all restaurants in the sample. For example, to discretize the ratings in the entire dataset into 3 intervals, we rank the ratings of all restaurants, and classify the highest 33.3% of the ratings into the high-quality bin. Ratings between 33.3% and 66.6% are in the medium-quality bin, and the remainder is in the low-quality bin.[19] Breaking the ratings into 3 intervals creates cutoff ratings for each bin as 4.1+ for high quality, between 3.7 and 4.1 for medium quality, and below 3.7 for low quality. The new entrants with ratings that fall into these cutoffs are classified into their corresponding bins. To verify the robustness of our findings to the number of bins, we run our analyses with 5 and 7 bins. For each market type, we do a separate discretization based on the rating distributions of restaurants in that market type.

For the fourth measure, the action of exit, we code an indicator variable that is 1 if a restaurant exits, and 0 if the restaurant stays in the market. The month of exit is informed by the out-of-

[19] We choose discretization based on quantile instead of equal distance in rating scale because the latter can generate disproportionally large bins. For example, most restaurants have a rating greater than 3.5 on Google. If we use equal distance in rating scale to create bins, then the cutoff ratings would be 1.67, 1.67 to 3.33, and above 3.33. In this case, almost all restaurants will be classified into high-quality bins, leaving very few in the other two bins.

Table 3: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	Number of Observations
Average # of entries for independent per county per month	0.44	1.35	0	19	23,288
Average # of entries for chain per county per month	0.11	0.43	0	10	23,288
Average # of exits for independent per county per month	0.32	1.05	0	15	23,288
Average # of exits for chain per county per month	0.048	0.32	0	19	23,288
Number of months in operation	74.6	71.3	1	265	15,417
Average Google rating for independent	3.88	0.33	1.5	5.0	12,089
Average Google rating for chain	3.65	0.42	1.9	4.8	3,328

business (OOB) date in the Texas mixed beverage dataset. The OOB date is the official date of business closure, at which point, the business no longer needs to report taxable earnings. However, there are some restaurants that begin to report zero revenues prior to the OOB date, suggesting that they exited prior to the official closure date. In those cases, we define the timing of exit as the month after which a restaurant's revenue is 0. In our sample, 8,594 restaurants exit, of which about 5,199 exit after March 2005, the start of online platforms' penetration.[20]

Table 3 provides the key summary statistics of our data. In terms of entry, the number of entries in each market varies from 0 to 19 with a mean of 0.44 for independent restaurants, and ranges from 0 to 10 with a mean of 0.11 for chain restaurants. In terms of exit, Table 3 shows that the number of exits in each market ranges from 0 to 15 with a mean of 0.32 for independent restaurants, and extends from 0 to 19 with a mean of 0.048 for chain restaurants. The data show significantly more exits for independent restaurants than chain restaurants during the sample period. In fact, of the 8,594 restaurants that exit, only 1,125 were chains. The average lifespan of a restaurant is about 75 months or 6 years. With regard to ratings, the average rating for independent restaurants is 3.88, much higher than the average of 3.65 for chain restaurants.

5.1.2 Market Types

We break markets in our sample into three different types: college-town markets, highway-exit markets, and other markets. College-town markets are characterized by lower entry costs *and* a large difference between high and low qualities. We use this type of market to test Hypothesis 1. We define college-town markets as counties that include restaurants with two miles of a college outside

[20]For both entry and exit definitions, we adjust for regular or major renovation. If some restaurants temporarily withdrew from the market and then reappeared after less than two months, then we treat this as regular renovation and do not count the withdrawal as exit nor the reappearance as a new entry. However, if a restaurant disappeared from the market for longer than two months, then we treat this as a major renovation that likely changes the restaurant substantially. We count this incidence in both exit and new entry. We do not, however, change the Google rating before and after the major renovation due to lack of data. Given that there are very few major renovation incidents (less than 10), this treatment of the rating should not cause significant bias in the empirical testing.

of large metropolitan areas (i.e. Austin, Houston, Dallas, Fortworth and San Antonio).[21] If some restaurants also fall within 500 meters of the interstate highway exits, we exclude them from this market type. College towns are often suburban areas with universities and colleges as the biggest employers. We believe it is reasonable to assume that the entry costs for restaurants in college towns are usually lower than those in large urban areas. Although we do not have measures of entry cost, such as expenses for hiring construction workers, completing legal paperwork to establish a business, and so on, residential rent often serves as a reliable indicator of demand, economic conditions, and costs of doing business. Areas with higher residential rents signal increased demand for both living and commercial spaces, which in turn raises construction and renovation costs due to competition for space and higher material and labor costs. According to DeVon (2024) and Reuther (2024), the average rent for a two-bedroom apartment in large cities like Austin was about $2,800 per month in 2023, whereas the rent in College Station, where Texas A&M University is located, was only $772. Because students, faculties and staff from the universities are a constant flow of regular customers with a wide range of needs (e.g. quick lunches and professional dinners), the restaurants in these areas are likely to vary widely in quality. This feature of a relatively large quality range is supported by the data. As shown in Table 4, the ratings in college-town markets extend from 1.9 to 5, covering a wider range than that in the highway-exit markets, although smaller than other markets. In fact, the range of ratings in college-town markets being smaller than that in other markets is not an issue for testing Hypothesis 1. The reason is that Hypothesis 1 requires *both* a large quality difference and a low entry cost. Even though the range of ratings in other markets is large, those markets where most of the restaurants in our sample operate are in large urban cities, which likely also have high entry costs. Therefore, the conditions stipulated by Hypothesis 1 are unlikely to be met in other markets. In total, we have 5,242 unique restaurants in college-town markets.

The second type of market, highway-exit markets, are dominated by low-quality restaurants and a relatively small difference between high and low qualities. We use this type of market to test Hypothesis 2. Specifically, following Fang (2022), we define areas within 500 metres of interstate highway exits as highway-exit markets, but we exclude those markets that go through large city downtowns, such as Austin downtown.[22] The rationale behind our choice is that restaurants around highway exits outside large urban downtowns tend to serve mostly travellers, who do not try out restaurants repeatedly. As a result, lower-quality restaurants may achieve higher profits in these areas, leading to their predominance in such markets.[23] The features of low-quality dominance

[21] More specifically, the counties in these large metropolitan areas include Bexar, Dallas, Harris, Tarrant, and Travis.

[22] We exclude downtowns in Austin, Houston, Dallas, Fort-worth and San Antonio. The boundaries of city downtowns are based on each city's planning departments' definitions.

[23] This pattern has been documented in prior research, such as Mazzeo (2004) and Blair and Lafontaine (2005), which find that the quality of restaurants or hotels in locations with transient consumers tends to be lower than that in locations with repeat consumers.

Table 4: Google Rating Distribution by Market Type

Variable	Min	Median	Mean	Max.	Number of Restaurants	Number of Observations
All markets	1.50	3.81	3.76	5.00	15,394	23,288
College-town markets	1.90	3.83	3.79	5.00	5,242	13,204
Highway-exit markets	2.00	3.75	3.70	4.90	1,524	8,483
Others	1.50	3.84	3.81	4.90	8,628	9,615

Note: Rating information is based on imputed data. An observation is a county-month pair.

and a small quality range are also supported by the data. As shown in Table 4, the highway-exit markets have a lower mean and median (3.70 and 3.75 respectively) compared to all markets in Texas (3.76 and 3.81 respectively). The quality range is also smaller, 2 to 4.9 v.s. 1.5 to 5.0 in all markets. We aggregate the highway-exit markets to the county-month level in our empirical analysis to account for the fact that some highway exits in the same county are fairly close to each other and to reduce the noise in entry and exit measures.

All other markets outside highway-exit and college-town markets are labeled as the "others" market type.[24] They are counties covering a much larger geographic space, including the large urban counties. We use this market type to test Hypothesis 3, which is the intuitive scenario. There are 8,628 unique restaurants in the "others" markets, the highest among all market types. Because most of these restaurants are in large urban cities, their entry costs are likely to be higher than those in the other two types of markets. Furthermore, as displayed in Table 4, the average quality in this type of market, 3.81, is higher than those in the other two types. The quality range is also the widest, extending from 1.5 to 5.0. As shown by our numerical examples in Figure 1a, when the entry cost is relatively high and the quality range large, the intuitive scenario described in Hypothesis 3 is likely to occur. Therefore, this market type provides a good setting for testing the intuitive scenario.

5.2 Identification Strategy

The identification strategy we employ to test the model predictions is a triple difference-in-difference (DiD) approach. We exploit the variation in the penetration of online review platforms across geographic regions and time to tease out their effects on the entry and exit of restaurants. Because online reviews disclose quality of restaurants, for a given quality, the change (increase or decrease) in new entries or exits should be larger in regions with greater amount of consumer review activity than in those with a lower amount, and the change should vary with restaurant quality, chain affiliation, and age. Below we discuss in detail the econometric specifications to examine new entries and exits.

[24]For counties that contain highway-exit markets, we treat the areas outside of the highway-exit markets as their own markets. Therefore, one county can appear in both highway-exit and "others" markets, but the restaurants in the two types of markets are mutually exclusive.

Entry To test the effect on new entries, we use a Poisson regression. Poisson DiD regressions yield consistent estimates as long as the conditional mean of the dependent variable is correctly specified (Wooldridge, 2023). The main econometric specification is as follows:

$$\begin{aligned}\log(E(N_{qfct}|\mathbf{X}, \boldsymbol{\theta}, \ldots)) =&(\theta_y + \theta_{yr} Rating_{qfct}) \log(Yelp_{ct})(1 - D_f^{ch}) + (\theta_y^{ch} + \theta_{yr}^{ch} Rating_{qfct}) \log(Yelp_{ct}) D_f^{ch} \\ &+ \mathbf{X_{ct}} \boldsymbol{\theta_x} + \theta_{qfc} + \theta_{Mt} + \theta_{Mt}^{ch} + \varepsilon_{qfct}, \end{aligned} \tag{4}$$

where the subscript ct presents county c at time t. N_{qfct} is the number of new entries by quality bin q and chain affiliation f in county c at time t. $Rating_{qfct}$ is the numerical average of Google ratings of all new entries in the quality bin q and chain affiliation f. As we described earlier, for the bulk of our analyses we use 3 quality bins: high, medium and low. The notation q indicates which bin a new restaurant falls into, and $Rating_{qfct}$ is the group average of Google ratings for all new entries in bin q.[25] The variable $\log(Yelp_{ct})$ is the natural logarithm of the penetration measure in county c at time t. $\theta_y, \theta_{yr}, \theta_y^{ch}$, and θ_{yr}^{ch} are the coefficients related to $\log(Yelp_{ct})$.[26] D_f^{ch} is an indicator for chain affiliation. If f indicates chain, then $D_f^{ch} = 1$; otherwise, 0. $\mathbf{X_{ct}}$ is county-level market characteristics such as population, age, race, and income. The variable θ_{qfc} is the fixed effect of new entries with quality bin q, chain affiliation f, and county c. This fixed effect is to capture that restaurants with certain quality and chain affiliation are more likely to enter in some counties than in others. θ_{Mt} is a metro-region-time fixed effects, and θ_{Mt}^{ch} is metro-region-time fixed effects for chains. A metro region is a core-based statistical area (CBSA) as defined by the US Census Bureau. The fixed effects θ_{Mt} and θ_{Mt}^{ch} are to control for the varying time trends in restaurant entry across different regions in Texas.

In this regression, the key parameters of interest are $\theta_y, \theta_{yr}, \theta_y^{ch}$, and θ_{yr}^{ch}. In particular, the composite coefficient $\theta_y + \theta_{yr} Rating_{qfct}$ represents the impact of online review platforms on the entries of independent restaurants with rating $Rating_{qfct}$, and $\theta_y^{ch} + \theta_{yr}^{ch} Rating_{qfct}$ represents that for chain restaurants. They can be interpreted as the percentage change in the expected number of entries when the activity on online review platforms increases by 100%.

The sample for this regression includes all entries after March 2005, the start time of online review platforms' penetration. New entries before this date should not be affected by online reviews. We first estimate regression as specified in Equation (4) by pooling all markets (county-month level) together and assuming that the effects of online review platforms in all markets share the same direction. This helps us gauge the overall directions of the effects in majority of the markets. Then, we conduct the analysis by market type by interacting $\log(Yelp_{ct})$ with market-type indicators. The

[25] Here we use the numerical rating to represent quality instead of quality-bin indicators because the coefficient associated with the numerical rating offers greater flexibility in interpretation. As discussed in the Results section, we calculate the effect of review platform penetration on restaurant entry for each Google star rating (i.e, 1-star, 2-star, and so on). It is more straightforward to compute these effects using the coefficient for the numerical rating than the coefficients for quality bin indicators. In addition, using quality-bin indicators increases the number of coefficients to estimate, especially when the number of bins rises to 5 or 7, leading to potential issues with statistical power.

[26] The subscript y indicates Yelp, and r represents rating.

estimated interaction effects will measure how the impact of online review platforms on entry varies across different market types. These interaction terms will be used to test Hypotheses 1 and 2, which specify how the impact of quality disclosure may lead to counterintintuive scenarios in some of the market types. It should be noted that these interactions are exhaustive, and we do not have a baseline.

The identifying assumptions for this DiD approach include (1) the treatment — online review platforms' penetration — is exogenous to the change in the outcome, and (2) the new entries in the treated markets should follow the same trend as those in the non-treated markets in the absence of the treatment. With these two assumptions, the effect of online review platforms estimated from the regression can be interpreted as causal, conditional on our controls, which include observed market characteristics and various sets of fixed effects. As discussed in Subsection 5.1, a significant fraction of variation in online review platform penetration likely arises from the availability of broadband internet. One might still argue that our penetration measure may be correlated with county-time specific demand shocks; for example, county fairs may boost demand, increase new entry and generate a lot of reviews at the same time. We deal with this potential endogeneity by controlling for observable market characteristics as well as a rich set of fixed effects. For market characteristics, we control for the demographic information at the county level, such as population, income, age, and race. For fixed effects, we use metro-region-time fixed effects to account for unobserved county-time specific demand shocks. Given that one county's demand shocks are likely to spill over to an entire metro region, using metro-region-time fixed effects is appropriate. Additionally, we allow the metro-region-time fixed effect to vary by chain affiliation. To capture idiosyncratic shocks in each county for a specific quality category and chain affiliation, we include the quality-bin-chain-affiliation-county fixed effect.

Furthermore, to check the assumption of parallel trend, we conduct a placebo test where we moved the penetration of online review platforms to 1995, ten years prior to the actual start of penetration. If there are inherent differences in the trends across markets, which would lead to a spurious relationship between the change in entry and exit of restaurants and the penetration of online review platforms, the effect of penetration should also be significant in the placebo test. None of the coefficients associated with the penetration measure are significant, providing support for the parallel trend assumption (Online Appendix Table F.7).

Exit To examine the effect of online review platforms' penetration on exit, we estimate a linear probability regression:

$$\begin{aligned} a_{jt} =& (\theta_y + \theta_{yr} Rating_j) \log(Yelp_{ct})(1 - D_j^{ch}) + (\theta_y^{ch} + \theta_{yr}^{ch} Rating_j) \log(Yelp_{ct}) D_j^{ch} \\ &+ \mathbf{X_{jt}} \boldsymbol{\theta_x} + \theta_{jmnh} + \theta_{Mt} + \theta_{Mt}^{ch} + \varepsilon_{jt}, \end{aligned} \tag{5}$$

where a_{jt} is an indicator for the exit action, which is 1 if restaurant j decides to exit at time t, and 0 otherwise. $D_j^{ch} \equiv \mathbb{1}\{\text{restaurant } j \text{ is affiliated with a chain.}\}$ is an indicator for chain affiliation.

$\mathbf{X_{jt}}$ includes restaurant characteristics as well as county-level demand and cost conditions at time t. θ_{jmnh} is the restaurant-month fixed effect, where mnh denotes the calendar months from January to December. It controls for restaurant-specific seasonality because different types of restaurants may be affected by seasonality to varying degrees, and they may decide to exit in particular months of the year. The rest of the variables are defined the same way as those in Equation (4).

The identifying assumption in this test is the same as that for the entry. In particular, the identification comes from comparing the changes in exit behaviors of restaurants in high-penetration areas to those in low-penetration areas. As mentioned previously, the placebo test on exit supports the parallel trend assumption (Online Appendix Table F.8). The sample for the exit regression is limited to the periods when a restaurant received treatment from online reviews; that is, the periods after a restaurant had received its first online review on Yelp, TripAdvisor or Google.[27]

We use this regression to test Hypothesis 4. Because Hypothesis 4 emphasizes the effect on young restaurants, we interact $\log(Yelp_{ct})$ in Equation (5) with an age indicator for a restaurant being young. If a restaurant is less than 12 years old, we label that as young; otherwise, old. We omit this indicator from Equation (5) for cleaner presentation. We choose this age cutoff based on the findings from Fang (2022), who shows that the effect of online review platform penetration on restaurant revenue diminishes with restaurant age and disappears after a restaurant reaches 12 years old.[28] We also use different age cutoffs of 6 and 10 years old as robustness checks, finding similar results. For this test on exit, we pool all market types together instead of separating them as in the entry analysis because all markets exhibit the same pattern in effects. All the empirical results on entry and exit are shown in Sections 5.3 and 5.4, respectively.

5.3 Results: Entry

For the results on entry, we present two sets of results: one is from pooling all markets together, and the other is by market type. These results confirm Hypotheses 1 to 3.

5.3.1 Pooling All Markets

Table 5 presents the results of our first set of regressions (Equation (4)), where the number of new entries is the dependent variable. The first column shows the estimated coefficients related to the penetration of online review platforms when we use 3 quality bins. The second and third columns present estimates using 5 or 7 quality bins, respectively. As shown in the table, the

[27]For Yelp, we use the exact date since we have access to the historical review data for each restaurant listed on the platform. For TripAdvisor and Google, however, we lack historical review data, so we assume January 2011 as the date of the first review, following the methodology used in Fang (2022). If a restaurant opened after this date, we assume that reviews began as soon as it opened. If a restaurant is listed on multiple platforms, the earliest review date across platforms is used as the first review date.

[28]Our model simulates the aggregate average of the learning behavior of the entire population. The empirical evidence from Fang (2022) shows that it takes the entire population 12 years to completely learn the true quality of a restaurant.

coefficients for chains (on the 3rd and 4th rows) are almost all statistically insignificant, indicating that online review platform penetration had very little effect on the entry of chain restaurants. In the first two rows of Table 5, we present the estimated main and quality interaction coefficients of review platform penetration for independent restaurants. The coefficients for independent are all significant at the 1% level regardless of which number of quality bins we use. The coefficient for log(Yelp) in the table represents the percentage change in the expected number of entries in a market when rating is 0 as consumers' review activity increases by 100%. It being negative implies that for very low quality levels, the number of new entries declines with the penetration of online review platforms. The positive and significant coefficient for log(Yelp)×rating in the table implies that the effect of review platform penetration increases with quality. In particular, the magnitudes of the main and interaction effects imply that for very high quality levels, the overall effect from review platform penetration on entry would be positive.

To illustrate how the impact of quality disclosure differs across restaurant quality, in Table 6 we present the composite coefficient $\theta_y + \theta_{yr} Rating$ for four different Google star ratings. We use the estimates from the regression specification with 3 quality bins when constructing the estimates in Table 6. As shown, for restaurants with 2- and 3-star ratings, the numbers of new entries decline by 28% and 12.1%, respectively, as consumers' review activity doubles. For restaurants with a 5-star rating, the effect is the opposite, with new entries increasing by 19.6%. For restaurants with a 4-star rating, which is close to the average quality of all independent restaurants, the effects are small and insignificant. The results presented in Table 6 indicate that high-quality restaurants' entry is encouraged by quality disclosure from online reviews, whereas low-quality restaurants' entry is discouraged. In summary, our findings indicate that most markets seem to be of the type where Hypothesis 3 arises.

5.3.2 Across Different Market Types

For the analysis by market type, we use 3 quality bins instead of 5 or 7 bins as using a larger number of bins would segment the data too finely, introducing excessive noise in the measurement of entry.[29] Table 7 presents the estimated regression coefficients from an expanded version of Equation (4), where we include interactions between the coefficients on online review penetration, chain affiliation, and market type. Again, all the interactions with the chain affiliation and the market type are exhaustive without a baseline. The first three columns present the interaction coefficients for independent restaurants, and the fourth through sixth columns present those for chains. As shown, most coefficients for independent restaurants are statistically significant at the 5% level, but almost all coefficients for chains are insignificant, except for one coefficient under

[29] For example, as shown in Table 4, highway-exit markets contain a total of 1,524 restaurants, with only 890 entries spreading across 8,483 county-month pairs after Yelp penetration. Dividing the number of entries per county and month into three quality bins already results in many observations with just 0 or 1 entry. Further breaking down the data into 5 or 7 bins would overly fragment the data, introducing excessive noise.

Table 5: Effect on Entry by Chain Affiliation

	(1)	(2)	(3)
	Entry	Entry	Entry
log(Yelp) (independent)	-0.597***	-0.580***	-0.571***
	(0.106)	(0.108)	(0.0896)
log(Yelp)×rating (independent)	0.159***	0.156***	0.144***
	(0.0249)	(0.0264)	(0.0225)
log(Yelp) (chain)	0.0553	0.116	0.0417
	(0.0779)	(0.0818)	(0.0542)
log(Yelp)×rating (chain)	0.0191	-0.00489	0.0212**
	(0.0163)	(0.0189)	(0.00998)
Controls	✓	✓	✓
Group FE	✓	✓	✓
Year× Month× Metro× Chain FE	✓	✓	✓
N	3,798	4,504	4,882
N of Clusters	49	49	48
Number of Quality Levels	3	5	7

Controls include population, income, age, race, and population density. Group FE is quality-bin×chain-affiliation×county fixed effect. All standard errors are clustered at the county level. They are shown in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

the college-town market type (0.176), which is only marginally significant with a p-value of 0.091. Taken together, the estimates in the last three columns of Table 7 indicate that the effects of online review platforms' penetration on chain restaurants' entry are not significant in all types of markets. Therefore, we focus our discussion of results on independent restaurants.

We present the marginal impact of online review platforms' penetration by Google star rating in Table 8, where the marginal impact is computed from the regression coefficients in Table 7, using a similar approach as that taken when computing the effects in Table 6. Below, we discuss the results for each market type individually.

College-town Markets As shown in Table 7, for independent restaurants in college-town markets, the estimated coefficient for log(Yelp) is negative and significant, while the estimated coefficient for log(Yelp)×rating is positive but insignificant. The fact that the main effect is strongly negative and dominates the positive coefficient of the interaction implies that independent restaurants from all quality levels are potentially discouraged from entry. To see this, notice that the

Table 6: Effect on Entry by Google Star Rating (Independent)

Star Rating	2	3	4	5
Effect	-0.2800***	-0.1214***	0.0371	0.1956***
	(0.0591)	(0.0392)	(0.0287)	(0.0368)

All standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 7: Effect on Entry by Chain Affiliation and Market Type

Dependent variable	Entry					
	Independent			Chain		
	College-town	Highway-exit	Others	College-town	Highway-exit	Others
log(Yelp)	-0.287**	-0.179	-0.541***	0.176*	0.197	-0.203
	(0.140)	(0.111)	(0.0921)	(0.104)	(0.160)	(0.317)
log(Yelp)×rating	0.0591	0.109***	0.157***	0.00524	-0.00678	0.0480
	(0.0384)	(0.0260)	(0.0176)	(0.0268)	(0.0488)	(0.0708)
Controls	✓					
Group FE	✓					
Year×Month×Metro×Chain × Market-Type FE	✓					
N	3349					
N of Clusters	47					

Group FE is quality-bin× chain-affiliation× county× market-type fixed effect. Controls include population, income, age, race, and population density. All standard errors are clustered at the county level and are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

composite coefficient, $\theta_y + \theta_{yr} Rating$, only becomes positive once a restaurant's rating exceeds 4.85, which is significantly higher than the mean rating of 3.79 in college-town markets (Table 4). As a result, even high-quality restaurants are discouraged from entry in college-town markets. The first row of Table 8, which shows the marginal impact of quality disclosure for restaurants in college-town markets by Google star rating, also illustrates the previous point. Independent restaurants with 2- or 3-star ratings are 16% and 11% less likely to enter if consumers online review activities increase by 100%, and these effects are statistically significant. The impact of quality disclosure on restaurants with 4-star ratings is negative, although insignificant. For those with 5-star ratings, the impact of quality disclosure on entry is statistically insignificant and close to zero. In summary, the analysis of college-town markets supports Hypothesis 1.

Highway-Exit Markets Column 2 in Table 7 presents the effect of online review platforms' penetration on the entry of independent restaurants in the highway-exit markets. The coefficient for log(Yelp) is statistically insignificant, but the coefficient for log(Yelp)×rating in the table is positive and statistically significant. Both coefficients have the same respective signs as those for the college-town markets, implying that the impact of online review platforms' penetration is negative for very low-quality restaurants but positive for higher-quality restaurants. In particular, for highway-exit markets, the predicted impact of quality disclosure on entry is negative for restaurants with ratings

Table 8: Effect on Entry by Google Star Rating and Market Type (Independent)

Star Rating	2	3	4	5
College-town	-0.1683**	-0.1092**	-0.0501	0.0090
	(0.0692)	(0.0421)	(0.0413)	(0.0678)
Highway-exit	0.0400	0.1494**	0.2587***	0.3681***
	(0.0743)	(0.0639)	(0.0633)	(0.0726)
Others	-0.2260***	-0.0686	0.0889**	0.2463***
	(0.0627)	(0.0508)	(0.0430)	(0.0417)

All standard errors are shown in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

below 1.64, and positive for those with ratings above. Notice that the cutoff rating for where the effect becomes positive is very low relative to the average quality rating in this market, which is 3.7 (Table 4). As a result, our estimates imply that quality disclosure increases entry for all independent restaurants in highway-exit markets, even those with low quality. This finding is further supported by the second row of Table 8. As can be seen, the predicted marginal impacts of quality disclosures are large, positive and statistically significant for restaurants with ratings of 3 or more stars. In particular, the entries of restaurants with 3-, 4- and 5-star ratings will increase by 15%, 26% and 37%, respectively, as consumers' review activity increases by 100%. For restaurants with 2-star ratings, the estimated impact is statistically insignificant, but still positive in sign. In summary, our analysis of highway-exit markets supports Hypothesis 2.

Other Markets Column 3 in Table 7 displays the estimated main and interaction effects of quality disclosure for other markets. The estimated coefficients are very similar in sign and magnitude to those presented in Table 5, where all markets are pooled together. This similarity is likely due to the fact that this type of market contains the highest number of restaurants compared to the other two types (see Table 4). Correspondingly, the composite coefficients presented in the third line of Table 8 show that high-quality independent restaurants are encouraged to enter the market, and low-quality restaurants are discouraged. Again, as with the analyses of the entire market, the estimated results from other markets support Hypothesis 3, and we expect this because most restaurants are in the markets of the "others" type.

5.4 Results: Exit

We present the estimated coefficients of interest from Equation (5) in Table 9. Similar to the results on entry, the coefficients for chains are all statistically insignificant. For independent restaurants, the coefficients associated with established restaurants are all insignificant as well. Because the quality of established restaurants is known to consumers even without online reviews, these results are in line with our prior expectations. In contrast, the coefficients for young independent restaurants are all statistically significant at the 1% level. In particular, the coefficient for log(Yelp) in the table being positive indicates that when a restaurant is very low-quality, the penetration of online review platforms will speed up its exit. The coefficient for log(Yelp)×rating being negative implies that as quality increases, the positive impact of review platform penetration on exit diminishes. The magnitudes of the coefficients show that for restaurants with sufficiently high quality, the impact of review platform penetration on exit will in fact be negative, meaning that the penetration of online review platforms in fact encourages high-quality restaurants to stay in the market for longer.

To show how the impact of review platform penetration on exit varies with Google star rating, we present the composite coefficients for young independent restaurants by Google star rating in Table 10. The estimated effects imply that when consumers' review activity doubles, the exit rates with 2- and 3-star ratings will increase by 0.66% and 0.32%, respectively, and these effects are statistically significant. In contrast to the effects on low-quality restaurants, for a 5-star young independent restaurant, the exit rate will decrease by 0.37%. For medium-quality restaurants with 4-star ratings, the effects are very small and statistically insignificant. In addition to using 12 years as the age cutoff for young and established restaurants, we also conducted the analysis using alternative cutoffs of 6 and 10 years for robustness checks. The findings are consistent with those obtained using the 12-year cutoff. Overall, our findings support Hypothesis 4: the penetration of online review platforms discourages the exit of high-quality young independent restaurants, but speeds up the exit of low-quality ones.

Table 9: Effect on Exit by Chain Affiliation and Age

Dependent variable	Exit			
	Independent		Chain	
	Established	Young	Established	Young
log(Yelp)	-0.000470	0.0134***	-0.00168	0.000234
	(0.00314)	(0.00415)	(0.00110)	(0.00102)
log(Yelp)×rating	-0.000468	-0.00343***	0.000201	-0.000214
	(0.000790)	(0.00103)	(0.000344)	(0.000272)
Controls	✓			
Restaurant×Month FE	✓			
Year×Month×Metro×Chain FE	✓			
N	408381			
N of Clusters	98			

Established restaurants are those that are older than 12 years. Controls include population, income, age, race, and population density. All standard errors are clustered at the county level and are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 10: Effect on Exit by Google Star Rating (Young Independent)

Star Rating	2	3	4	5
Effect	0.0066***	0.0032**	-0.0003	-0.0037***
	(0.0022)	(0.0013)	(0.0008)	(0.0014)

All standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

5.5 Implications on Quality Distribution

Up to this point, our analysis has focused on entry and exit rates, but not the quality distribution of incumbent firms. Since consumer welfare is also impacted by the quality of firms operating in the market, we analyze that outcome in this section. The results from the entry and exit analyses can be translated into market dynamics on quality distribution. In particular, given that higher-quality firms are encouraged to enter and exit less frequently in most markets,[30] the overall average quality of the stock of restaurants in the market should improve as a result of the penetration of online review platforms.

Specifically, we estimate a similar regression as Equation (4), with different dependent variables: the first is the mean rating of restaurants within a county in each month; the second is the percentage of high-quality restaurants within a county in each month.[31] The results are presented in Table F.6 in the Online Appendix F.2. Our results suggest that the penetration of online review platforms had little impact on the average quality of the stock of chain restaurants in the market. This finding is consistent with the results from the entry and exit regressions. In contrast, we find positive and significant effect of online review platform on the overall quality of the stock of independent restaurants. In particular, mean rating of the stock of independent restaurants in a market will increase by 0.012 when consumers' online review activity doubles. Moreover, the share of high-quality restaurants will increase by almost two percentage points when consumers' review activity increases by 100%. This represents a substantial improvement in quality. For example, in Harris county, our estimates suggest that the share of high-quality independent restaurants increased by 8 percentage points from the start of review platform penetration to the end of the sample period. These results show that by affecting the entry and exit decisions of firms, online review platforms have a significant impact on the improvement of the quality of incumbents in the market.

[30]Note that even in highway-exit and college-town markets, high-quality firms are always more encouraged to enter the market *relative* to low-quality firms due to quality disclosure.

[31]We define high-quality restaurants as the highest quality bin in the 3-bin quality classification as shown in Table 5.

6 Conclusion

This paper examines how quality disclosure affects the dynamics of firm entry and exit in markets with many small firms. We develop a novel theoretical model that highlights two key forces through which quality disclosure drives market dynamics: (1) the direct effect, which pertains to changes in consumers' preference for a firm upon learning its quality, and (2) the competition effect, which relates to changes in a firm's competitive environment once consumers learn its competitors' qualities. Depending on which force dominates, several scenarios can emerge. In some instances, quality disclosure intensifies competition to the extent that high-quality firms are discouraged from entry, while in some other cases, it reduces competition so significantly that even low-quality firms are encouraged to enter.

We empirically test our model predictions using a unique dataset that tracks the entry and exit of restaurants and consumers' online review activities in Texas from 1995 to 2015. Our results show that increased quality disclosure in different market types in Texas gives rise to different scenarios. In college-town markets, which are characterized by low entry costs and large variations in restaurant quality, the penetration of online review platforms discourages the entry of both high- and low-quality firms. Conversely, in highway-exit markets, which are dominated by low-quality firms, the penetration of online review platforms encourages the entry from all firms, even those of low quality. In other markets, the penetration of online review platforms encourages the entry of high-quality firms while deterring low-quality ones. Regarding exits, our results show that quality disclosure helps young high-quality independent firms stay in the market longer, while accelerates the exit of young low-quality ones. No significant impact is found for chain or established independent restaurants. Overall, these findings suggest that the penetration of online review platforms contributes to the overall improvement in restaurant quality in the market.

Our results shed light on quality disclosure's broader implications, beyond matching consumers with products. From a policy perspective, our paper highlights the complex nature of quality disclosure. While it generally enhances the average quality of firms in the market, it can also discourage the entry of high-quality restaurants in certain market segments, particularly those with lower entry costs and large quality difference. Policymakers need to consider the broader economic and social impacts when designing and implementing quality disclosure regulations. Our paper also guides businesses to navigate the challenges of information availability and increased competition. For instance, high-quality firms should be aware of the heightened competition in markets where quality information is readily accessible to consumers. Conversely, lower-quality firms might find opportunities in markets where the competitive pressures from quality disclosure are reduced.

References

Anderson, M. and Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, 122(563):957–989.

Anenberg, E., Kuang, C., and Kung, E. (2019). Social learning and local consumption amenities: Evidence from yelp. *The Journal of Industrial Economics*.

Benassy, J.-P. (1996). Taste for variety and optimum production patterns in monopolistic competition. *Economics Letters*, 52(1):41–47.

Berry, S. and Waldfogel, J. (2010). Product quality and market size. *The Journal of Industrial Economics*, 58(1):1–31.

Blair, R. D. and Lafontaine, F. (2005). *The economics of franchising*. Cambridge University Press.

Chen, Y. and Xu, D. Y. (2023). A structural empirical model of R&D, firm heterogeneity, and industry evolution. *The Journal of Industrial Economics*, 71(2):323–353.

Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3):345–354.

Chintagunta, P. K., Gopinath, S., and Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, 29(5):944–957.

Dafny, L. and Dranove, D. (2008). Do report cards tell consumers anything they don't already know? the case of medicare hmos. *The RAND Journal of Economics*, 39:790–821.

Dendorfer, F. and Seibel, R. (2024). The cost of the "cold-start" problem on airbnb. *Working Paper on SSRN*.

DeVon, C. (2024). The 10 most affordable college towns in the u.s.-half are in texas. *CNBC*.

Dhar, V. and Chang, E. A. (2009). Does chatter matter? the impact of user-generated content on music sales. *Journal of Interactive marketing*, 23(4):300–307.

Donati, D. (2022). The end of tourist traps: A natural experiment on the impact of tripadvisor on quality upgrading. *CESifo Working Paper*.

Dranove, D. and Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4):935–963.

Dranove, D. and Sfekas, A. (2008). Start spreading the news: A structural estimate of the effects of new york hospital report cards. *Journal of Health Economics*, 27:1201–7.

Ericson, R. and Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, 62(1):53–82.

Fang, L. (2022). The effects of online review platforms on restaurant revenue, consumer learning, and welfare. *Management Science*, 68(11):8116–8143.

Farronato, C. and Zervas, G. (2022). Consumer reviews and regulation: Evidence from nyc restaurants. Technical report, National Bureau of Economic Research.

Forman, C., Ghose, A., and Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information systems research*, 19(3):291–313.

Godes, D. and Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science*, 23(4):545–560.

Guo, L. and Zhao, Y. (2009). Voluntary quality disclosure and market interaction. *Marketing Science*, 28(3):488–501.

Hollenbeck, B. (2018). Online reputation mechanisms and the decreasing value of chain affiliation. *Journal of Marketing Research*, 55(5):636–654.

Hopenhayn, H. and Saeedi, M. (2023). Optimal information disclosure and market outcomes. *Theoretical Economics*, 18(4):1317–1344.

Hotz, V. J. and Xiao, M. (2013). Strategic information disclosure: The case of multi-attribute products with heterogeneous consumers. *Economic Inquiry*, 51(1):865–881.

Hui, X., Saeedi, M., Spagnolo, G., and Tadelis, S. (2023). Raising the bar: Certification thresholds and market outcomes. *American Economic Journal: Microeconomics*, 15(2):599–626.

Jacobs, J. (1962). *The Death and Life of Great American Cities*. Random House, New York.

Jin, G. Z. and Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451.

Jin, G. Z. and Sorensen, A. T. (2006). Information and consumer choice: The value of publicized health plan ratings. *Journal of Health Economics*, 25(2):248–275.

Li, X. and Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information systems research*, 19(4):456–474.

Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp.com. *Harvard Business School NOM Unit Working Paper*, (12-016).

Mazzeo, M. J. (2004). Retail contracting and organizational form: Alternatives to chain affiliation in the motel industry. *Journal of Economics & Management Strategy*, 13(4):599–615.

Moe, W. W. and Trusov, M. (2011). The value of social dynamics in online product ratings forums. *Journal of marketing research*, 48(3):444–456.

Newberry, P. and Zhou, X. (2019). Heterogeneous effects of online reputation for local and national retailers. *International Economic Review*, 60(4):1565–1587.

Porter, M. (2015). The competitive advantage of the inner city. In *The city reader*, pages 358–371. Routledge.

Reuther, K. (2024). The best texas college towns for fun & affordability. *TurboTenant*.

Saeedi, M. (2019). Reputation and adverse selection: Theory and evidence from ebay. *The RAND Journal of Economics*, 50(4):822–853.

Texas, C. (2014). The broadband landscape in the state of texas: Assessment at a state, regional & local level, and recommendations for broadband expansion. *Connected Texas Reports*.

Tripadvisor.com (2018). Influences on diner decision-making. Technical report, Tripadvisor.com.

van Gelder, K. (2023). How many reviews do you typically read before you make a decision to purchase? *Statista.com*.

Vellodi, N. (2018). Ratings design and barriers to entry. *Available at SSRN 3267061*.

Waxman, L. (2006). The coffee shop: Social and physical factors influencing place attachment. *Journal of Interior Design*, 31(3):35–53.

Weintraub, G. Y., Benkard, C. L., and Van Roy, B. (2008). Markov perfect industry dynamics with many firms. *Econometrica*, 76(6):1375–1411.

Wooldridge, J. M. (2023). Simple approaches to nonlinear difference-in-differences with panel data. *The Econometrics Journal*, 26(3):C31–C66.

Wu, C., Che, H., Chan, T. Y., and Lu, X. (2015). The economic value of online reviews. *Marketing Science*, 34(5):739–754.

Zhao, Y., Yang, S., Narayan, V., and Zhao, Y. (2013). Modeling consumer learning from online product reviews. *Marketing Science*, 32(1):153–169.

Zhu, F. and Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, 74(2):133–148.

Appendix

A Equilibrium Conditions and Model Solutions

It is easier to write out the equilibrium conditions in terms of firms' entry and exit probabilities and the integrated value functions. Given that ε and ϕ are private shocks, firms' strategies can be represented by entry and exit probabilities. Let $\mathbf{P}^* = (\mathbf{P^E}, \mathbf{P^X})$ denote the equilibrium probabilities associated with the optimal strategy σ^*, where the superscript E indicates entry and X represents exit. Let $V(s, \boldsymbol{n}|\mathbf{P}^*) \equiv \mathbb{E}\left(\widetilde{V}(s, \boldsymbol{n}, \varepsilon, \phi, \sigma^*(s, \boldsymbol{n}, \varepsilon, \phi)|\sigma^*)\right)$ denote an integrated value function.[32] Then an OE should satisfy the following conditions:

$$P^E(q, D) = Prob\left(\pi(q, 0, D) - \kappa - \varepsilon + \beta V(q, 1, D|\mathbf{P}^*) > \beta V(q, 0, D|\mathbf{P}^*)\right), \tag{A.2}$$

$$P^X(q, g, D) = Prob(\pi(q, g > 0, D) + \beta V(q, 2, D|\mathbf{P}^*) < \phi), \tag{A.3}$$

Here in the profit and value functions, we write out the components in $s = (q, g, D)$ to differentiate firms' values at various stages of their life. Firms make entry and exit decisions based on which action gives them the optimal payoff at each state. In Equation (A.2), a potential entrant chooses to enter the market if the value of entering now is greater than that of waiting till the next period. Similarly in Equation (A.3), an incumbent chooses to exit now if the scrap value is greater than the value the firm obtains by staying.

In addition, the time-invariant long-run average of the industry state $\boldsymbol{n}$ in an OE should satisfy the following law of motion:

$$n(q, 0, I) = N(q, I)P^E(q, I); \tag{A.4}$$

$$n(q, 1, I) = n(q, 0, I)(1 - P^X(q, 1, I)); \tag{A.5}$$

$$n(q, 2, I) = (n(q, 1, I) + n(q, 2, I))(1 - P^X(q, 2, I)); \tag{A.6}$$

$$n(q, c) = N(q, c)P^E(q, c) + n(q, c)(1 - P^X(q, c)) \tag{A.7}$$

where $N(q, D)$ is N_H if $q = \bar{q}$ and $D = I$, N_L if $q = \underline{q}$ and $D = I$, N_{Hc} if $q = \bar{q}$ and $D = c$, and N_{Lc} if $q = \underline{q}$ and $D = c$. Equations (A.4) to (A.6) show how the number of independent restaurants at each perceived quality level evolves. The equilibrium conditions in Equations (A.2) to (A.7) yield a system of non-linear equations, of which the solution $\mathbf{P}^*$ is a fixed point.[33] For any given value

[32]The integrated value function can be expressed in terms of the entry and exit probabilities:

$$V(s, \boldsymbol{n}|\mathbf{P}^\sigma) = \begin{cases} P^E_\sigma(s, \boldsymbol{n})\left(\pi(s, \boldsymbol{n}) - \kappa(q, D) - \mathbb{E}(\varepsilon|a = 1) + \beta V(q, 1, D, \boldsymbol{n}'|\mathbf{P}^\sigma)\right) + (1 - P^E_\sigma(s, \boldsymbol{n}))\beta V(s, \boldsymbol{n}'|\mathbf{P}^\sigma) & \text{if } g = 0, \\ (1 - P^X_\sigma(s, \boldsymbol{n}))\left(\pi(s, \boldsymbol{n}) + \beta V(q, 2, D, \boldsymbol{n}'|\mathbf{P}^\sigma)\right) + P^X_\sigma(s, \boldsymbol{n})\mathbb{E}(\phi|a = 0) & \text{if } g > 0, \end{cases} \tag{A.1}$$

The first expression in Equation (A.1) represents the integrated value function for potential entrants ($g = 0$). The second expression in Equation (A.1) is the integrated value function for the incumbents ($g > 0$).

[33]Specifically, if we substitute Equations (A.4)-(A.7) into Equations (A.2) and (A.3), we obtain a system of equations with the equilibrium probabilities on both sides. Here the number of equations is equal to the number of unknown probabilities, each of which lies between 0 and 1. By applying Brouwer's Fixed Point Theorem, we can demonstrate the existence of an equilibrium.

of $\gamma \in [0,1]$, an equilibrium solution $\mathbf{P}^*$ exists.

As described in the main text, a restaurant experiences 3 stages: new, young, and established. The perceived quality of a restaurant at the young stage is heavily influenced by the level of information about quality disclosed. Given different perceptions of quality, restaurants at different stages will have different flow profits. Following Weintraub et al. (2008), we can compute the integrated value functions for restaurants at different stages. For notation simplicity, following the main text, we convert the arguments inside the integrated values to subscripts and write them as V_{Tg} and V_{Tc} for the independent and chain restaurants at stage g respectively. For example, the integrated function for high-quality independent young restaurant $V(\bar{q}, 1, I)$ will be denoted as V_{H1}. To establish the existence of OE solution, we require the extreme value of error term, b_0 and b_1 to be high enough. In particular, the following conditions should hold:

$$-2\beta b_1 < -A_T + 2\sqrt{b_1}B_T < 0 \tag{A.8}$$

$$b_1 > \frac{1}{(1-\beta)^2}\max\{B_T, B_{Tc}\} \tag{A.9}$$

$$b_0 > (1-\beta)\max\{E_T, E_{Tc}\}, \tag{A.10}$$

where $A_T = 2b_1 + \beta(\pi_{T1} - \pi_{T2})$, $B_T = \sqrt{(b_1 - \beta b_1 - \beta\pi_{T2})}$, $B_{Tc} = \sqrt{(b_1 - \beta b_1 - \beta\pi_{Tc})}$, $E_T = -\kappa_T + \beta V_{T1} + \pi_{T0}$, and $E_{Tc} = -\kappa_{Tc} + \beta V_{Tc1} + \pi_{Tc}$ $\forall T \in \{H, L\}$. Based on the definition of OE (Weintraub et al., 2008), along with integrated value functions computed in Table A.1, it is straightforward to derive the OE solutions, which are functions of the equilibrium entry and exit probabilities and long-run invariant distributions of restaurant types.

Table A.1: Value Functions

Symbol	Expression
V_{T0}	$(-b_0(-2+\beta) + \beta E_T - 2\sqrt{b_0(1-\beta)(b_0+\beta E_T)}/\beta^2$
V_{T1}	$(A_T - 6\sqrt{b_1}B_T)(A_T + 2\sqrt{b_1}B_T)/(4b_1\beta^2)$
V_{T2}	$(\sqrt{b_1} - \sqrt{b_1}B_T)^2/\beta^2$
V_{Tc0}	$(-b_0(-2+\beta) + \beta E_{Tc} - 2\sqrt{b_0(1-\beta)(b_0+\beta E_{Tc})}/\beta^2$
V_{Tc1}	$(\sqrt{b_1} - \sqrt{b_1}B_{Tc})^2/\beta^2$

Note: $A_T = 2b_1 + \beta(\pi_{T1} - \pi_{T2})$, $B_T = \sqrt{(b_1 - \beta b_1 - \beta\pi_{T2})}$, $B_{Tc} = \sqrt{(b_1 - \beta b_1 - \beta\pi_{Tc})}$, $E_T = -\kappa_T + \beta V_{T1} + \pi_{T0}$, and $E_{Tc} = -\kappa_{Tc} + \beta V_{Tc1} + \pi_{Tc}$ $\forall T \in \{H, L\}$

Table A.2: Model Solution

Variable	Meaning	Expression
P_T^E	entry probabilities for independent new entrants ($g=0$) of quality type $T \in \{H, L\}$	$1+\frac{-\sqrt{b_0}+\sqrt{(1-\beta)(b_0+\beta E_T)}}{\beta\sqrt{b_0}}$
P_{T1}^X	exit probabilities for young independent restaurants ($g=1$) of quality type $T \in \{H, L\}$	$1+\frac{-A_T+2\sqrt{b_1}B_T}{2\beta b_1}$
P_{T2}^X	exit probabilities for established independent restaurants ($g=2$) of quality type $T \in \{H, L\}$	$1+\frac{-\sqrt{b_1}+B_T}{\beta\sqrt{b_1}}$
P_{Tc}^E	entry probabilities for chain new entrants ($g=0$) of quality type $T \in \{H, L\}$	$1+\frac{-\sqrt{b_0}+\sqrt{(1-\beta)(b_0+\beta E_{Tc})}}{\beta\sqrt{b_0}}$
P_{Tc}^X	exit probabilities for chain restaurants ($g=1,2$) of quality type $T \in \{H, L\}$	$1+\frac{-\sqrt{b_1}+B_{Tc}}{\beta\sqrt{b_1}}$

Note: $A_T = 2b_1+\beta(\pi_{T1}-\pi_{T2})$, $B_T = \sqrt{(b_1-\beta b_1-\beta\pi_{T2})}$, $B_{Tc} = \sqrt{(b_1-\beta b_1-\beta\pi_{Tc})}$, $E_T = -\kappa_T+\beta V_{T1}+\pi_{T0}$, and $E_{Tc} = -\kappa_{Tc}+\beta V_{Tc1}+\pi_{Tc}$ $\forall T \in \{H, L\}$.

B Assumption and Statement of Propositions

B.1 Assumption

To eliminate extreme cases of the effect of γ, we make an explicit assumption below:

Assumption B.1. The market size in dollar value M is large enough such that

1. $\left(1+\frac{(\sqrt{b_1}-B_H)}{B_H}\exp(\bar{q}-\hat{q}_{H1})\right)\frac{-(n_{L1}\exp(\hat{q}_{L1})-n_{H1}\exp(\hat{q}_{H1}))}{\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')+1} < 1$ when $\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')/\partial\gamma > 0$;
2. $\left(1+\frac{(\sqrt{b_1}-B_L)}{B_L}\exp(\underline{q}-\hat{q}_{L1})\right)\frac{n_{L1}\exp(\hat{q}_{L1})-n_{H1}\exp(\hat{q}_{H1})}{\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')+1} < 1$ when $\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')/\partial\gamma < 0$,

where $B_T = \sqrt{(b_1-\beta b_1-\beta\pi_{T2})}, \forall T \in \{H, L\}$.

The conditions in Assumption B.1 can easily hold when there are many firms in the market, such that $\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')+1$ is very large. This assumption is to eliminate cases where the number of firms from every perceived quality type is going up (down) with γ, but the competition faced by firms, $\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')$, is going down (up). That is, we eliminate the cases where the decrease (increase) in the perceived quality of competitors is so large that it overturns the increase in the number of competitors from all perceived-quality categories. This assumption ensures that when the number of firms increases in every perceived-quality level (as a result of quality disclosure), the competition, $\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')$, also increases.

B.2 Statements of Propositions 2-4 and Corollary 1

Proposition 2. The sign of the effect of γ on competition, $\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')/\partial\gamma$, depends on the sum of $\bar{F}_L$ and $\bar{F}_H$. In particular,

(1) if $\bar{F}_L + \bar{F}_H > 0$, then $\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')/\partial\gamma > 0$;

(2) if $\bar{F}_L + \bar{F}_H = 0$, then $\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')/\partial\gamma = 0$, and

(3) if $\bar{F}_L + \bar{F}_H < 0$, then $\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')/\partial\gamma < 0$.

The proof of Proposition 2 is shown in Online Appendix C.2.

Corollary 1.

(1) When $\bar{F}_L + \bar{F}_H > 0$, $CE(P^E_T) < 0$, $CE(P^X_{T1}) > 0$, $CE(P^X_{T2}) > 0$, $CE(P^E_{Tc}) < 0$, $CE(P^X_{Tc}) > 0, \forall T \in \{H, L\}$.

(2) When $\bar{F}_L + \bar{F}_H = 0$, $CE(P) = 0, \forall P$.

(3) When $\bar{F}_L + \bar{F}_H < 0$, $CE(P^E_T) > 0$, $CE(P^X_{T1}) < 0$, $CE(P^X_{T2}) < 0$, $CE(P^E_{Tc}) > 0$, $CE(P^X_{Tc}) < 0, \forall T \in \{H, L\}$.

The proof of Corollary 1 is trivial and thus omitted.

Proposition 3. When $\bar{F}_L + \bar{F}_H > 0$, the signs of the effects of γ on all equilibrium probabilities are certain except for P^E_H and P^X_{H1}. In particular, $\partial P^E_L/\partial\gamma < 0$, $\partial P^X_{L1}/\partial\gamma > 0$, $\partial P^E_{Tc}/\partial\gamma < 0$, $\partial P^X_{T2}/\partial\gamma > 0$, and $\partial P^X_{Tc}/\partial\gamma > 0, \forall T \in \{H, L\}$. As for P^E_H and P^X_{H1}, only two cases arise: (1) $\partial P^E_H/\partial\gamma > 0$ and $\partial P^X_{H1}/\partial\gamma < 0$. (2) $\partial P^E_H/\partial\gamma < 0$ and $\partial P^X_{H1}/\partial\gamma < 0$.

The proof of Proposition 3 is shown in Online Appendix C.3.

Proposition 4. When $\bar{F}_L + \bar{F}_H < 0$, the signs of the effects of γ on all equilibrium probabilities are certain except for P^E_L and P^X_{L1}. In particular, $\partial P^E_H/\partial\gamma > 0$, $\partial P^X_{H1}/\partial\gamma < 0$, $\partial P^E_{Tc}/\partial\gamma > 0$, $\partial P^X_{T2}/\partial\gamma < 0$, and $\partial P^X_{Tc}/\partial\gamma < 0, \forall T \in \{H, L\}$. As for P^E_L and P^X_{L1}, only two cases arise: (1) $\partial P^E_L/\partial\gamma > 0$ and $\partial P^X_{L1}/\partial\gamma > 0$. (2) $\partial P^E_L/\partial\gamma < 0$ and $\partial P^X_{L1}/\partial\gamma > 0$.

The proof of Proposition 4 is shown in Online Appendix C.4.

ONLINE APPENDICES [Intended to be made available online.]

C Proofs of Propositions

C.1 Proposition 1 Proof

Proof. For chain restaurants, the perceived quality does not change with respect to γ, and therefore, the DEs on chain restaurants are 0. For independent restaurants, we need to show the following:

$$\frac{\partial P_H^E}{\partial \pi_{H1}} \frac{\partial \pi_{H1}}{\partial M \exp(\hat{q}_{H1})} \frac{\partial M \exp(\hat{q}_{H1})}{\partial \hat{q}_{H1}} \frac{\partial \hat{q}_{H1}}{\partial \gamma} > 0, \tag{C.1}$$

$$\frac{\partial P_{H1}^X}{\partial \pi_{H1}} \frac{\partial \pi_{H1}}{\partial M \exp(\hat{q}_{H1})} \frac{\partial M \exp(\hat{q}_{H1})}{\partial \hat{q}_{H1}} \frac{\partial \hat{q}_{H1}}{\partial \gamma} < 0, \tag{C.2}$$

$$\frac{\partial P_{H2}^X}{\partial \pi_{H1}} \frac{\partial \pi_{H1}}{\partial M \exp(\hat{q}_{H1})} \frac{\partial M \exp(\hat{q}_{H1})}{\partial \hat{q}_{H1}} \frac{\partial \hat{q}_{H1}}{\partial \gamma} = 0, \tag{C.3}$$

$$\frac{\partial P_L^E}{\partial \pi_{L1}} \frac{\partial \pi_{L1}}{\partial M \exp(\hat{q}_{L1})} \frac{\partial M \exp(\hat{q}_{L1})}{\partial \hat{q}_{L1}} \frac{\partial \hat{q}_{L1}}{\partial \gamma} < 0, \tag{C.4}$$

$$\frac{\partial P_{L1}^X}{\partial \pi_{L1}} \frac{\partial \pi_{L1}}{\partial M \exp(\hat{q}_{L1})} \frac{\partial M \exp(\hat{q}_{L1})}{\partial \hat{q}_{L1}} \frac{\partial \hat{q}_{L1}}{\partial \gamma} > 0, \tag{C.5}$$

$$\frac{\partial P_{L2}^X}{\partial \pi_{L1}} \frac{\partial \pi_{L1}}{\partial M \exp(\hat{q}_{L1})} \frac{\partial M \exp(\hat{q}_{L1})}{\partial \hat{q}_{L1}} \frac{\partial \hat{q}_{L1}}{\partial \gamma} = 0. \tag{C.6}$$

Equations (C.3) and (C.6) are obvious because $\pi_{T1}, T \in \{H, L\}$ does not enter the exit probabilities at the "established" stage of the firm. It is also evident that $(\partial \pi_{T1}/\partial M \exp(\hat{q}_{T1}))(\partial M \exp(\hat{q}_{T1})/\partial \hat{q}_{T1}) = \pi_{T1} > 0, \forall T \in \{H, L\}$. Furthermore, $\partial \hat{q}_{H1}/\partial \gamma = (\bar{q} - \underline{q})/2 > 0$ and $\partial \hat{q}_{L1}/\partial \gamma = (\underline{q} - \bar{q})/2 < 0$. The only work left to show is the signs of the derivatives of the probabilities with respect to the "young" stage flow profit. Below we derive these derivatives.

$$\frac{\partial P_T^E}{\partial \pi_{T1}} = \frac{\partial P_T^E}{\partial V_{T1}} \frac{\partial V_{T1}}{\partial \pi_{T1}} > 0, \tag{C.7}$$

$$\frac{\partial P_{T1}^X}{\partial \pi_{T1}} = -\frac{1}{2b_1} < 0, \forall T \in \{H, L\} \tag{C.8}$$

For the entry probabilities, as shown in the expression for entry probabilities in Table A.2, $\partial P_T^E/\partial V_{T1} > 0$, and as shown in the expression for V_{T1} in Table A.1, $\partial V_{T1}/\partial \pi_{T1} > 0$. Therefore, $\partial P_T^E/\partial \pi_{T1} > 0$. With the signs of $\partial P_T^E/\partial \pi_{T1}$ and $\partial P_{T1}^X/\partial \pi_{T1}$ in hand, the signs of the expressions C.1 to C.6 follow naturally. □

C.2 Proposition 2 Expressions and Proof

C.2.1 Expressions of $\bar{F}_L$ and $\bar{F}_H$

$$\bar{F}_L = N_L \left[\exp(\hat{q}_0)\frac{\partial P_L^E}{\partial\gamma} + \exp(\hat{q}_{L1})\left(\frac{\partial P_L^E(1-P_{L1}^X)}{\partial\gamma}\right) + \exp(\underline{q})\left(\frac{\partial P_L^E(1-P_{L1}^X)(1/P_{L2}^X-1)}{\partial\gamma} + \frac{N_{Lc}}{N_L}\frac{\partial P_{Lc}^E(1/P_{Lc}^X-1)}{\partial\gamma}\right)\right]$$
$$+N_L P_L^E(1-P_{L1}^X)\exp(\hat{q}_{L1})\frac{(\underline{q}-\bar{q})}{2}. \quad \text{(C.9)}$$

$$\bar{F}_H = N_H \left[\exp(\hat{q}_0)\frac{\partial P_H^E}{\partial\gamma} + \exp(\hat{q}_{H1})\left(\frac{\partial P_H^E(1-P_{H1}^X)}{\partial\gamma}\right) + \exp(\bar{q})\left(\frac{\partial P_H^E(1-P_{H1}^X)(1/P_{H2}^X-1)}{\partial\gamma} + \frac{N_{Hc}}{N_H}\frac{\partial P_{Hc}^E(1/P_{Hc}^X-1)}{\partial\gamma}\right)\right]$$
$$+N_H P_H^E(1-P_{H1}^X))\exp(\hat{q}_{H1})\frac{(\bar{q}-\underline{q})}{2}. \quad \text{(C.10)}$$

As shown in these equations, $\bar{F}_L$ and $\bar{F}_H$, represent the separate effects of γ on high- and low-quality firms in the market, respectively. For each term, the effect of γ propagates through two channels: one is by changing the perceived quality of firms. This channel is reflected in the last terms of Equations (C.9) and (C.10), i.e. $\exp(\hat{q}_{L1})\frac{(\underline{q}-\bar{q})}{2}$ and $\exp(\hat{q}_{H1})\frac{(\bar{q}-\underline{q})}{2}$, which are the derivatives of $\hat{q}_1$ with respect to γ. It is evident that this channel yields a negative sign in $\bar{F}_L$, but a positive one in $\bar{F}_H$. The other channel is through affecting the distribution of firms over perceived quality levels, i.e. $n(\hat{q}')$. This channel is captured by the first three items of Equations (C.9) and (C.9) through the derivatives of entry and exit probabilities with respect to γ. These two channels together determine the total change in $\bar{F}_L$ and $\bar{F}_H$, which in turn determines the total change in competition. Because $\bar{F}_L$ is weighted by N_L and $\bar{F}_H$ is weighted by N_H, and because the first channel is negative in $\bar{F}_L$ and positive in $\bar{F}_H$, it is easy to infer that if $N_L >> N_H$, competition is likely to decline with γ. We use this insight in choosing settings for the numerical examples in Section 4.3.

C.2.2 Proposition 2 Proof

Proof. As shown in Definition 2, the key determining factor of the directions of CEs is $\partial \sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}')/\partial\gamma$. In particular, when $\partial \sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}')/\partial\gamma = 0$, the CEs are 0. Below we expand $\partial \sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}')/\partial\gamma$, and derive the expressions of $\bar{F}_H$ and $\bar{F}_L$.

$$\frac{\partial \sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}')}{\partial\gamma} = \sum_{\hat{q}'}\left[\frac{\partial n(\hat{q}')}{\partial\gamma}\exp(\hat{q}') + n(\hat{q}')\frac{\partial \exp(\hat{q}')}{\partial\gamma}\right]$$
$$= \sum_{\hat{q}'}\frac{\partial n(\hat{q}')}{\partial\gamma}\exp(\hat{q}') - (n_{L1}\exp(\hat{q}_{L1}) - n_{H1}\exp(\hat{q}_{H1}))\frac{(\bar{q}-\underline{q})}{2}, \text{ where} \quad \text{(C.11)}$$

$$\sum_{\hat{q}'}\frac{\partial n(\hat{q}')}{\partial\gamma}\exp(\hat{q}') = \left(\frac{\partial n_{L0}}{\partial\gamma} + \frac{\partial n_{H0}}{\partial\gamma}\right)\exp(\hat{q}_0) + \frac{\partial n_{L1}}{\partial\gamma}\exp(\hat{q}_{L1}) + \frac{\partial n_{H1}}{\partial\gamma}\exp(\hat{q}_{H1})$$
$$+\left(\frac{\partial n_{L2}}{\partial\gamma} + \frac{\partial n_{Lc}}{\partial\gamma}\right)\exp(\underline{q}) + \left(\frac{\partial n_{H2}}{\partial\gamma} + \frac{\partial n_{Hc}}{\partial\gamma}\right)\exp(\bar{q}) \quad \text{(C.12)}$$

We can express the equilibrium $n(\hat{q})$ based on the model solution (A) and the Equations (A.4) to (A.7), we can write

$$\frac{\partial n_{L0}}{\partial \gamma} = N_L \frac{\partial P_L^E}{\partial \gamma} \tag{C.13}$$

$$\frac{\partial n_{H0}}{\partial \gamma} = N_H \frac{\partial P_H^E}{\partial \gamma} \tag{C.14}$$

$$\frac{\partial n_{L1}}{\partial \gamma} = N_L \frac{\partial P_L^E (1 - P_{L1}^X)}{\partial \gamma} \tag{C.15}$$

$$\frac{\partial n_{H1}}{\partial \gamma} = N_H \frac{\partial P_H^E (1 - P_{H1}^X)}{\partial \gamma} \tag{C.16}$$

$$\frac{\partial n_{L2}}{\partial \gamma} = N_L \frac{\partial P_L^E (1 - P_{L1}^X)(1/P_{L2}^X - 1)}{\partial \gamma} \tag{C.17}$$

$$\frac{\partial n_{H2}}{\partial \gamma} = N_H \frac{\partial P_H^E (1 - P_{H1}^X)(1/P_{H2}^X - 1)}{\partial \gamma} \tag{C.18}$$

$$\frac{\partial n_{Lc}}{\partial \gamma} = N_L \frac{N_{Lc}}{N_L} \frac{\partial P_{Lc}^E (1/P_{Lc}^X - 1)}{\partial \gamma} \tag{C.19}$$

$$\frac{\partial n_{Hc}}{\partial \gamma} = N_H \frac{N_{Hc}}{N_H} \frac{\partial P_{Hc}^E (1/P_{Hc}^X - 1)}{\partial \gamma} \tag{C.20}$$

Substituting Equations (C.13) to (C.20) into Equation (C.12) and then (C.11), we have the effect of γ on competition as follows:

$$\frac{\partial \sum_{\hat{q}'} n(\hat{q}') \exp(\hat{q}')}{\partial \gamma} = \bar{F}_L + \bar{F}_H, \text{ where} \tag{C.21}$$

$$\tag{C.22}$$

$\bar{F}_L$ and $\bar{F}_H$ are expressed in Equation (C.9) and (C.9), respectively. □

C.3 Proposition 3 Proof

Proof. It is easy to show the statement in the first sentence. As shown in Proposition 1 and Corollary 1, when $\bar{F}_L + \bar{F}_H > 0$, DE and CE work in the same direction for all equilibrium probabilities, except for P_H^E and P_{H1}^X. In particular, $\partial P_L^E/\partial\gamma < 0$, $\partial P_{L1}^X/\partial\gamma > 0$, $\partial P_{L2}^X/\partial\gamma > 0$, $\partial P_{H2}^X/\partial\gamma < 0$, $\partial P_{Hc}^E/\partial\gamma < 0$, $\partial P_{Lc}^E/\partial\gamma < 0$, $\partial P_{Hc}^X/\partial\gamma > 0$, and $\partial P_{Lc}^X/\partial\gamma > 0$. Four cases are possible in terms of the signs of $\partial P_H^E/\partial\gamma$ and $\partial P_{H1}^X/\partial\gamma$. We examine each case one by one and eliminate those that lead to contradictions with $\bar{F}_L + \bar{F}_H > 0$ or other conditions of the model.

Case 1 $\partial P_H^E/\partial\gamma > 0$ and $\partial P_{H1}^X/\partial\gamma < 0$. This case requires DEs dominate CEs for both probabilities, which is possible as long as DEs are large enough for both probabilities. In particular,

$\partial P_H^E/\partial\gamma$ and $\partial P_{H1}^X/\partial\gamma$ can be written as follows:

$$\begin{aligned}\frac{\partial P_H^E}{\partial\gamma} &= \frac{\partial P_H^E}{\partial\pi_{H1}}\frac{\partial\pi_{H1}}{\partial M\exp(\hat{q}_{H1})}\frac{\partial M\exp(\hat{q}_{H1})}{\partial\hat{q}_{H1}}\frac{\partial\hat{q}_{H1}}{\partial\gamma}\\ &+\left(\frac{\partial P_H^E}{\partial\pi_{H0}}\frac{\partial\pi_{H0}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}+\frac{\partial P_H^E}{\partial\pi_{H1}}\frac{\partial\pi_{H1}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}+\frac{\partial P_H^E}{\partial\pi_{H2}}\frac{\partial\pi_{H2}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}\right)\frac{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}{\partial\gamma} \qquad (C.23)\\ \frac{\partial P_{H1}^X}{\partial\gamma} &= \frac{\partial P_{H1}^X}{\partial\pi_{H1}}\frac{\partial\pi_{H1}}{\partial M\exp(\hat{q}_{H1})}\frac{\partial M\exp(\hat{q}_{H1})}{\partial\hat{q}_{H1}}\frac{\partial\hat{q}_{H1}}{\partial\gamma}\\ &+\left(\frac{\partial P_{H1}^X}{\partial\pi_{H1}}\frac{\partial\pi_{H1}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}+\frac{\partial P_{H1}^X}{\partial\pi_{H2}}\frac{\partial\pi_{H2}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}\right)\frac{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}{\partial\gamma} \qquad (C.24)\end{aligned}$$

Using Equations (C.23) and (C.24), we can derive that $\partial P_H^E/\partial\gamma > 0$ and $\partial P_{H1}^X/\partial\gamma < 0$ imply the following conditions:

$$\begin{aligned}\tilde{\pi}_{H1}\left(\frac{1}{2}(\bar{q}-\underline{q})-\Lambda\right) &> \frac{2b_1}{(A_H-2\sqrt{b_1}B_H)}\tilde{\pi}_{H0}\Lambda+\left(-1+\frac{b_1(A_H+6\sqrt{b_1}B_H)}{(A_H-2\sqrt{b_1}B_H)\sqrt{b_1}B_H}\right)\tilde{\pi}_{H2}\Lambda \qquad (C.25)\\ \tilde{\pi}_{H1}\left(\frac{1}{2}(\bar{q}-\underline{q})-\Lambda\right) &> \frac{(\sqrt{b_1}-B_H)}{B_H}\tilde{\pi}_{H2}\Lambda, \qquad (C.26)\end{aligned}$$

where $\tilde{\pi}_{Hg} = \pi_{Hg} + C_H, \forall g \in \{0, 1, 2\}$, and

$$\Lambda = \frac{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}{\partial\gamma}\frac{1}{\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')+1}$$

Inequality C.25 comes from $\partial P_H^E/\partial\gamma > 0$, and uses the condition A.8. Inequality C.26 comes from $\partial P_{H1}^X/\partial\gamma < 0$.

It is easy to show that the right-hand-side (RHS) of inequality C.25 is larger than the RHS of inequality C.26. The rationale is that when $\bar{F}_L + \bar{F}_H > 0$, $\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')/\partial\gamma > 0$. Therefore, $\Lambda > 0$, and based on condition A.8, the first term on the RHS of inequality C.25, $\frac{2b_1}{(A_H-2\sqrt{b_1}B_H)}\tilde{\pi}_{H0}\Lambda > 0$.

In addition, because $b_1(A_H + 6\sqrt{b_1}B_H) > b_1(A - 2\sqrt{b_1}B_H)$, which always holds as long as $B_H > 0$, the second term on the RHS of inequality C.25 must be greater than the RHS of inequality C.26, i.e.,

$$\left(-1+\frac{b_1(A_H+6\sqrt{b_1}B_H)}{(A_H-2\sqrt{b_1}B_H)\sqrt{b_1}B_H}\right)\tilde{\pi}_{H2}\Lambda > \frac{(\sqrt{b_1}-B_H)}{B_H}\tilde{\pi}_{H2}\Lambda,$$

Therefore, all of the RHS of inequality C.25 is always larger than the RHS of inequality C.26. As long as inequality C.25 holds, inequality C.26 also holds. Case 1 is possible.

Now we can check whether this case satisfies the assumption $\bar{F}_L + \bar{F}_H > 0$. We know that low-quality restaurants are less likely to enter and more likely to exit as γ increases, therefore, there should be fewer low-quality restaurants. Furthermore, because both high-quality chain and established independent restaurants are more likely to exit, there should also be fewer high-quality chain or established independent restaurants. The decrease in the number of competitors can reduce competition. However, because high-quality independent restaurants are more likely to enter and less likely to exit at the young stage, there should be more new and young high-quality independent restaurants, which can offset the loss in the number of restaurants from other types.

Case 2 $\partial P_H^E/\partial\gamma > 0$ and $\partial P_{H1}^X/\partial\gamma > 0$. This case requires that $DE(P_H^E)$ dominate $CE(P_H^E)$, but $DE(P_{H1}^X)$ be dominated by $CE(P_{H1}^X)$, implying that inequality C.25 and the reverse of inequality C.26 hold at the same time. It is impossible because the RHS of inequality C.25 is always larger than the RHS of inequality C.26. $\tilde{\pi}_{H1}\left(\frac{1}{2}(\bar{q}-\underline{q})-\Lambda\right)$ cannot be simultaneously bigger than a large number and smaller than a small number.

Case 3 $\partial P_H^E/\partial\gamma < 0$ and $\partial P_{H1}^X/\partial\gamma < 0$. This case requires$DE(P_H^E)$ be dominated by $CE(P_H^E)$, but $DE(P_{H1}^X)$ dominate $CE(P_{H1}^X)$, meaning that the reverse of inequality C.25 and inequality C.26 hold. That is,

$$\frac{2b_1}{(A_H-2\sqrt{b_1}B_H)}\tilde{\pi}_{H0}\Lambda+\left(-1+\frac{b_1(A_H+6\sqrt{b_1}B_H)}{(A_H-2\sqrt{b_1}B_H)\sqrt{b_1}B_H}\right)\tilde{\pi}_{H2}\Lambda > \tilde{\pi}_{H1}\left(\frac{1}{2}(\bar{q}-\underline{q})-\Lambda\right) > \frac{(\sqrt{b_1}-B_H)}{B_H}\tilde{\pi}_{H2}\Lambda$$

This condition in the above inequality can hold. Now we only need to check whether this case complies with $\bar{F}_L+\bar{F}_H > 0$. In this case, the numbers of restaurants of all types decline with γ, except for the numbers of young and established high-quality restaurants (i.e. n_{H1} and n_{H2}). The changes in n_{H1} and n_{H2} are uncertain because, although high-quality independent restaurants are less likely to enter, they are also less likely to exit at the young stage, potentially leading to an increase in the number of young high-quality independent restaurants, i.e.n_{H1} can increase. In addition, as more young high-quality independent restaurants transition to the established stage, n_{H2} could potentially increase as well, despite that established restaurants are more likely to exit. The potential increases in n_{H1} and n_{H2} can make up for the loss in the number of restaurants from other types. Therefore, competition can increase, which is consistent with $\bar{F}_L+\bar{F}_H > 0$. This case is possible.

Case 4 $\partial P_H^E/\partial\gamma < 0$ and $\partial P_{H1}^X/\partial\gamma > 0$. This case requires DEs be dominated by CEs for both probabilities, implying that the numbers of restaurants from all types are declining as γ increases. That is, $\frac{\partial n(\hat{q}')}{\partial\gamma} < 0$, and $U \equiv \sum_{\hat{q}'}\frac{\partial n(\hat{q}')}{\partial\gamma}\exp(\hat{q}') < 0$.

In addition, this case also requires both inequality C.25 and inequality C.26 be reversed. Given that the RHS of inequality C.25 is larger than that of inequality C.26, we only need the reverse of inequality C.26 to be true, meaning

$$\frac{\bar{q}-\underline{q}}{2} < \left(1+\frac{(\sqrt{b_1}-B_H)}{B_H}\exp(\bar{q}-\hat{q}_{H1})\right)\frac{U+\Delta}{\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')+1}, \tag{C.27}$$

where $\Delta \equiv -\left(n_{L1}\exp(\hat{q}_{L1})-n_{H1}\exp(\hat{q}_{H1})\right)\frac{(\bar{q}-\underline{q})}{2}$.

Based on Assumption B.1, $\left(1+\frac{(\sqrt{b_1}-B_H)}{B_H}\exp(\bar{q}-\hat{q}_{H1})\right)\frac{\Delta}{\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')+1} < \frac{\bar{q}-\underline{q}}{2}$. In addition, because $U < 0$, inequality C.27 cannot hold. Therefore, Case 4 is impossible. □

C.4 Proposition 4 Proof

Proof. The proof for this proposition is very similar to that for Proposition 3. First, it is easy to show the statement in the first sentence of the proposition. Based on Proposition 1 and Corollary

1, when $\bar{F}_L + \bar{F}_H < 0$, DE and CE work in the same direction for all equilibrium probabilities, except for P_L^E and P_{L1}^X. In particular, $\partial P_H^E/\partial\gamma > 0$, $\partial P_{H1}^X/\partial\gamma < 0$, $\partial P_{H2}^X/\partial\gamma < 0$, $\partial P_{L2}^X/\partial\gamma < 0$, $\partial P_{Hc}^E/\partial\gamma > 0$, $\partial P_{Lc}^E/\partial\gamma > 0$, $\partial P_{Hc}^X/\partial\gamma < 0$, and $\partial P_{Lc}^X/\partial\gamma < 0$. Second, we can iterate over the four possible combinations of the signs of $\partial P_L^E/\partial\gamma$ and $\partial P_{L1}^X/\partial\gamma$.

Case 1 $\partial P_L^E/\partial\gamma < 0$ and $\partial P_{L1}^X/\partial\gamma > 0$. This case require DEs dominate CEs for both probabilities. This condition is possible as long as DEs are large enough for both probabilities. We can write $\partial P_L^E/\partial\gamma < 0$ and $\partial P_{L1}^X/\partial\gamma > 0$ as

$$\begin{aligned}\frac{\partial P_L^E}{\partial\gamma} = {} & \frac{\partial P_L^E}{\partial\pi_{L1}}\frac{\partial\pi_{L1}}{\partial M\exp(\hat{q}_{L1})}\frac{\partial M\exp(\hat{q}_{L1})}{\partial\hat{q}_{L1}}\frac{\partial\hat{q}_{L1}}{\partial\gamma} \\ & + \left(\frac{\partial P_L^E}{\partial\pi_{L0}}\frac{\partial\pi_{L0}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')} + \frac{\partial P_L^E}{\partial\pi_{L1}}\frac{\partial\pi_{L1}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')} + \frac{\partial P_L^E}{\partial\pi_{L2}}\frac{\partial\pi_{L2}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}\right)\frac{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}{\partial\gamma} \end{aligned} \quad \text{(C.28)}$$

$$\begin{aligned}\frac{\partial P_{L1}^X}{\partial\gamma} = {} & \frac{\partial P_{L1}^X}{\partial\pi_{L1}}\frac{\partial\pi_{L1}}{\partial M\exp(\hat{q}_{L1})}\frac{\partial M\exp(\hat{q}_{L1})}{\partial\hat{q}_{L1}}\frac{\partial\hat{q}_{L1}}{\partial\gamma} \\ & + \left(\frac{\partial P_{L1}^X}{\partial\pi_{L1}}\frac{\partial\pi_{L1}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')} + \frac{\partial P_{L1}^X}{\partial\pi_{L2}}\frac{\partial\pi_{L2}}{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}\right)\frac{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}{\partial\gamma} \end{aligned} \quad \text{(C.29)}$$

Using Equations (C.28) and (C.29), we can derive the following conditions that satisfy $\partial P_L^E/\partial\gamma < 0$ and $\partial P_{L1}^X/\partial\gamma > 0$:

$$\tilde{\pi}_{L1}\left(\frac{1}{2}(\underline{q}-\bar{q})-\Lambda\right) < \frac{2b_1}{(A_L-2\sqrt{b_1}B_L)}\tilde{\pi}_{L0}\Lambda + \left(-1+\frac{b_1(A_L+6\sqrt{b_1}B_L)}{(A_L-2\sqrt{b_1}B_L)\sqrt{b_1}B_L}\right)\tilde{\pi}_{L2}\Lambda \quad \text{(C.30)}$$

$$\tilde{\pi}_{L1}\left(\frac{1}{2}(\underline{q}-\bar{q})-\Lambda\right) < \frac{(\sqrt{b_1}-B_L)}{B_L}\tilde{\pi}_{L2}\Lambda, \quad \text{(C.31)}$$

where $\tilde{\pi}_{Lg} = \pi_{Lg} + C_L, \forall g \in \{0,1,2\}$, and

$$\Lambda = \frac{\partial\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')}{\partial\gamma}\frac{1}{\sum_{\hat{q}'}n(\hat{q}')\exp(\hat{q}')+1}$$

Inequality C.30 comes from $\partial P_L^E/\partial\gamma < 0$, and uses the condition A.8. Inequality C.31 comes from $\partial P_{L1}^X/\partial\gamma > 0$.

These two inequalities have a very similar form compared with inequalities C.25 and C.26 from the proof for Proposition 3. In particular, it is easy to show that the RHS of inequality C.30 is smaller than that of inequality C.31. When $\bar{F}_L + \bar{F}_H < 0$, $\Lambda < 0$. Therefore, $\frac{2b_1}{(A_L-2\sqrt{b_1}B_L)}\tilde{\pi}_{L0}\Lambda < 0$, and

$$\left(-1+\frac{b_1(A_L+6\sqrt{b_1}B_L)}{(A_L-2\sqrt{b_1}B_L)\sqrt{b_1}B_L}\right)\tilde{\pi}_{L2}\Lambda < \frac{(\sqrt{b_1}-B_L)}{B_L}\tilde{\pi}_{L2}\Lambda.$$

Then we have

$$\frac{2b_1}{(A_L-2\sqrt{b_1}B_L)}\tilde{\pi}_{L0}\Lambda + \left(-1+\frac{b_1(A_L+6\sqrt{b_1}B_L)}{(A_L-2\sqrt{b_1}B_L)\sqrt{b_1}B_L}\right)\tilde{\pi}_{L2}\Lambda < \frac{(\sqrt{b_1}-B_L)}{B_L}\tilde{\pi}_{L2}\Lambda.$$

For Case 1 to be possible, we only need inequality C.30 to hold.

Now we can check if this case complies with the condition $\bar{F}_L + \bar{F}_H < 0$. Although both the number of high-quality independent restaurants and the number of chain restaurants from all quality types will increase as γ increases, the number of low-quality independent restaurants decreases. This decrease can offset the increase in the number of restaurants from other types, leading to a decrease in competition, consistent with $\bar{F}_L + \bar{F}_H < 0$.

Case 2 $\partial P_L^E/\partial\gamma < 0$ **and** $\partial P_{L1}^X/\partial\gamma < 0$**.** This case requires $DE(P_L^E)$ dominate $CE(P_L^E)$, but $DE(P_{L1}^X)$ be dominated by $CE(P_{L1}^X)$, implying that inequality C.30 and the reverse of inequality C.31 hold. This case is impossible because the RHS of inequality C.30 is smaller than that of inequality C.31.

Case 3 $\partial P_L^E/\partial\gamma > 0$ **and** $\partial P_{L1}^X/\partial\gamma > 0$**.** This case requires the $DE(P_L^E)$ be dominated by $CE(P_L^E)$, but $DE(P_{L1}^X)$ dominate $CE(P_{L1}^X)$, meaning that the reverse of inequality C.30 and inequality C.31 hold. This case is possible.

We can now check if it contradicts with $\bar{F}_L + \bar{F}_H < 0$. In this case, the numbers of restaurants from all types increase, except for the numbers of young and established low-quality restaurants (i.e., n_{L1} and n_{L2}). The changes in n_{L1} and n_{L2} are uncertain because, although low-quality independent restaurants are more likely to enter, they are also more likely to exit at the young stage, potentially leading to a decrease in the number of young low-quality independent restaurants, i.e., n_{L1} could decline. Furthermore, as fewer young low-quality independent restaurants transition to the established stage, n_{L2} could potentially decrease as well, even though established restaurants are less likely to exit. The potential decreases in n_{L1} and n_{L2} might offset the increase in the number of restaurants from other types. Therefore, competition can decrease, which is consistent with $\bar{F}_L + \bar{F}_H < 0$. This case is possible.

Case 4 $\partial P_L^E/\partial\gamma > 0$ **and** $\partial P_{L1}^X/\partial\gamma < 0$**.** This case requires DEs be dominated by CEs for both probabilities, implying that the numbers of restaurants from all types will increase.

This case also requires the reverse of inequality C.31 to hold, meaning

$$\frac{\bar{q} - \underline{q}}{2} < \left(1 + \frac{(\sqrt{b_1} - B_L)}{B_L} \exp(\bar{q} - \hat{q}_{L1})\right)(-\Lambda) \tag{C.32}$$

This condition C.32 contradicts with Assumption B.1. Therefore, this case is impossible. □

Table C.1: Summary of Effect of Quality Disclosure on Entry and Exit Probabilities of Chain Restaurants

Condition	**Case**	**High-quality**		**Low-quality**	
		P_{Hc}^E	P_{Hc}^X	P_{Lc}^E	P_{Lc}^X
$\bar{F}_L + \bar{F}_H > 0$	Proposition 3	↓	↑	↓	↑
$\bar{F}_L + \bar{F}_H = 0$	Proposition1	-	-	-	-
$\bar{F}_L + \bar{F}_H < 0$	Proposition 4	↑	↓	↑	↑

Note: - means no change.

D Numerical Example Setup

D.1 Parameter Space Setup for Numerical Examples

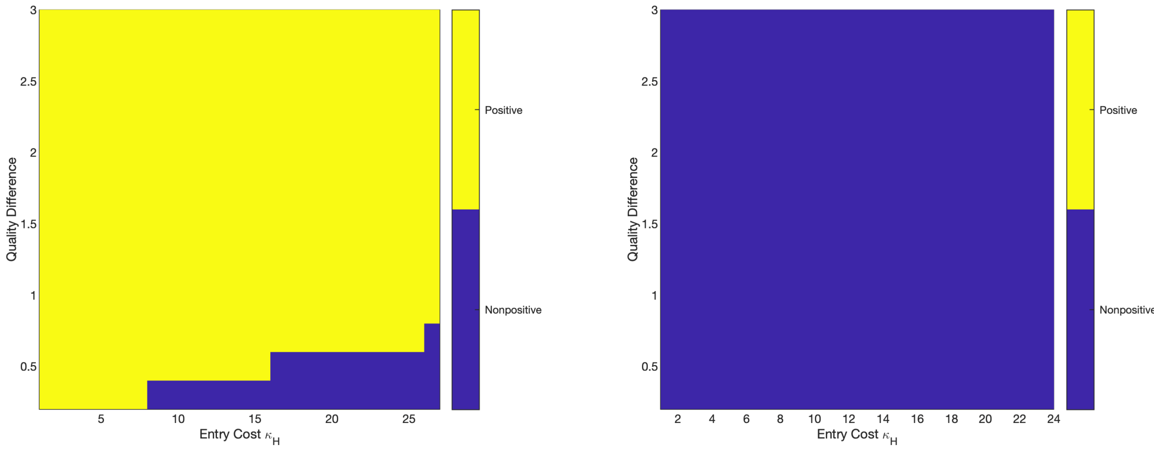
The Case of $\mathbf{N_L = N_H}$ For this case, the specifications of all parameters of the model are shown in Table D.1. For the market size parameter M and the numbers of potential entrants (e.g. N_H), we set them to relatively large numbers because the OE solution concept requires a very high market size and large numbers of potential entrants. We set the number of potential entrants for chain restaurants at a relatively smaller level than those for independent restaurants to reflect the fact that in reality, chain restaurants are harder to come by. Altering the market size and potential entrants parameters would not make a difference in our solutions or the properties of these solutions. Therefore, we set them constant.

For the per-period fixed cost parameters (e.g. C_H), we set them relatively small in order to ensure that the flow-profits in each period for all firms are positive. Again this is a requirement in the OE solution concept (See the third item in Assumption 1 in Weintraub et al. (2008)). Moreover, we set the fixed costs for chain restaurants to be lower than those for independent restaurants (e.g. $C_{Hc} < C_H$) to reflect that fact that chain restaurants may benefit from economies of scale, hence facing lower costs. Regarding the discount factor β, we set it to a relatively small number 0.8 to capture the fact that consumer learning make take a long time. A smaller β makes the profits from established stage of independent restaurants less important and therefore allows a greater effect of γ. For b_0 and b_1, we set them equal for simplicity, and we set them large enough such that the model has a real solution. Smaller b's also amplify the effect of γ on equilibrium entry and exit probabilities.

Table D.1: Parameter Specifications in the Numerical Example

Variable	Value	Variable	Value	Variable	Value
M	1000	β	0.8	$\underline{q}$	0
N_H	1000	C_H	0.4	$\bar{q}$	$[0.2, 3]$
N_L	1000	C_L	0.2	κ_H	$[1, 24]$
N_{Hc}	200	C_{Hc}	0.2	κ_L	$0.7\kappa_H$ $[0.7, 18.9]$
N_{Lc}	200	C_{Lc}	0.1	$\kappa_{Hc} = \kappa_{Lc}$	$0.8\kappa_H$, $[0.8, 21.6]$
				$b_0 = b_1$	55

The Case of $\mathbf{N_L >> N_H}$ For the case with a larger number of potential low-quality entrants, we use the same parameter specification shown in Table D.1 except that $N_L = 3000$ and $N_H = 500$.



(a) $N_L = N_H$ (b) $N_L >> N_H$

Figure D.1: Change in Competition

D.2 Effect on Competition

E Prediction of Missing Ratings with Machine Learning

To predict the missing ratings, we use all observable restaurant characteristics and market characteristics, including restaurant cuisine types, prices, metro location, revenue changes, demographics, etc. Including all these variables, especially location dummies, led to a high-dimensional set of vectors (119), significantly slowing down the random forest algorithm we employ to predict ratings. To address this issue, we conduct a principal component analysis (PCA) to reduce the dimension of the predictive variables. The PCA transforms the predictive variables into a set of principal components that are linear combinations of the predictive variables. Each principal component (PC) is associated with an eigen value, which indicates how much variation of the original vectors a PC captures. The higher the eigen value, the greater variation a PC captures. We select the top 76 PCs with the highest eigen values. They explain over 80% of the variation in the original set of predictive vectors. We use these PCs as the predictive variables in the random forest algorithm to forecast ratings.

To predict ratings, we first train a random-forest prediction model using all the restaurants with known Google ratings. More specifically, we conduct 4-fold cross-validation when training and fine-tuning the model. That is, we break down the data with known ratings into four equal quarters and use three quarters to train the model and the last quarter to validate the model. We rotate the training and validation sets to fine tune the model. Furthermore, we use a minimum node size of 5 to avoid over-fitting. The random forest model performs well with an RMSE of 0.14 and R-squared at 0.924. We then apply the trained model to the data with unknown ratings to obtain the prediction.

Table E.1: Google Rating Distribution (with Missing Ratings)

	Min	Median	Mean	Max	Number of Restaurants
Average Google rating for independent	1.5	3.9	3.89	5.0	6,350
Average Google rating for chain	1.9	3.6	3.60	4.8	2,668
All markets	1.5	3.8	3.78	5.0	9,018
College-town markets	1.9	3.8	3.77	5.0	2,901
Highway-exit markets	2.0	3.7	3.68	4.9	1,139
Others	1.5	3.9	3.81	4.9	4,978

Note: Rating information was collected in November 2016. Number of restaurants represents the number of restaurants with Google ratings.

The distribution of the imputed Google ratings is very similar to that of the original data. Table E.1 shows the Google rating distribution from the original data with missing ratings. As can be seen, they are very close to those from the imputed ratings in Table 3 and Table 4.

F Robustness Checks

F.1 Entry and Exit Analysis with Missing Ratings

In this section, we provide robustness check to our analysis using the original data with missing Google ratings. As mentioned in the main text, there is potential endogeneity associated with the missing data. Therefore, we need additional controls to address the potential endogeneity.

Entry The main econometric specification is as follows:

$$\begin{aligned}\log(E(N_{qfct}|\mathbf{X},\boldsymbol{\theta},\ldots)) =& (\theta_y + \theta_{yr} Rating_{qfct}) \log(Yelp_{ct})(1 - D_f^{ch}) + (\theta_y^{ch} + \theta_{yr}^{ch} Rating_{qfct}) \log(Yelp_{ct}) D_f^{ch} \\ &+ \mathbf{X_{ct}}\boldsymbol{\theta_x} + \theta_\alpha \log(\alpha_{fct}) + \theta_{lt} + \theta_{lt}^{ch} + \theta_{qfc} + \theta_{Mt} + \theta_{Mt}^{ch}, \qquad \text{(F.1)}\end{aligned}$$

Compared to the regression specification (Equation (4)) in the main paper, we include three more variables to deal with the potential endogeneity caused by missing ratings: $\log(\alpha_{fct})$, θ_{lt} and θ_{lt}^{ch}. Here, α_{fct} is the share of new entries with observable Google ratings out of all new entrants by chain affiliation f in each county c and time t. The variables θ_{lt} and θ_{lt}^{ch} are the low-quality-time fixed effects for independent and chain restaurants, respectively. They control for a separate time trend of low-quality restaurants apart from other quality levels. We define low-quality as the lowest quality bin in the 3-bin quality classification and the lowest 2 bins in the 5 and 7 quality-bin classifications. The reasons for including these additional controls are as follows.

An important challenge to identification in the entry analysis with the missing ratings is that we do not observe the quality of all new entrants during our sample period. Although we observe all new entries and their chain affiliation in each market, we cannot tell the quality of a new restaurant if it did not have a rating on Google. To approximate N_{qfct} in Equation (F.1), we use the number of new entries that have ratings on Google instead; however, they represent only a portion of

N_{qfct}. About 60% of the new entrants around March, 2005 had ratings on Google. This share increased over time. By the end of the sample period, about 85% of the new entrants have ratings on Google. This increasing trend is likely due to two processes: (1) one is that a smaller proportion of restaurants were listed online when online review platforms first started; (2) the other is that many restaurants exited long before the data on ratings were collected. These two processes could cause bias in our estimates if they are correlated with online review platforms' penetration.

For the first process, if the portion is uncorrelated with online review platforms' penetration, then the metro-region×time fixed effect will absorb the effect of the changing ratio because the logarithm of the ratio comes into Equation (F.1) linearly. However, if the ratio is correlated with online review platform's penetration, then not accounting for it will cause omitted variable bias. Therefore, on the right-hand side of Equation (4), we also add the natural logarithm of the share of the new entries listed on Google by chain affiliation in each county and each month ($\log(\alpha_{fct})$).[34]

For the second process related to exit, it is likely that more lower-quality restaurants had exited than higher-quality restaurants by the end of the sample period. Therefore, in terms of the quality-makeup, the new entries with Google ratings from the earlier years are likely to have a lower share of low-quality restaurants than those from the later years. If exits are uncorrelated with online review platforms' penetration, then by controlling for a common trend in the entries of low-quality restaurants across high- and low-penetration areas, we can account for this systematic under-representation of low-quality entries observed in the older years. For this reason, we also add low-quality-time-chain-affiliation fixed effects to the right-hand side of Equation (F.1). However, based on our theoretical results in Section 4, we know that exits should be affected by online review platforms' penetration. In particular, there is a greater under-counting of low-quality entries in high-penetration areas in the older years than in low-penetration areas because online reviews speed up the exit of low-quality restaurants (especially young independent restaurants). In this case, the change in low-quality entries in high-penetration area will be dampened, leading to a downward bias on the magnitude of the estimated effect of online review platforms' penetration. In this regard, our estimates from Equation (F.1) can be seen as a lower bound.

The entry results from pooling all markets are shown in Table F.1 and the results for different market types are shown in Table F.2. Following the practice in the main text, we present the composite coefficients for different market types in Table F.3. These results are qualitatively the same as those from the sample with imputed ratings.

Exit The regression specification is the same as the main paper. The only difference is that restaurants without Google ratings are not included in the regression. The results are shown in Tables F.4 and F.5. Again, the results are consistent with those in the main text with imputed ratings.

[34]Note that it is not possible to have this ratio by quality level because we simply do not observe the quality of all restaurants.

Table F.1: Effects of Online Review Platforms on Entry (Missing Ratings)

	(1)	(2)	(3)
	entry	entry	entry
log(Yelp) (independent)	-0.415***	-0.413***	-0.398***
	(0.113)	(0.103)	(0.0946)
log(Yelp)×rating (independent)	0.105***	0.101***	0.107***
	(0.0281)	(0.0273)	(0.0200)
log(Yelp) (chain)	0.116	-0.0612	-0.178
	(0.120)	(0.152)	(0.149)
log(Yelp)×rating (chain)	-0.0128	0.0456	0.0473
	(0.0263)	(0.0349)	(0.0351)
Controls	✓	✓	✓
Group FE	✓	✓	✓
Year× Month× Metro× Chain FE	✓	✓	✓
Year× Month× Low-quality× Chain FE	✓	✓	✓
N	2,734	3,126	3,301
N of Clusters	41	35	34
Number of Quality Levels	3	5	7

Controls include population, income, age, race, and population density, ratio of new entries that are listed on Google. Group FE is quality-bin × chain-affiliation × county fixed effects. Low-quality includes the lowest 2 levels when 5 and 7 quality levels are used, and only the lowest level when 3 levels are used. All standard errors are two-way clustered at the county level and time level. They are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table F.2: Effects of Online Review Platforms on Entry by Chain Affiliation and Market Types (Missing Ratings)

Dependent variable	Entry					
	Independent			Chain		
	College-town	Highway-exit	Others	College-town	Highway-exit	Others
log(Yelp)	-0.609***	0.0792	-0.663***	0.129	-0.0683	0.155
	(0.128)	(0.0960)	(0.119)	(0.124)	(0.0674)	(0.119)
log(Yelp)×rating	0.124***	-0.0145	0.177***	-0.0103	-0.0153	-0.0355
	(0.0313)	(0.0252)	(0.0191)	(0.0334)	(0.0182)	(0.0385)
Controls	✓					
Group FE	✓					
Year×Month×Metro×Chain × Market-type FE	✓					
Year×Month×Low-quality×Chain×Market-type FE	✓					
N	2152					
N of Clusters	35					

Controls include population, income, age, race, and population density, ratio of new entries that are listed on Google. Group FE is quality-bin×chain-affiliation×county fixed effects. Low-quality is the lowest level in the 3-bin quality classification. All standard errors are clustered at the county level and areshown in parenthese * p<0.10, ** p<0.05, *** p<0.01

Table F.3: Effects on Entry by Google Star Rating and Market Type (Independent) (Missing Ratings)

Star Rating	2	3	4	5
College-town	-0.3617***	-0.2380***	-0.1144*	0.0093
	(0.0803)	(0.0661)	(0.0651)	(0.0779)
Highway-exit	0.0502	0.0356	0.0211	0.0065
	(0.0629)	(0.0569)	(0.0615)	(0.0748)
Others	-0.3018***	-0.1252	0.0515	0.2281***
	(0.0902)	(0.0791)	(0.0717)	(0.0691)

All standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table F.4: Effects on Exit by Chain Affiliation and Age (Missing Ratings)

Dependent variable	Exit			
	Independent		Chain	
	Established	Young	Established	Young
log(Yelp)	-0.000103	0.0103***	-0.00139	0.000838
	(0.00328)	(0.00363)	(0.00114)	(0.000980)
log(Yelp)×rating	-0.000299	-0.00260***	0.000202	-0.000321
	(0.000828)	(0.000895)	(0.000351)	(0.000262)
Controls	✓			
Restaurant×Month FE	✓			
Year×Month×Metro×Chain FE	✓			
N	363630			
N of Clusters	97			

Established restaurants are those that are older than 12 years. Controls include population, income, age, race, and population density. All standard errors are clustered at the county level and are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table F.5: Effects on Exit by Google Star Rating (Young Independent) (Missing Ratings)

Star Rating	2	3	4	5
Effect	0.0051**	0.0025**	-0.0001	-0.0027**
	(0.0019)	(0.0012)	(0.0009)	(0.0013)

All standard errors are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

F.2 Quality Improvement and Placebo Test

Table F.6: Effects on Quality by Restaurant Affiliation

	(1) Mean Rating	(2) Percentage of High Quality Resturants
log(Yelp) (independent)	0.0120***	0.0187***
	(0.00389)	(0.00587)
log(Yelp) (chain)	0.00365	-0.00275
	(0.00501)	(0.00593)
Controls	✓	✓
Group FE	✓	✓
Year× Metro× Chain FE	✓	✓
N	22,696	22,696
N of Clusters	110	110

Controls include population, income, age, race, and population density. Group FE is chain-affiliation × county fixed effect. All standard errors are clustered at the county level. They are shown in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table F.7: Placebo Test for Entry (3-Quality-Bin Classification)

	(1) entry
log(Yelp) (independent)	-0.148
	(0.116)
log(Yelp)×rating (independent)	0.0273
	(0.0214)
log(Yelp) (chain)	0.115
	(0.161)
log(Yelp)×rating (chain)	0.0320
	(0.0275)
Controls	✓
Year× Month× Metro × Chain FE	✓
Group FE	✓
N	1,448
N of Clusters	31

Controls include population, income, age, race, and population density. All standard errors are clustered at the county level. They are shown in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table F.8: Placebo Test for Exit

	(1) exit
log(Yelp) (independent)	0.00324
	(0.00416)
log(Yelp)×rating (independent)	-0.00136
	(0.00104)
log(Yelp) (chain)	-0.00435
	(0.00409)
log(Yelp)×rating (chain)	0.00138
	(0.00112)
Controls	✓
Restaurant× Month FE	✓
Year× Month× Metro× Chain FE	✓
N	235,107
N of Clusters	78

Controls include population, income, age, race, and population density. All standard errors are clustered at the county level. They are shown in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

G Derivation of Firms' Profit Function

The profit function is derived as follows. Consumer i takes the stock of restaurants in the market as given and chooses a restaurant (indexed by j) based on the expected quality they will get from dining at the restaurant:

$$U_{ijt} = \hat{\hat{q}}_{jt} - p_{jt} + e_{ijt}, \tag{G.1}$$

where U_{ijt} is consumer i's utility from dining at restaurant j at time t. $\hat{\hat{q}}_{jt}$ is the perceived "absolute" quality of restaurant j at time t; that is, it is unadjusted by price. p_{jt} is the price restaurant j charges at time t. e_{ijt} is an idiosyncratic taste shock of consumer i for restaurant j at time t. It follows an extreme value type 1 distribution with a location parameter 0 and scale parameter 1. It is identically independently distributed (i.i.d.) across i, j and t.

At both the new and established stages of a restaurant's life, the perceived quality is consistent across all consumers. At these stages, consumers will choose a restaurant with the probability $\frac{\exp(\hat{\hat{q}}_{jt}-p_{jt})}{\sum_{l\in\mathbf{J}}\exp(\hat{\hat{q}}_{lt}-p_{lt})+1}$. At the young stage, we make the simplifying assumption that in aggregate, the probability has the same logit form as those from the new and established stages with the perceived quality being at the population average. Under this assumption, the aggregate demand for restaurant j at any stage is

$$d_{jt} = \frac{M\exp(\hat{\hat{q}}_{jt} - p_{jt})}{\sum_{l\in\mathbf{J}}\exp(\hat{\hat{q}}_{lt} - p_{lt}) + 1}, \tag{G.2}$$

where d_{jt} denotes the aggregate demand, and $\mathbf{J}$ is the total set of restaurants in the market.

Given this demand, a firm will set its profit maximizing prices. It is easy to derive that the optimal price has the following form: $p_{jt} = c_j + \frac{1}{(1-s_{jt})}$, where c_j is the marginal cost of restaurant j, and $s_{jt} = \frac{\exp(\hat{\hat{q}}_{jt}-p_{jt})}{\sum_{l\in\mathbf{J}}\exp(\hat{\hat{q}}_{lt}-p_{lt})+1}$ is the market share of restaurant j. Because the market has many small firms, each restaurant's market share is very small, and therefore, $1 - s_{jt} \approx 1$. We then have the constant markup result, which yields a constant price $p_{jt} = c_j + 1$. This result is entirely reasonable, as competition with many small and differentiated firms is a monopolistic competition, and firms have limited pricing power (Benassy, 1996).

With the optimal price in hand, we can write the firm's profit as

$$\pi(\cdot) = M\frac{\exp(\hat{\hat{q}}_{jt} - c_j - 1)}{\sum_{l\in J}\exp(\hat{\hat{q}}_{lt} - c_l - 1) + 1} - C(\tilde{q}_j, D), \tag{G.3}$$

where $C(\tilde{q}_j, D)$ is the per-period fixed cost of restaurant j, and $\tilde{q}_j$ is the true "absolute" quality of restaurant j.

By assuming a constant marginal cost across firms, $c_j = c, \forall j \in \mathbf{J}$, and by setting $\hat{q}_{jt} = \hat{\hat{q}}_{jt} - c - 1$, where $\hat{q}_{jt} \in \{\hat{q}_0, \hat{q}_1, q\}$ and $q = \tilde{q}_j - c - 1$, we have the final profit function:

$$\pi(\cdot) = \frac{M\exp(\hat{q})}{\sum_{\hat{q}'} n(\hat{q}')\exp(\hat{q}') + 1} - C(q, D).$$