THREE ESSAYS IN INDUSTRIAL ORGANIZATION

by

Limin Fang

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Economics
University of Toronto

© Copyright 2018 by Limin Fang
Abstract

Three Essays in Industrial Organization

Limin Fang
Doctor of Philosophy
Graduate Department of Economics
University of Toronto
2018

My thesis includes three chapters that examine the dynamics of competition in the retail industry. Chapter One of the thesis demonstrates that online review platforms help consumers learn faster about product quality and improve consumer welfare as a result. To examine this, I use a novel dataset containing the universe of full-service restaurants in Texas, consumer search interest on major online review platforms and their online review information. I illustrate that online review platforms’ effects on learning show up in restaurant revenues and survival probabilities. Specifically, doubling consumers’ exposure to Yelp, the dominant platform, increases the revenue of a high-quality new independent restaurant by 8-20% and decreases that of a low-quality restaurant by about the same amount. Doubling Yelp exposure also raises the survival rate of a new high-quality independent restaurant by 7-19 basis points and reduces that of a low-quality restaurant by a similar level. Other platforms, especially Google, have similar effects but in smaller magnitude. In contrast, online platforms do not affect the revenues or survival rates of chains and old independent restaurants. Counterfactual analysis based on a structural demand model with consumer learning shows that online review platforms speed up the learning process by 0.5 to 2.5 years, increase consumer welfare by 5.4% and the total industry revenue by 5.9% during the period of 2005-2015.

Chapters Two and Three deal with the entry and exit dynamics of the retail industry, in particular, how chain retailers can pre-empt the entry of competitors by densely packing a geographic area with their outlets. Chapter Two develops a measure for preemptive motives in dynamic oligopoly games with entry and exit, and applies the measure in both theoretical and empirical studies. Chapter Three uses this measure in the fast casual dining industry in Texas to investigate if there is a trade-off between preemptive entry and survival of firms. The results show that under some conditions, preemption in fact helps the incumbent firm survive, while in other cases, preemption harms survival.
Contents

1 Evaluating the Effects of Online Review Platforms on Restaurant Revenues, Consumer Learning and Welfare
   1.1 Introduction .................................................. 2
   1.2 A Demand Model With Consumer Learning ...................... 5
      1.2.1 Learning About Restaurant Attributes ..................... 6
      1.2.2 Implications of the Learning Model ....................... 12
   1.3 Data ............................................................ 14
      1.3.1 Data Source and Measures .................................. 14
      1.3.2 Summary Statistics and Measures .......................... 17
   1.4 Reduced-Form Evidence ........................................ 20
      1.4.1 Revenue-Age Profiles ...................................... 22
      1.4.2 Bifurcating Effect of Exposure to Online Review Platforms 29
   1.5 Structural Estimation .......................................... 41
      1.5.1 Structural Model ............................................ 41
      1.5.2 Identification, Estimation Strategy and Structural Estimates 43
      1.5.3 Structural Estimates ........................................ 44
      1.5.4 Counterfactual Analysis .................................... 48
   1.6 Conclusion ...................................................... 53
   1.7 Indirect Utility in Constant Expenditure Model ................ 58
   1.8 Data Related to Market Demands and Costs ..................... 59
   1.9 Heckman’s Correction and Other Robustness Checks ............ 60
   1.10 Placebo Test .................................................. 60
      1.10.1 Other Channels of Yelp Effect Besides Learning ........... 61
   1.11 Identification of the Structural Model ......................... 67
   1.12 Derivation of the Likelihood Function ......................... 68
   1.13 Technical Issues Regarding Welfare Calculation .............. 70
   1.14 Tables ......................................................... 72

2 Measuring Preemption in Dynamic Oligopoly Games ............... 76
   2.1 Introduction .................................................. 76
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2 Model</td>
<td>80</td>
</tr>
<tr>
<td>2.2.1 Static Competition</td>
<td>80</td>
</tr>
<tr>
<td>2.2.2 Action Set and Markov States</td>
<td>80</td>
</tr>
<tr>
<td>2.2.3 Timing</td>
<td>80</td>
</tr>
<tr>
<td>2.2.4 Entry costs and scrap values</td>
<td>81</td>
</tr>
<tr>
<td>2.2.5 Flow payoff functions</td>
<td>81</td>
</tr>
<tr>
<td>2.2.6 Assumptions</td>
<td>82</td>
</tr>
<tr>
<td>2.2.7 Strategies and Conditional Choice Probabilities</td>
<td>82</td>
</tr>
<tr>
<td>2.2.8 Firms' Problems</td>
<td>83</td>
</tr>
<tr>
<td>2.2.9 Markov Perfect Equilibria</td>
<td>83</td>
</tr>
<tr>
<td>2.2.10 MPE in Probability Space</td>
<td>84</td>
</tr>
<tr>
<td>2.3 Definition of Preemptive Motives in a Simplified Model</td>
<td>85</td>
</tr>
<tr>
<td>2.3.1 Decomposition and Definition</td>
<td>85</td>
</tr>
<tr>
<td>2.3.2 Measuring the Contribution of Preemption to Firms' Behaviors</td>
<td>89</td>
</tr>
<tr>
<td>2.3.3 Extension to Models with Stochastic Demand, Multiple Firms and Multiple Locations</td>
<td>93</td>
</tr>
<tr>
<td>2.4 A Numerical Application</td>
<td>94</td>
</tr>
<tr>
<td>2.4.1 Parameters of the Simplified Model</td>
<td>95</td>
</tr>
<tr>
<td>2.4.2 How Sunk Costs Influence Preemption</td>
<td>95</td>
</tr>
<tr>
<td>2.5 An Empirical Application: Preemptive Race in Canadian Burger Industry</td>
<td>104</td>
</tr>
<tr>
<td>2.5.1 Data and Summary Statistics</td>
<td>104</td>
</tr>
<tr>
<td>2.5.2 Model Specification</td>
<td>107</td>
</tr>
<tr>
<td>2.5.3 Identification and Estimation</td>
<td>108</td>
</tr>
<tr>
<td>2.5.4 Effect of Preemption by Market Size</td>
<td>111</td>
</tr>
<tr>
<td>2.6 Comparison to Other Definitions in the Literature</td>
<td>115</td>
</tr>
<tr>
<td>2.7 Conclusion</td>
<td>118</td>
</tr>
<tr>
<td>1.1 Homotopy Method</td>
<td>123</td>
</tr>
<tr>
<td>1.2 Some Counterfactuals Are MPEs when ( EC = SV )</td>
<td>124</td>
</tr>
<tr>
<td>1.2.1 Proof of Proposition 1</td>
<td>125</td>
</tr>
<tr>
<td>1.2.2 Proof of Proposition 2</td>
<td>126</td>
</tr>
<tr>
<td>1.3 Extension of Definition to General Model</td>
<td>128</td>
</tr>
<tr>
<td>1.4 Extension of Definition to a Model With Two Locations</td>
<td>130</td>
</tr>
<tr>
<td>1.4.1 Model</td>
<td>130</td>
</tr>
<tr>
<td>1.4.2 Decomposition and Definition</td>
<td>132</td>
</tr>
<tr>
<td>1.4.3 Counterfactual</td>
<td>135</td>
</tr>
<tr>
<td>3 Aggressive Growth in Retail: A Trade-off Between Deterrence and Survival?</td>
<td>137</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>138</td>
</tr>
<tr>
<td>3.2 Model</td>
<td>140</td>
</tr>
<tr>
<td>3.2.1 Actions and Payoffs</td>
<td>140</td>
</tr>
</tbody>
</table>
Chapter 1

Evaluating the Effects of Online Review Platforms on Restaurant Revenues, Consumer Learning and Welfare
1.1 Introduction

Over the past two decades, online review platforms, such as Yelp, TripAdvisor, Amazon, and Google, have provided consumers with an enormous amount of information on products. As The Economist put it, information and data have surpassed oil and become the “world’s most valuable resource” (The Economist, 2017). Online review platforms have enabled consumers to learn from each other’s experiences, have reduced the information asymmetry between consumers and firms, and have helped consumers make better decisions when choosing products. Given their increasing role in the economy, it is important to understand the economics of online review platforms; in particular, the value of public goods they generate for consumers. A good understanding of this issue provides the basis for government policies and regulations on online platforms in the digital age.

This paper quantifies the welfare effect of online review platforms. In particular, I examine the effects of major online review platforms on revenues and consumer welfare in the restaurant industry in Texas. These platforms include Yelp, TripAdvisor and Google. I ask two research questions: (1) Do online review platforms help consumers learn faster about restaurant quality? (2) If they do, what are their effects on consumer welfare and industry dynamics? To answer these questions, I collected a novel dataset containing the universe of full-service restaurants in Texas, search interest on Yelp and TripAdvisor, and online review information from all three platforms. Using these data, I first demonstrate that consumer learning is important in the restaurant industry, and then test if online review platforms speed up this learning process. Next I build a structural demand model with social learning to quantify the effects of online reviews on consumer welfare and industry dynamics. I find that by helping consumers discover the underlying true quality of a restaurant quicker, online review platforms increase both the revenues and survival probabilities of high-quality restaurants and have the reverse effects on low-quality restaurants. Interestingly, these effects hold only for young independent restaurants, not chain or old established restaurants. My analysis from the structural estimation shows that by helping consumers choose better restaurants, online review platforms improve both consumer welfare and industry revenues. Of all the main platforms under examination, Yelp’s effect is larger and more salient compared to those from others, confirming the general observation that Yelp is the dominant review platform in the restaurant industry.

To examine if consumer learning is present in the restaurant industry, I construct restaurant revenue-age profiles. These profiles show how revenues evolve with restaurant age. I find that the age profile of higher-quality restaurants first slopes upwards, whereas the profile for lower-quality restaurants slopes downwards. These patterns are consistent with the learning theory in which consumers discover gradually the unknown quality of a restaurant over time.

To test if online review platforms speed up learning, I then investigate their effects on the slopes of the revenue-age profiles. In particular, I employ a triple difference approach that uses variation

\footnote{For example, the recent European Union’s ruling against Google with a $2.7 billion fine for malpractice in displaying search results has marked a new era of regulatory intent for digital platforms, and the increasing debate on regulating digital platforms across the globe calls for a better understanding of the economics of online platforms. See European Parliament briefing Online Platforms: How to Adapt Regulatory Framework to the Digital Age? \url{http://www.europarl.europa.eu/RegData/etudes/BRIE/2017/607333/IPOL_BRI(2017)607333_EN.pdf}}
in restaurant quality, restaurant type (age and chain affiliation), and online platforms' penetration by geographic region to examine their effects on revenues and survival probabilities. I first examine Yelp’s effect alone on restaurants listed on Yelp, then extend the analysis to include the effect of other platforms on restaurants that are listed on all three platforms. I find that doubling Yelp exposure (as measured by search interest on Google) increases the revenue of a high-quality new independent restaurant by 8-20% and decreases that of a low-quality restaurant by about the same amount. Doubling Yelp exposure also raises the survival rate of a new high-quality independent restaurant by 7-19 basis points and reduces that of a low-quality restaurant by a similar level. I call this differential effect on restaurants of various quality levels a “bifurcating” effect. Other platforms also have a bifurcating effect on revenue, but smaller in magnitude. In contrast, none of the platforms affect the revenues or survival rates of chains and old established independent restaurants. All of these results are consistent with the implications of the theory that online review platforms speed up consumer learning. An interesting result from this analysis is also that Yelp’s effect seems to overpower other platforms’ effects. For restaurants listed on Yelp, Yelp’s penetration is the only force that has a significant effect on revenue, regardless of whether the restaurants are also listed on other platforms.

In this reduced-form empirical analysis, I control for a rich set of fixed effects (FEs): calendar time FEs, calendar time×chain FEs, restaurant FEs and metro×month FEs. These fixed effects account for various unobserved variables that affect revenues and may be correlated with the measure of the penetration of online platforms. Thus, my key identifying assumption is that during the diffusion process of these main platforms, there were no other time-varying region-specific unobserved factors that increased the revenues of high-quality young independent restaurants, but reduced those of low-quality restaurants, and at the same time, had no effect on chain or old independent restaurants. To eliminate the concern of potential endogeneity coming from inherent difference in learning speed across geographic regions, I conduct a Placebo test to see if online platforms’ penetration would have an effect on restaurant revenues that predate the founding of online review platforms. The test shows that that source of endogeneity is not important.

Based on this reduced-form evidence, I develop a structural demand model with consumer learning to quantify the effects of online review platforms on the speed of learning, market shares and consumer welfare. My structural model combines the aggregate learning model from Ching (2010b) and the constant expenditure demand model with differentiated products from Björnerstedt and Verboven (2016). I further extend these models by incorporating social learning from both online platforms and other information sources. Using the estimated model, I evaluate online platforms’ effect on industry dynamics and consumer welfare through a counterfactual analysis, in which I remove the existence of online reviews. By comparing how long it takes consumers’ beliefs to reach a steady level in the counterfactual world and that in the real world, I find that learning slows down by about 0.5 to 2.5 years in the counterfactual world. The counterfactual analysis also shows that online review platforms redistribute demand from low-quality restaurants to high-quality restaurants, and in doing so, they improve consumer welfare by $15.2 million, which is equivalent
to a 5.4% increase in consumer’s monthly budget on eating out during 2005-2015. This percentage roughly translates into an $11 coupon on eating out per consumer per month. Furthermore, online review platforms also increase the overall industry revenue by $23.2 million, which constitutes a 5.9% gain over the same period.

Recent research has examined the impact of online reviews through their effects on sales, learning, competition and consumer welfare. For example, Chevalier and Mayzlin (2006) find that higher ratings online lead to an increase in book sales. Zhu and Zhang (2010) investigate if the effect of online reviews on sales is influenced by product and consumer characteristics, and find that online reviews have a greater impact on the sales of less popular video games and those games whose players are more experienced internet users. Zhao et al. (2013) use a structural learning model to study the effect of online product reviews on consumer purchases of books and find that consumers learn more from online reviews than from their own experiences of reading similar books in the past. Gutt et al. (2016) explore the effects of competition on the distribution of online reviews. My paper adds to this literature by showing that consumers’ exposure to online reviews has an effect on restaurant sales.

This paper is most closely related to Luca (2016)\(^2\). Luca employs a regression discontinuity design (RDD) to evaluate the causal effect of Yelp rating changes on restaurant revenues. He finds that an increase in rating by one star causes the revenues of independent restaurants to increase by 5%-9%, but does not affect those of chain restaurants. Luca also discusses if the way consumers use reviews is consistent with Bayesian learning. Compared to Luca’s work, my paper differs in a number of aspects: (1) First, my paper examines not only Yelp but also other online review platforms, including TripAdvisor and Google; (2) Second, my paper focuses on the question of whether online platforms speed up consumer learning by examining the bifurcating effects of their penetration on restaurants’ revenues, while Luca focuses on the effect of change in online displayed ratings on revenue without emphasizing the channel of the effects; (3) Third, I look at online review platforms’ effects not only on revenue, but also on restaurants’ survival probabilities; (4) Last, but not the least, I develop a structural model to quantify the effect of online review platforms on consumer welfare and industry dynamics.

This paper is also closely connected to Wu et al. (2015) and Lewis and Zervas (2016). Wu et al. (2015) examine the economic value of online reviews for consumers and restaurants using data on web browsing and reviews for seven restaurants on Dianping.com, a leading Chinese website with user-generated reviews. Lewis and Zervas (2016) assess the welfare effects of online reviews in the hotel industry using data on hotel revenues and reviews from a number of crowd-sourcing websites including TripAdvisor and Expedia. Compared to their work, my paper estimates a social learning model where both learning from online reviews and learning from other information sources are considered. This allows me to design a counterfactual that more accurately captures the welfare effect of online reviews. More specifically, my paper assumes that without online reviews, consumers

\(^2\)Anderson and Magruder (2012) examine the effect of Yelp ratings on customer flows also using a RDD approach. It is very closely related to Luca (2016), but different in that it does not use revenue data and does not investigate the aspect of consumer learning.
can still learn from other information sources and discover the true quality of products eventually, whereas the previous work assumes that consumers would never know the true quality of a product without online reviews. By assuming there is no learning without online reviews, the authors are likely to overestimate the welfare effects.


In the context of the literature, the contribution of my paper is three-fold: (1) First, this paper is the first that quantifies the welfare value of online platforms through the perspective of social learning. By incorporating learning from both online reviews and other information sources, this paper captures an important mechanism through which online review platforms improve welfare. (2) Second, my paper is the first to document a number of empirical patterns, including the effect of online reviews on the exit behaviors of firms and the bi-furcating effects on products of various quality levels; (3) Third, I develop a novel structural demand model with social learning that utilizes the aggregate revenue data at the product level. This combination of a social learning model and a constant expenditure model has not been done before. Furthermore, my learning model also has the novel feature of differentiating learning from online reviews from learning from other information sources.

The paper proceeds as follows. Section 3.2 introduces a demand model of differentiated products with consumer learning and describes the implications of online review platforms in the context of the model. Section 3.4 describes the data and discusses the main variables and preliminary patterns evident in the data. Section 3.6 discusses the empirical strategy for the reduced-form analysis and presents the results. Section 3.5 describes the structural model, its estimation strategy and the counterfactual analysis used to quantify the effects of online review platforms on consumer welfare and industry dynamics. Finally, Section 3.7 concludes by summarizing the key results and discussing the limitations and possible future extensions.

### 1.2 A Demand Model With Consumer Learning

This section describes a demand model with social learning. The social learning model is based on Ching (2010b), and the demand model is a constant expenditure model based on Björnerstedt and Verboven (2016). Both models lend themselves very well to modelling consumer choices over differentiated products by using aggregate product level data. The constant expenditure model is especially convenient for expressing consumer choices in the form of revenues. Given that the main
The data source in this paper is in the form of restaurant revenue, the constant expenditure model is convenient for my application.

In the social learning model, I assume that every restaurant’s quality is normally distributed around a mean. The quality can vary across time or consumers, but the mean quality stays constant. Consumers know that restaurant’s quality follows a normal distribution, but does not know the true mean quality of a new restaurant, and would need to learn about this mean quality from their own experiences and other people’s experiences. They first form a prior belief of a restaurant’s true mean quality, and then update these beliefs using what they have learned from their own and others’ experiences. Overtime, they find out the true mean quality of the restaurant. In this model, consumers’ belief updating process is a Bayesian updating process. Their own experiences and experiences from others are quality signals that are drawn from the normal distribution of the true quality. They use these signals to update their beliefs. Information shared on Yelp and other review platforms can be seen as experience signals shared by many consumers, and the more popular online review platforms are, the more experience signals are shared among consumers, and the faster consumers’ belief updating, and the quick the diffusion of the information about a restaurant’s true mean quality.

As consumers update their beliefs about restaurants’ quality, they make decisions on which restaurant to go to based on their new beliefs. As a result, restaurants that have a high true mean quality will experience an increase in their revenue, and those with a low true mean quality will see their revenues decline. These are the implications of the learning model and are testable using restaurant revenue data. Subsection 1.2.1 below describes the demand model with social learning, and Subsection 1.2.2 discusses the testable implications of the model.

1.2.1 Learning About Restaurant Attributes

In this demand model with learning, I assume that consumers get to decide how much they want to eat at a restaurant once they have chosen the restaurant. This assumption is the key difference between the constant expenditure demand model and the standard logit demand model, where the demand is assumed to be unit demand. This assumption of flexible quantity gives the constant expenditure the feature that consumers always spend a constant share of their income on eating out. In this model, consumers choose a restaurant based on the indirect utility that they receive at a restaurant. The indirect utility of consumer $i$ from visiting restaurant $j$ depends on her income, demographic characteristics, other market and restaurant characteristics, and the price. It has the following functional form:

$$U_{ijt} = X_{jt}'\theta + \alpha^{-1}\log(y_i) - \alpha\log(p_{jt}) + \tilde{A}_{ijt} + \xi_{jt} + \epsilon_{ijt}$$ (1.2.1)

---

3Although the price data are available for restaurants listed on Yelp, TripAdvisor and Google, these prices are in broad ranges. Dividing restaurant revenues by these gross price measures does not give a very good approximation for quantity. Therefore, in this application, I opt away from the standard logit demand model that requires quantity data.
where $X_{jt}$ includes restaurant characteristics that are observable to consumers initially and do not require learning as well as market characteristics that capture demand factors such as demographic attributes. $\theta_x$ are coefficients associated with $X_{jt}$. $\log(y_i)$ is the natural log of consumer $i$’s income; $\gamma$ is the budget share of consumer $i$’s income spent on eating at restaurants; $\log(p_{jt})$ is the natural log of the average price of a meal at restaurant $j$, and $\alpha$ is the price coefficient. $\tilde{A}_{ijt}$ is the realized experience of consumer $i$ from dining at restaurant $j$; it is a random draw from a restaurant’s quality distribution. $\xi_{jt}$ is an aggregate demand shock for restaurant $j$ at time $t$. It is observed by consumers, but not observed by the econometrician; $e_{ijt}$ is an idiosyncratic taste shock that is i.i.d. across consumers, restaurants and time, and it follows the extreme value type I distribution. The key feature of a constant expenditure model in terms of the functional form is that prices come into consumers’ indirect utility functions in log forms instead of the linear form. A detailed derivation of this indirect utility function can be found in Section 1.1 of the Appendix.

In this utility function, there are both horizontal differentiation and vertical differentiation. I assume that horizontal differentiation among consumers’ tastes is reflected mostly in the initially observable characteristics of restaurants, in particular, the cuisine types. For example, Asians may prefer a sushi restaurant to a gastro pub, while white Americans may prefer the latter. To capture the horizontal differentiation, I include in $X_{jt}$ the interaction terms of consumer demographics and cuisine types. Conditional on these observable characteristics, I assume that restaurants are vertically differentiated in the quality $\tilde{A}_{ijt}$, the mean of which requires learning. For example, there are two seemingly similar Italian restaurants. One uses fresher ingredients and has a better chef compared to the other one. This quality difference is unknown to consumers until consumers have tried the restaurants themselves or heard recommendations from others through word-of-mouth communication. This quality requires learning and is vertically differentiated since most people would agree on the first Italian restaurant being better than the second one.

As mentioned before, $\tilde{A}_{ijt}$ is the realized experience of consumer $i$ from dining at restaurant $j$. It is a random draw from restaurant $j$’s quality distribution. In particular, it can be written as

$$\tilde{A}_{ijt} = A_j + \delta_{ijt} \quad (1.2.2)$$

where $A_j$ is the true mean quality of the restaurant $j$. $\delta_{ijt}$ is random deviation from the mean, and is assumed to follow a normal distribution $\delta_{ijt} \sim N(0,\sigma^2_{\delta})$, and distributed i.i.d. across consumers, restaurants and time.

The reason why we do not assume a constant quality of a restaurant is that there is inherent variability in products being offered and also in consumers’ experiences of the product from time to time. For example, restaurant services could vary from one visit to the next, and consumers’ satisfaction over food could also change across visits. In this regard, $\sigma^2_{\delta}$ represents the experience variability.

Consumers know the value of $\sigma^2_{\delta}$, but do not know the mean $A_j$, and therefore, they use experience signals they receive from their own visits and others’ experiences shared through online platforms or face-to-face communication to update their prior beliefs about the restaurant’s true
mean quality. This updating process follows Bayes’ rule, which incorporates the values of the signals and the number of signals. More specifically, let consumers’ initial prior of $A_j$ be a normal distribution with mean $A$ and variance $\sigma_A^2$. The Bayes’ rule implies that the updated belief at every period after incorporating new information from the experience signals, i.e. the posterior belief, also has a normal distribution. Let $\eta_{jt+1}$ denote the mean of the posterior distribution and $\sigma_{jt+1}^2$ the variance of the distribution. Bayes’ rule stipulates that $\eta_{jt+1}$ and $\sigma_{jt+1}^2$ have the following recursive updating form (DeGroot, 1970):

$$\eta_{jt+1} = (1 - \beta_{jt})\eta_{jt} + \beta_{jt}\bar{A}_{jt}$$  \hspace{1cm} (1.2.3)

where $\eta_{jt}$ is the prior mean and $\beta_{jt}$ is

$$\beta_{jt} = \frac{1}{\sigma_{jt}^2 + \frac{1}{n_{jt}\sigma_j^2}}$$  \hspace{1cm} (1.2.4)

in which $n_{jt}$ is the number of experience signals that are released to consumers through their own trials or social media at time $t$, and $\bar{A}_{jt}$ is the sample mean of these experience signals. It follows a conditional normal distribution: $\bar{A}_{jt}|n_{jt} \sim N(A_j, \sigma_j^2/n_{jt})$. The precision or the inverse of the variance of the posterior is

$$\frac{1}{\sigma_{jt+1}^2} = \frac{1}{\sigma_{jt}^2} + \frac{n_{jt}}{\sigma_j^2}$$  \hspace{1cm} (1.2.5)

where $\sigma_j^2$ is the prior variance.

Note that in the above system of equations, I omit the notation $i$. This is due to two reasons: one is that $A_j$ is vertically differentiated and is the same to everyone; second, I assume that everyone in a market is exposed to the same set of experience signals, and therefore, the society in a market updates the prior belief as a whole. Since the mechanism of learning here is social learning, where the knowledge from one person gets disseminated quickly to the rest of the society, this assumption is reasonable. Furthermore, since the society updates beliefs as a whole, this social learning process can also be seen as the diffusion of knowledge of restaurants throughout the society.

As can be seen from equations 1.2.3 to 1.2.5, the posterior distribution of beliefs depends on the average of the values of the experience signals that are released to consumers in period $t$, i.e. $\bar{A}_{jt}$, and also the number of experience signals, $n_{jt}$. Although consumers observe these signals when updating their beliefs, I as the researcher do not observe either $\bar{A}_{jt}$ or $n_{jt}$ in my data. Therefore, I model these variables to capture the learning and belief updating process. First, I assume that the number of signals $n_{jt}$ is proportional to consumers’ exposure to online review platforms in a market and the restaurant’s past quantity, which represents roughly how many consumers have visited the restaurant in the past period. It has the following functional form:

$$n_{jt} = (\kappa + \lambda Yelp_{mt})q_{jt}$$  \hspace{1cm} (1.2.6)
where Yelp represents how much consumers are exposed to online review platforms in market \( m \) at time \( t \). Note that Yelp as a notation here is only for simplicity purposes. It represents not only Yelp but also other online review platforms. \( \kappa \) captures the portion of information that consumers receive from other sources such as their own trials, interactions with friends or other media platforms. \( \lambda \) is simply the coefficient for Yelp. Both \( \kappa \) and \( \lambda \) are assumed to be positive.

If we take a closer look at the mean updating equation \([1.2.3]\), we can see that the belief updating process is very intuitive. The posterior mean \( \eta_{jt+1} \) is simply a weighted average of the prior mean \( \eta_{jt} \) and the average of the new signals. This weight is \( \beta_{jt} \), which depends on the precision of the prior and that of new signals. If at time \( t \), the prior \( \eta_{jt} \) is already very precise, then \( \frac{1}{\sigma_{jt}^2} \) would be large, and the updating process would put more weight on the prior and less on the new signals \( \bar{A}_{jt} \). If, however, the prior is not precise and \( \frac{1}{\sigma_{jt}^2} \) is small, which implies a precise average of the new signal \( \bar{A}_{jt} \), then the updating process would place a greater weight on \( \bar{A}_{jt} \) and less on the prior. Typically, as \( t \) gets large, the prior would become precise; that is, \( \frac{1}{\sigma_{jt}^2} \) grows larger; as a result, new signals would contribute less and less to consumers’ beliefs.

In this model, I make the assumption that consumers only learn about independent restaurants’ quality, not that of chain restaurants. This assumption is very reasonable given that chain restaurants benefit from centralized advertising and multiple locations with a uniform menu, pricing and service. As will be shown later, this assumption generates testable implications that help reinforce the learning story in the data generating process.

Given the assumptions and the learning model described above, we can write down consumers’ demand for each restaurant. Consumers choose which restaurant to go to based on their prior beliefs. Since this belief has a distribution, consumers form an expectation of the indirect utility from visiting a restaurant, and then choose the restaurant with the maximum expected indirect utility. Taking the expectation of \( \bar{A}_{ijt} \) in equation \([1.2.1]\) conditional on the information set \( I(t) \) at time \( t \) gives

\[
E[U_{ijt}|I(t)] = X_{jt}\theta_x + \alpha\gamma^{-1}\log(y_i) - \alpha \log(p_{jt}) + \eta_{jt} + \xi_{jt} + e_{ijt}, \forall j \in J_I
\]  

for independent restaurants, where \( J_I \) denotes the set of independent restaurants in a market. For chain restaurants, since there is no learning, we have

\[
U_{ijt} = X_{jt}\theta_x + \alpha\gamma^{-1}\log(y_i) - \alpha \log(p_{jt}) + A_j + \xi_{jt} + e_{ijt}, \forall j \in J_C
\]  

where \( J_C \) is the set of chain restaurants in a market. In additional to chain and independent restaurants, consumers also have an outside option, which includes restaurants or stores that offer food but not alcoholic drinks. The indirect utility associated with the outside option is normalized to

\[
U_{i0t} = \alpha\gamma^{-1}\log(y_i) - \theta_t + e_{i0t}
\]

where \( \theta_t \) represents changes in consumer taste towards eating out at full-service restaurants relative
to the outside option over time. Note that the term $\alpha \gamma^{-1} \log(y_i)$ shows up in the indirect utility in all restaurants in the choice set, and as will be shown later, it will cancel out in consumers’ choice probability function.

In this model, consumers are assumed to be myopic, and therefore I abstract away from consumers making a sacrifice today to try out some new restaurants so that they can gain better knowledge in restaurants’ quality and make a better decision tomorrow. This assumption of consumers being myopic is reasonable given that the gain from trying out a restaurant today versus tomorrow may not be very large, and this gain from trying restaurants out oneself is particularly small when there is a large scale of social learning, where consumers benefit from others’ experiences. Also this modelling assumption is consistent with that used in Zhao et al. (2013), who model consumer learning in the experienced goods industry. Furthermore, Erdem and Keane (1996) also did not find much difference in the estimates of structural parameter and model predictions between models with myopic consumers and forward-looking consumers in the laundry detergent product category. Building on the findings from the exiting literature, I also assume that consumers are myopic and only maximize their current utility. This assumption also simplifies the model and estimation substantially.

Given the setup of consumers’ expected indirect utility function, we can write a consumer’s probability of choosing a restaurant $j$ as

$$s_{ijt} = \frac{\exp(\Delta_{ijt} - \Delta_{i0t})}{1 + \sum_j \exp(\Delta_{ijt} - \Delta_{i0t})}$$

where $\Delta_{ijt}$ denote the mean utility from dining at a restaurant $j$ at time $t$; it is $E[U_{ijt}|I(t)] - e_{ijt}$ for independent restaurants and $U_{ijt} - e_{ijt}$ for chain restaurants. Since $\gamma^{-1} \log(y_i)$ cancels out in this expression, we can remove the $i$ subscript from $s_{ijt}$. The log ratio of $s_{jt}$ over $s_{0t}$ is then

$$\log\left(\frac{s_{jt}}{s_{0t}}\right) = \Delta_{jt} - \Delta_{0t}$$

Note that this share $s_{jt}$ is different from the market share in the standard discrete choice models. It represents how many consumers will choose restaurant $j$ out of one unit mass of consumers, but does not represent the quantity sold at restaurant $j$ like in the standard discrete choice demand model. To obtain quantity, one can use Roy’s identity to derive a consumer’s demand for restaurant $j$, $d_j(y_i)$, conditional on her having chosen $j$:

$$d_j(y_i) = -(\partial E[U_{ijt}|I(t)]/\partial p_{jt})/(\partial E[U_{ijt}|I(t)]/\partial y_i)$$

$$= \gamma \frac{y_i}{p_{jt}}$$

where $y_i$ is consumer $i$’s income. By integrating over the income distribution in a market and multiplying the average individual demand by the market size, we have that the aggregate demand
$q_{jt}$ is

\[
q_{jt} = \int s_{jt} d_j(y) dF_y(y) M_{mt}
= s_{jt} \int d_j(y) dF_y(y) M_{mt}
= s_{jt} \gamma Y_{mt} / p_{jt}
\]

(1.2.12)

where $M_{mt}$ is the size of the population in market $m$ at time $t$; $F_y(y)$ is the income distribution of the population in the market; $Y_{mt}$ is the total income of consumers in market $m$ at time $t$, and $\gamma$, as mentioned before, is the budget share of consumers’ income spent on eating out. This equation implies that $s_{jt}$ can be written as a function of the revenue:

\[
s_{jt} = \frac{p_{jt} q_{jt}}{\gamma Y_{mt}}
= \frac{\text{Rev}_{jt}}{\gamma Y_{mt}}
\]

(1.2.13)

where $\text{Rev}_{jt}$ is the revenues of independent restaurant $j$ at time $t$. Substituting equation (1.2.13) into equation (1.2.10) we have

\[
\log(\frac{\text{Rev}_{jt}}{\text{Rev}_{0t}}) = \log(\frac{s_{jt}}{s_{0t}})
= X_{jt} \theta_x - \alpha \log(p_{jt}) + \eta_{jt} + \theta_t + \xi_{jt}
\]

(1.2.14)

This equation relates restaurant revenues to consumers’ mean utility. Since the revenue data I have includes only alcoholic drinks sales, $\text{Rev}_{ajt}$, I express it as a proportion $g_{jt}$ of the total restaurant revenue ($\text{Rev}_{jt}$) (henceforth, alcohol-to-total-revenue ratio)\textsuperscript{4} that is, $\text{Rev}_{ajt} = \text{Rev}_{jt} \cdot g_{jt}$. Equation (1.2.14) is then

\[
\log(\text{Rev}_{ajt}) = \log(g_{jt}) - \alpha \log(p_{jt}) + X_{jt} \theta_x + \eta_{jt} + \log(\text{Rev}_{0t}) + \theta_t + \xi_{jt}
\]

(1.2.15)

For simplicity of notation, I use $\text{Rev}_{jt}$ to represent the revenues from alcoholic beverages in the rest of the paper.

Equation (1.2.15) provides a baseline specification for both the reduced-form and the structural analyses. In the reduced-form analysis, instead of using $\text{Rev}_{0t}$, whose construction often requires strict assumptions, I use the number of rivals to approximate the effect of $\text{Rev}_{0t}$ in the revenue

\textsuperscript{4}Although this ratio changes over time, it fluctuates around a mean that depends mostly on a restaurant’s type. It is common knowledge in the restaurant industry that bars and bistros have a much higher alcohol-to-revenue ratio on average than a breakfast joint. This ratio also depends on the demographic characteristics of the market; for example, a bar in an area with lots of young college students is likely to have a higher alcohol ratio than one in an area with mostly young families. These variations in the alcohol-to-total-revenue ratio across restaurants can be captured by restaurants’ type and demographic characteristics. Therefore, although $g_{jt}$ is unobservable, I can restaurant fixed effects, which includes restaurant cuisine types, market and demographic characteristics and time fixed effects to control for this ratio in the empirical analysis. This will be discussed in greater detail in the results section.
function. In the structural estimation, I construct \( Rev_{0t} \) by collecting consumer expenditure data. The estimating equations and estimation strategy will be discussed in detail in Sections 3.6 and 1.5.

This demand model with social learning has a number of testable implications. In the following subsection, I discuss these implications by using a numerical illustrative example.

### 1.2.2 Implications of the Learning Model

The key evidence of learning in the model described above lies in how log revenues \( \log(Rev_{jt}) \) change with restaurants’ age. After controlling for all the demand shifters and competitive effects, log revenues will depend only on \( \eta_{jt} \). Therefore, it is sufficient to examine how \( \eta_{jt} \) evolves with age; in particular, how consumers’ exposure to online reviews (Yelp\(_{mt}\)) affects the path of \( \eta_{jt} \) over age. To illustrate this path, I use a simple numerical example with a set of predetermined structural parameters (see Table 1.1) and plot in Figure 1 \( \eta_{jt} \) against age for both chain restaurants and independent restaurants under the scenarios where consumers are exposed to online reviews (with Yelp\(_{mt} = 5\)) and where they are not (with Yelp\(_{mt} = 0\)).

In this numerical example, there are one chain restaurant, one high-quality independent restaurant, and one low-quality independent restaurant in the market. Their true mean qualities are 6, 5 and -5 respectively. Other parameters of the model are shown in Table 1.1. Although this numerical example is generated from this specific set of parameters, the insight it provides is very general and applies to a wide range of parameters.

#### Table 1.1: Structural Parameters for the Numerical Example

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain ( A_j )</td>
<td>6</td>
<td>( A )</td>
<td>0</td>
</tr>
<tr>
<td>high ( A_j ) (independent)</td>
<td>5</td>
<td>low ( A_j ) (independent)</td>
<td>-5</td>
</tr>
<tr>
<td>( \sigma^2_A )</td>
<td>1</td>
<td>( \sigma^2_A )</td>
<td>1</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>0.0025</td>
<td>( \lambda )</td>
<td>0.0005</td>
</tr>
<tr>
<td>( \xi_{jt} )</td>
<td>0</td>
<td>Yelp(_{mt} )</td>
<td>5</td>
</tr>
<tr>
<td>( \log(g_{jt}) - \alpha \log(p_{jt}) + X_{jt}\theta_x + \log(Rev_{0t}) )</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To obtain the evolutionary paths of \( \eta_{jt} \), I simulate the paths 100 times and take the average. Simulation is necessary because \( \eta_{jt} \) is random, and it depends on the past experience signals. These average paths are plotted in Figure 1 for the age between 0 and 60. Given that these paths are effectively the profiles of log revenues over age conditional on the demand shifters and competition, I refer to these paths as "revenue-age profiles" of restaurants from now on.

As shown in the figure, the revenue-age profile of the chain restaurant (the dark line) is flat at its quality level 6 because there is no learning. Those of the high-quality independent restaurant (the red lines) start from the prior mean quality 0, and then slope up and converge eventually to

\[ \bar{A}_{jt} \]

As shown in equation 1.2.3, \( \bar{A}_{jt} \) is the average of the random experience signals that enters the updating process of consumers’ beliefs. Furthermore, \( \eta_{jt} \) is path-dependent in the sense that past experience signals can set a restaurant on to a different trajectory. To reduce the effect of the randomness in past experience signals, I take the average of the 100 simulations for each type of restaurant. The average can be seen as the path that \( \eta_{jt} \) is expected to follow.
Chapter 1

Figure 1: Revenue-Age Profiles: Numerical Example

The true mean quality at 5. The revenue-age profile under the scenario with online reviews (the dotted red line) has a steeper slope and converges earlier than that under the scenario without, indicating that online reviews help consumers learn faster about restaurant qualities. For the low-quality independent restaurant, the revenue-age profiles (the blue lines) slope downwards as their low qualities are discovered by consumers. Compared to those of the high-quality restaurants, the slopes of the revenue-age profiles of the low-quality restaurants are much more gentle, suggesting that learning is slower. This pattern of more gently sloped profiles is consistent with the learning model discussed above. Since lower-quality restaurants have a smaller demand, learning about their qualities is naturally slower because there are not as many consumers visiting these restaurants compared to the higher quality restaurants.

This figure illustrates a number of implications of the learning model, and these implications can be summarized as follows:

1. Revenue-age profiles of independent restaurants systematically slope upwards or downwards initially (depending on the quality) and then level off overtime.
2. If online reviews help consumers learn, they should have a negative effect on the slope of the revenue-age-profile of lower quality independent restaurants and a positive effect on those of higher quality, and the effects are asymmetric.
3. Online reviews’ effects depend on restaurants’ age. The effects are stronger in the earlier years of a restaurant’s life than later years.
4. Chain restaurants’ revenues are relatively stable over age.
5. Online reviews should not affect chain restaurants’ revenues much through learning.

In the next sections, I will discuss my main data source and measures for the various components in the learning model, and then test these implications using these data.

### 1.3 Data

This section describes the data and the measures for the variables in the learning model and provides descriptive statistics of the data. Section 1.3.1 discusses the data sources, and Section 1.3.2 illustrates the summary statistics of the data and measures for consumers’ exposure to online review platforms and those for restaurant quality.

#### 1.3.1 Data Source and Measures

My data come from a variety of sources: (1) The first one is the restaurant mixed-beverage revenue data from the Mixed Beverage Tax Information Records held by the Office of the Comptroller of Public Accounts in Texas; (2) The second set of data include demographics data from the census, American Community Survey, visitor and tourism spending data from Texas Economic Development and Tourism, average daily traffic flow data from the Texas Department of Transportation, consumer expenditure on eating out from the Consumer Expenditure Survey, and wage data from Occupational Employment Statistics; these data inform the market demand and cost characteristics. (3) The third data source is the Google Trends data on the search term Yelp and the website TripAdvisor; this data reflect the geographic penetration of Yelp and TripAdvisor over time and are used as the main measure for consumers’ exposure to online platforms. As mentioned previously, Yelp is the dominant platform in the restaurant industry. My paper focuses on Yelp’s effect on restaurants’ revenue and survival rates, and uses TripAdvisor as the main control for the penetration of other platforms. (4) The fourth data source is detailed restaurant level data collected from the Yelp, TripAdvisor and Google websites, including restaurant overall ratings, pricing, and restaurant cuisine category on each platform. For restaurants listed on Yelp especially, I collected the rating history with time stamp and star rating for each review for each restaurant on Yelp. Below I describe each data source in greater detail.

The first dataset, the restaurant revenue data, is the gross receipts of alcoholic drinks sold at each restaurant in Texas. The volumes of these gross receipts are reported to the State of Texas Comptroller Office on a monthly basis by mixed-beverage permit holders. This dataset started in 1993 and contains information on a restaurant’s monthly gross receipts from liquor, wine and beer beverages, restaurant address, owner names and addresses, business location name and address, permit number, permit issue date and out of business date. The dataset covers all establishments that hold (or held) a mixed-beverage permit, including short-term events, such as festivals, and non-restaurant entities like convention centers. To obtain a homogeneous set of sample, I select full-service restaurants based on their NAICS code and their names. I exclude primarily night life

---

6 Although the mixed-beverage receipt dataset does not include NAICS code, I merged it with another dataset,
oriented places such as night clubs or sport bars, and I also remove restaurants at major hotels and airports as they are often monopolies at their locations and are subject to a unique set of regulations. To avoid potential irregularities associated with the earlier effort at collecting this data, I set my sample period to January 1995 - December 2015. This elimination process leaves 15,417 restaurants in my dataset, about 37% of all establishments recorded in the Mixed Beverage Tax Information Records.

In my analysis, I choose the market definition as zipcode×month. The geographic area at the zipcode level is appropriate for modeling the competition between full-service restaurants given that most consumers for restaurants are local. To complement the restaurant dataset with market level information, I collected the second dataset related to market demand and costs. They include demographic characteristics, income, tourist spending, traffic volume data, food service industry wage data, and consumer expenditure data. A detailed account of these data sources and how I interpolate some data to each market is shown in Section 2 of the Appendix.

The third dataset, the Google Trends data, tracks the search interest on Google for the term “Yelp” and the TripAdvisor website at the metropolitan area level. The data is monthly, dates from January 2004 to November 2016, and spans across 20 metropolitan areas in Texas. For each metro area, the data is a time series of a normalized index ranging from 0 to 100. The normalization is carried out by dividing the search volume for the term “Yelp” or the TripAdvisor website for a given metro region at a particular time by the total searches for all terms in that geography at that time. This results in a proportion of the search volume for Yelp or TripAdvisor out of the total search volumes for a given month in a metro area. Among these proportions, the highest one for a metro area is normalized to 100. All the other proportions are scaled accordingly. This normalization is to control for the total population in each metro area, so that the metro area with the highest population does not always rank the first for consumers’ exposure to Yelp or TripAdvisor. For example, Austin ranks the highest by the normalized search interest for Yelp, but it is only the 4th largest metro region by population and the 3rd by the number of restaurants. Through this normalization, the Google Trends search interest can be interpreted as the attention paid to Yelp or TripAdvisor out of a person’s attention on all subjects on Google. This normalized search interest data provides an aggregate measure of consumers’ exposure to Yelp and TripAdvisor and reflects both demand and supply of information on these platforms. It also tracks the geographic penetration of these platforms over time, which provides an important source of variation for identifying the effect of these online platforms on restaurants’ revenues.

The fourth data source zooms in at the restaurant level. Of all restaurants included in my dataset, I identified those that are listed on Yelp, TripAdvisor and Google using Yelp API, Google API and automated Bing search on Microsoft Azure. As mentioned before, this information includes restaurant overall ratings, pricing and cuisine type on each platform. The data for Yelp and Google were collected in November 2016, and TripAdvisor data were collected in June 2018. The distribution

the Sales Permit dataset, which contains the NAICS code information.

Metropolitan areas are the finest geography for which the data is available. Although these metropolitan areas are delineated by Google, they are very close to the census definition of metropolitan statistical areas.
of restaurants across the platforms are illustrated in Figure 2. As shown in the figure, Google covers
the highest number of restaurants, and most restaurants that are listed on Yelp and TripAdvisor are
also listed on Google. Overall, there are 9,024 restaurants listed on Google, of which 5,307 are also
listed on Yelp and 4,726 on TripAdvisor, and 3,591 are listed on all three platforms. TripAdvisor
and Yelp alone have about 3,817 overlapping restaurants. In total, 9,900 restaurants have online
presence; they account for about 64.2% of all restaurants in the revenue dataset, and 96.5% of all
restaurants that were active as of December, 2015, the end of the sample period. Basically by the
end of the sample period, almost all restaurants had some type of online presence. Note that these
platforms retain a record of restaurants that are closed because they draw traffic to their websites.

All these platforms provide services to both businesses and consumers. For businesses, they offer
business analytics and advertising services, and for consumers, they perform not only as a business
directory but also a platform that aggregate consumers’ opinions on a business. For both Yelp and
TripAdvisor, anyone can create a profile for a business. For Yelp, business owners can then choose
to claim the business or not. TripAdvisor does not offer this "claim" option. For Google, since it
has the Google Maps GPS system, it covers almost all businesses. This explains why Google has the
highest coverage for restaurants in my dataset. All these platforms offer a discrete 1-5 star rating
system, and have prices shown in ranges, such as “$” or “$$”. There is also a description of the
restaurant cuisine type on each of the platforms. This detailed restaurant-level information allows
me to control for restaurant characteristics in my analysis.

For restaurants listed on Yelp, I also collected the rating history for each restaurant, including
the time stamp and star rating for each review of each restaurant. I use this data to complement
the previous data regarding online review platforms in a number of ways: (1) First, I use the rating
history to pinpoint the start date for Yelp’s penetration. As mentioned before, Yelp is the main
player of online review platforms in the restaurant industry, and the focus of the paper. Having
a precise information regarding its penetration is of particular interest. According to the rating
history data, the very first review on Yelp was written on March 29th, 2005 for an Italian restaurant
“Carrabba’s Italian Grill” in Plano, Texas. I use this date as the start of Yelp’s penetration in my
analysis. (2) Second, I deploy this rating history information to construct an additional measure
of Yelp’s penetration over time and geographic regions and compare that to the measure from the
Google Trends data. As will be discussed in the next subsection, this alternative measure gives a
more concrete interpretation of the normalized index provided by Google Trends, and supports the
validity of Google Trends’ normalized index as a measure for platforms’ geographic penetration.
(3) Third, I also employ this data to construct various measures of restaurant quality. Knowing
each review’s date and rating for each restaurant, I can calculate a restaurant’s average rating
at a particular time. This average rating can predate the revenue data. For example, I can use
restaurants’ first 10 reviews to construct a predetermined quality measure for each restaurant, and
then examine how exposure to Yelp affects their revenues after the first 10 reviews date. This
measure is more exogenous compared to the overall ratings displayed on each platform at the date
of data collection (i.e. November 2016 or June 2018), which postdates the sample period. Though
having the advantage of being a more exogenous measure, this “predetermined” measure limits the number of observations on restaurant revenues. It is used alongside the overall ratings collected in November 2016 and June 2018 as robustness checks. (4) Last, the rating history data also informs how restaurants’ quality might have evolved over time. It is used in the examination of other major channels through which these platforms may have an effect on restaurant revenues.

Combining all these sources, I obtain a panel that covers monthly mixed beverage sales for 15,417 restaurants during the period of January 1995 to December 2015, demographic and income information for each zipcode tabulation areas, traffic information based on each outlet’s location, the monthly Yelp popularity measure for each metro area in Texas, the number of monthly reviews and average rating for each Yelp listed restaurant, and the overall rating, price range and cuisine types for all restaurants that are listed on Yelp, Google and TripAdvisor. In total, the dataset includes 1.14 millions of observations in 846 geographic markets, which span over 482 cities and places, 113 counties and 20 metropolitan areas. The key features of this data and some stylized facts exhibited by this data are summarized in the next section.

1.3.2 Summary Statistics and Measures

This section discusses the summary statistics of the data described above, highlights some key features of the data, and talks about the measures for consumers’ exposure to online review platforms and restaurant quality. Table 1.2 summarizes the data. On average restaurants make $23,421 per month from mixed beverages. The highest amount of revenue is over $1.67 million. This indicates that there is a large variation in restaurant revenues. Some restaurants have survived a long time, as old as 22 years (or 265 months). There are also many new restaurants that just opened. The average age of a restaurant in my dataset is about 6 years (or 74 months). These restaurants spread over 163,774 zipcode×month markets in total.

At the market level, on average, about 2.32 chain and 4.61 independent restaurants exist in each
market. The average population in a market is over 26,000, and the maximum is about 118,000. The population in Texas is predominantly white between the age of 35 to 64. The average household income is about $56,000 in each market, and the highest is over $201,000. The average visitor spending per county and month is $31 million, and the highest and the lowest are $893 million and $0.07 million respectively. Traffic volume passing by a restaurant is an important demand indicator. The data show that daily traffic passing by a restaurant is about 76,000 vehicles on average and over 383,000 at the maximum.

Table 1.2: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>Number of Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly revenue ($)</td>
<td>$23,421</td>
<td>32,567</td>
<td>0</td>
<td>1,668,320</td>
<td>1,135,220</td>
</tr>
<tr>
<td>Number of chain restaurants per market</td>
<td>2.32</td>
<td>3.46</td>
<td>0</td>
<td>38</td>
<td>163,774</td>
</tr>
<tr>
<td>Number of independent restaurants per market</td>
<td>4.61</td>
<td>5.87</td>
<td>0</td>
<td>66</td>
<td>163,774</td>
</tr>
<tr>
<td>Total population per market</td>
<td>26,750</td>
<td>16,057</td>
<td>549</td>
<td>118,056</td>
<td>163,774</td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>0.30</td>
<td>0.081</td>
<td>0.047</td>
<td>0.99</td>
<td>163,774</td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>0.37</td>
<td>0.056</td>
<td>0.016</td>
<td>0.68</td>
<td>163,774</td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>0.10</td>
<td>0.046</td>
<td>0</td>
<td>0.58</td>
<td>163,774</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>0.33</td>
<td>0.25</td>
<td>0</td>
<td>1.0</td>
<td>163,774</td>
</tr>
<tr>
<td>Share of White population</td>
<td>0.81</td>
<td>0.18</td>
<td>0.045</td>
<td>1.0</td>
<td>163,774</td>
</tr>
<tr>
<td>Share of Black population</td>
<td>0.11</td>
<td>0.13</td>
<td>0</td>
<td>0.95</td>
<td>163,774</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>0.037</td>
<td>0.048</td>
<td>0</td>
<td>0.55</td>
<td>163,774</td>
</tr>
<tr>
<td>Average household income per market ($)</td>
<td>$56,303</td>
<td>24,554</td>
<td>11,906</td>
<td>201,255</td>
<td>163,774</td>
</tr>
<tr>
<td>Daily traffic counts</td>
<td>79,142</td>
<td>80,973</td>
<td>0</td>
<td>383,229</td>
<td>1,135,220</td>
</tr>
<tr>
<td>Visitor spending per county×month ($ millions)</td>
<td>36.6</td>
<td>108.0</td>
<td>0.3</td>
<td>893.2</td>
<td>23,374</td>
</tr>
<tr>
<td>Number of months a restaurant survived to</td>
<td>74.6</td>
<td>71.3</td>
<td>1</td>
<td>265</td>
<td>15,417</td>
</tr>
<tr>
<td>Yelp search interest</td>
<td>4.5</td>
<td>9.2</td>
<td>0</td>
<td>100</td>
<td>5,040</td>
</tr>
<tr>
<td>TripAdvisor search interest</td>
<td>10.8</td>
<td>15.3</td>
<td>0</td>
<td>100</td>
<td>5,040</td>
</tr>
<tr>
<td>Yelp Rating for chain as of November 2016</td>
<td>2.96</td>
<td>0.55</td>
<td>1</td>
<td>5</td>
<td>1,935</td>
</tr>
<tr>
<td>Yelp Rating for independent as of November 2016</td>
<td>3.52</td>
<td>0.59</td>
<td>1</td>
<td>5</td>
<td>3,995</td>
</tr>
<tr>
<td>Google Rating for chain as of November 2016</td>
<td>3.63</td>
<td>0.46</td>
<td>1.9</td>
<td>4.8</td>
<td>2,686</td>
</tr>
<tr>
<td>Google Rating for independent as of November 2016</td>
<td>3.93</td>
<td>0.43</td>
<td>1.5</td>
<td>5</td>
<td>6,356</td>
</tr>
<tr>
<td>TripAdvisor Rating for chain as of June 2018</td>
<td>3.75</td>
<td>0.51</td>
<td>1</td>
<td>5</td>
<td>1,587</td>
</tr>
<tr>
<td>TripAdvisor Rating for independent as of June 2018</td>
<td>4.0</td>
<td>0.47</td>
<td>1</td>
<td>5</td>
<td>3,617</td>
</tr>
</tbody>
</table>

Note: All monetary amounts are in December 2000 dollars.

The Google Trends Yelp search interest data is also summarized in Table 1.2. For periods before the start of Yelp’s penetration, March 2005, the search interest for Yelp is set to 0. Overall, during January 1995 to December 2015, the average Yelp interest for a metro area at a given month is 4.5, the standard deviation is 9.2. For TripAdvisor, the search interest is higher than Yelp, with an average of 10.8 and a standard deviation of 15.3. TripAdvisor as an online review platform, primarily aimed at travellers, was founded in February 2000, 4 years earlier than Yelp. It makes sense that consumers’ exposure to TripAdvisor is higher than Yelp on average. The overall trends in consumers’ search interest for both websites for the entire state of Texas are illustrated in Figure 3. As can be seen, TripAdvisor was active long before Yelp became popular in Texas. We can also

---

8The visitor spending data is at the annual frequency and has been interpolated to the monthly level between 1995 and 2015.

9Since a restaurant is at various distances to a traffic counting station, I weigh the daily traffic counts by the distance between a restaurant and the closest counting station to obtain a weighted number of daily traffic for my empirical analysis.
see that although Yelp started in Texas in March 2005, its popularity did not pick up until August 2007, from which point onwards, its popularity has surpassed that of TripAdvisor. It should be noted that despite its popularity, TripAdvisor is a platform for both hotels and restaurants, while Yelp concentrates mostly on restaurants. Nonetheless, the penetration of these two platforms over time is highly correlated, especially for years after 2010. This may be due to the increasing trend in the usage of the Internet and the Google search engine.

As mentioned before, besides the Google Trends data, I also use the rating history on Yelp to construct an alternative set of measure for online platforms’ penetration, particularly for Yelp. The rating history data gives the number of reviews written about a restaurant in each month. At the restaurant level, this data is highly endogenous with respect to the restaurant revenue because the higher the number of visitors to a restaurant, a higher the restaurant’s revenue, and the greater the number of reviews written about this restaurant. Therefore, I aggregate these number of reviews for each month at the metro region level and then divide them by the number of restaurants in that region to obtain an average number of reviews per month per restaurant. This measure is likely to even out the idiosyncratic shocks at the restaurant or market level, and can be used as an alternative representation of Yelp’s penetration. I plot this measure against time and compare the graph to that of the Google Trends data for major metro regions in Texas (including Austin, Houston, Dallas and Fort-Worth, and San Antonio). The juxtaposition of these graphs are shown in Figure 4. The left panel in the figure is the graph based on Google Trends data, and the right presents that using the average number of reviews per months per restaurant. It can be seen that these two graphs have almost identical shape, both indicating that Yelp is the most popular in Austin than the other

---

\[^{10}\text{I have also collected the Google Trends data on smaller platforms such as urbanspoon and zomato. However, their search interest is very small in comparison to that on Yelp and TripAdvisor, and therefore, not included in this graph or in the empirical analysis of the paper.}\]
Chapter 1

major regions. A simple calculation reveals that the correlation between these two measures is 0.85. Due to this high correlation, I choose only one measure in my empirical analysis, and that measure is the Google Trends data on consumers’ search interest because compared to the measure constructed from the number of reviews, this measure based on search interest includes not only supply of information on Yelp (which the number of reviews represents) but also the demand of information, i.e. consumers’ browsing the Yelp website to read reviews.

Table 1.2 also summarizes the restaurant level information collected from the Yelp, TripAdvisor and Google, in particular, restaurants’ overall average ratings. As can be seen, on average, chain restaurants receive a lower overall rating than their independent counterparts on all platforms: 2.96 vs. 3.52 on Yelp, 3.63 vs. 3.93 on Google, and 4.0 vs. 3.75 on TripAdvisor. The distribution of these ratings for each platform is shown in Figure 5.

As shown on the graph, all distributions are close to normal but skewed to the left. The distribution of Yelp ratings has a fatter tail and is more spread out compared to those for Google and TripAdvisor. Both Google and TripAdvisor ratings are also systematically higher than Yelp ratings. Their ratings’ averages are around 4.0, while Yelp’s is about 3.5. Nonetheless, ratings from all of these platforms are positively correlated. Their correlations are shown in Table 1.3. As can be seen, Yelp and Google ratings are more correlated with each other than with TripAdvisor. In particular, the correlation between Yelp and Google is 0.72, while those between them and TripAdvisor is only 0.58. This indicates that TripAdvisor is somehow different from the other two platforms. As explained earlier, TripAdvisor is primarily targeting travellers or tourists, whose taste and rating standards for restaurants may be systematically different from those for Yelp and Google, which are used mostly by the local population. Given this finding, I use Yelp and Google ratings as the primary quality measures in my paper since the focus of this paper is on consumers’ continuous learning activities and tourists, as one-time visitors, seldom learn accumulatively. In addition to these measures of quality, I also use the “predetermined” average ratings from Yelp, as has been discussed previously. Using these data and measures, I test the implications of the learning model in the next section.

Table 1.3: Correlation Between Ratings Across Platforms

<table>
<thead>
<tr>
<th></th>
<th>Yelp</th>
<th>Google</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>TripAdvisor</td>
<td>0.59</td>
<td>0.58</td>
</tr>
</tbody>
</table>

1.4 Reduced-Form Evidence

In the reduced-form analysis, I test the implications of the learning model. I present two main sets of results: (1) the first one is the revenue-profiles of restaurants. I show that the patterns of the revenue-age profiles are consistent with Implications 1 and 2, that is, high-quality (low-quality) independent restaurants have upward (downward) sloping revenue-age profiles while chains’ profiles
Figure 4: Yelp Search Interest and Number of Reviews per Restaurant for Selected Metro Areas, 2005-2015

Figure 5: Distribution of Restaurants over Average Ratings
stay constant. (2) the second is the bi-furcating effects of online review platforms, especially Yelp, on restaurants of high- and low-independent qualities. For high-quality independent restaurants, exposure to Yelp and other platforms increases their revenues and survival probabilities, whereas for low-quality independent restaurants, the exposure reduces the revenues and survival probabilities. Furthermore, exposure to Yelp and other online platforms has no effect on chain and old independent restaurants whose quality and reputation have been established. These results are consistent with Implications 2, 3 and 5. These two sets of results support that consumers are learning about restaurants’ quality through online review platforms.11

Section 1.4.1 below presents the results related to revenue-age profiles, and Section 1.4.2 discusses the bi-furcating results. In analyzing the bi-furcating effects of online platforms, I primarily focus on the effect from Yelp since it is the dominant platform. However, by controlling for the exposure to other online platforms, I also present the effect of other platforms on consumer learning and restaurant revenue. This analysis that includes other platforms utilizes all restaurants that are listed on some online review platform; that is, all restaurants included in Figure 2. This examination of the broader set of restaurants serves two main purposes: (1) one is that it tests if Yelp’s effect is limited only to restaurants listed on Yelp, not those that are not on Yelp but on other platforms. If the effects are not limited to restaurants on Yelp, then my measure of the Yelp exposure must have picked up the effects of other forces, especially the exposure to other platforms. The results show that the effect of Yelp exposure is limited to those restaurants listed on Yelp; the penetration of other online platforms affects those restaurants are not on Yelp. (2) The second one is that it quantifies the effect of other platforms on restaurant revenue. The results show that other platforms also contribute to consumer learning of restaurant quality, though the effects are not as long in magnitude as Yelp’s, confirming that Yelp is the dominant platform in the restaurant industry.

1.4.1 Revenue-Age Profiles

In this subsection, I first demonstrate the shape of the revenue-age profiles of restaurants and see if they comply with the prediction of the learning model. To produce the revenue-age profiles, I use the following regression:

\[
\log(Rev_{jt}) = X_{jt}\theta_x + \theta_n n_{mt} + \theta_ac n_{mt}^C + \sum_{a=1}^{a_{max}} 1\{a_{jt} = a\}(\theta_a + \theta_{ac} D_j^C) + \theta_t + \theta_{tc} D_j^C + \theta_m + \xi_{jt}
\]

(1.4.1)

where \(X_{jt}\) include market demand shifters and restaurant characteristics; \(n_{mt}^I\) and \(n_{mt}^C\) are the number of independent and chain rivals in a market at time \(t\); \(a_{jt}\) denotes the age for restaurant \(j\) at time \(t\), and \(a_{max}\) is the maximum age; \(D_j^C = 1\{\text{restaurant } j \text{ belongs to a chain}\}\) is a chain

11 As an additional check to see if consumers use Yelp ratings as quality signals, I carried out a regression discontinuity design (RDD) to evaluate the direct effect of Yelp displayed rating on revenues. The RDD is carried out in the same fashion as the one done in Luca (2016). The results confirm that consumers take Yelp ratings into consideration as quality signals, and are consistent with other results in this section. The results from the RDD analysis are available upon request.
dummy; $\theta$ are the coefficients associated with these variables; $\theta_t$ and $\theta_{tc}$ are calendar fixed effects for independent and chain restaurants respectively; $\theta_m$ is the market fixed effect; $\xi_{jt}$ is an aggregate demand shock for restaurant $j$ and has mean 0.

It is easy to see that this specification is based on equation 1.2.15 from Section 3.2. Recall that there are an alcohol-to-total-revenue ratio ($g_{jt}$) and a price ($p_{jt}$) in equation 1.2.15 and both need to be controlled for in the reduced-form analysis. To control for $g_{jt}$, I use restaurant cuisine types and market demographic characteristics (both included in $X_{jt}$) as well as calendar time fixed effects. It is common knowledge in the restaurant industry that the alcohol-to-total-revenue ratio of a restaurant depends mostly on a restaurant’s cuisine type; for example, a bar has a much higher alcohol-to-revenue ratio than a breakfast joint on average. Furthermore, this ratio can also vary according to the demographic characteristics of the market; for instance, a bar in an area with lots of young college students is likely to have a higher alcohol-to-food ratio than one in an area with mostly young families. Given these considerations, I can use restaurant cuisine types, market characteristics and time fixed effects to control for $g_{jt}$. Restaurant cuisine type information is obtained from the Yelp and TripAdvisor platforms. For those restaurants that are not listed on those platforms, I use the restaurant names and the description on Google to classify their type.

For prices, I use the price information provided by all three platforms: Google, Yelp and TripAdvisor. Since all prices are shown in ranges on these platforms, I use the midpoint of the price range as the menu price. For those restaurants whose prices are available from the platforms, I use the average price of restaurants with the same cuisine and in the same geographic market to approximate their prices. The specific dollar amount for each price range is shown in Table 23 of the Appendix. To capture the horizontal differentiation among consumers over restaurant observed characteristics, I also include in variable $X_{jt}$ the interaction terms between a restaurant’s cuisine type and the demographic characteristics of a market.

In the revenue-age profile regression above, I do not control for the cohort effects, i.e. a fixed effect depending on which year a restaurant opens. I do so because it is well known that the effects of age, cohort and calendar time cannot be separately identified from one another without assuming a nonlinear functional form for one of the effects, usually age (See Mehta et al. (2010), Hall et al. (2002)). By not including cohort effects in this regression, I want to be able to non-parametrically examine the shape of the revenue-age profiles, so that I can choose an appropriate functional form for them in later analysis. At that point, I can also control for restaurant fixed effects, which includes cohort effects. As will be seen later, I use log(age) to approximate the shape of the revenue-age profiles when examining restaurants by quality groups.

Here in equation 1.4.1, the age dummy coefficients form the shape of the revenue-age profiles. They represent how revenues evolve with respect to age after controlling for all demand variables and competitive effects in the market as well as aggregate trends in consumer taste that can affect chain and independent restaurants differently. The revenue-age profiles and their confidence intervals are

---

12 A detailed description of the types is shown in Table 22 of the Appendix

13 In the result tables in this paper, I do not report the actual estimates of these interaction terms since there are quite a large number of them.
plotted in Figure [6]. The red line in the figure represents the profile for chain restaurants and the blue line for independent restaurants.

As can be seen, chain restaurants enjoy a premium in revenue compared to independent restaurants, and their revenue-age profiles are fairly flat, except a spike at the beginning, which can be explained by that chains often do intensive promotions when opening a new outlet. Very different from chains, independent restaurants have a revenue-age profile that slopes upwards initially and flattens later on. The shapes of these revenue-age profiles are consistent with Implications 1 and 4. The profile for independent restaurants, in particular, conforms to a revenue-age profile for a high-quality restaurant, as shown by the redlines in Figure [1].

However, the speed of learning implied by this profile, is extremely slow: the profile starts to converge to a steady level after about 120 months, suggesting that learning about a restaurant’s quality takes 10 years to complete. This pattern may be the result of selection related to endogenous exit: restaurants of high quality and those that received persistent positive demand shocks are more likely to survive than those with low quality and those that experienced negative demand shocks. Those that exit early tend to be lower quality restaurants, and they participate in the estimation of the age coefficients for earlier years. The ones for later years are estimated from higher quality restaurants, which often have larger monthly revenues. Therefore, the upward sloping curve of the revenue-age profiles may be artificially created by this selection (Jovanovic, 1982).

To account for selection, I use Heckman’s correction and carry it out in two steps: first, I ran a probit model with the active status of a restaurant as the dependent variable. In addition to controls in $X_{jt}$, I add hourly wage in the food-service industry as an exclusion restriction since it relates to the cost of running a restaurant, but not demand. Then I obtain the predicted survival probability of a restaurant from this regression and construct the inverse Mills ratio. In the second step, I plug the inverse Mills ratio into regression 1.4.1 and re-estimate the revenue-age profiles. The profiles with Heckman’s correction are shown in Figure 7.

As can be seen, the blue line, the revenue-age profile for independent restaurants, no longer slopes upward steeply, although the gentle upward slope is still discerned for the age from 0 to 50. The shape of the profile for chain restaurants does not change much from the previous graph except that the estimates of the age coefficients seem to have become much more noisy after age 150. The initial spike on the chain’s profile still exists, though a bit subdued. These revenue-age profiles confirm that even after controlling for selection, the shapes of the revenue-age profiles are consistent with the implications of the consumer learning model.

To zoom in further on independent restaurants and examine if Implication 1 of the learning model holds, I break restaurants down into high- and low-quality classes based on their Google and Yelp ratings and then construct the revenue-age profiles for each quality class. According to Implication 1, the profiles of low-quality restaurants could slope systematically downwards and those

---

14 When estimating equation 1.4.1, I drop the last year’s observation before exit because the revenues recorded during that period are often sporadic and systematically much smaller than other periods.

15 For example, a chain may cut the meal prices steeply but maintain the prices of alcohol in the first couple of months, which could increase both the demand and revenues for alcoholic drinks.
Figure 6: Revenue-Age Profiles for All Restaurants

Figure 7: Revenue-Age Profiles for All Restaurants With Heckman Correction
of high-quality restaurants upwards. To test this implication, I use the following specification:

$$\log(\text{Rev}_{jt}) = [\theta_a + \theta_{ar} R_j] \times \log(a_{jt}) \times (1 - D_j^C) + [\theta_{ac} + \theta_{acr} \times R_j] \times \log(a_{jt}) \times D_j^C$$

$$+ X_{jt}\theta_x + \theta_{n1n_{int}} + \theta_{n2n_{int}} + \theta_{t1} + \theta_{t2}D_i^C + \theta_j + \xi_{jt}$$

(1.4.2)

where \(\log(a_{jt})\) is the log form of the age of a restaurant \(j\) at time \(t\), and the coefficients associated with it inform us of the slopes of the revenue-age profiles. \(R_j\) is restaurant quality, and it is interacted with \(\log(a_{jt})\), such that the coefficient \(\theta_a + \theta_{ar} R_j\) reflects the slope of the revenue-age profile of independent restaurants of a given quality \(R_j\). As an additional test for chain restaurants’ revenue-age profiles by quality class, I include parameters indicating the slopes of the profiles for chains as well: \(\theta_{ac} + \theta_{acr} R_j\). Since the age effect is now captured by a nonlinear function \(\log(a_{jt})\), I add a restaurant fixed effect, \(\theta_j\), to control for cohort effects and other time-invariant restaurant specific effects.

In this exercise, I use a parametric approach for two main reasons: (1) one is that the shapes of the revenue-age profiles as presented in Figures 6 and 7 suggest a logarithm functional form of age in the revenue equation. (2) the second is that breaking down restaurants by quality class reduces the sample size and the power of the test, and the key purpose of this exercise is to compare the slopes of the revenue-age profiles across quality classes. A parametric approach is therefore much more appealing than a non-parametric one. The additional benefit of this parametric approach is that I can add restaurant fixed effects to control for cohort effects, which could not be done in a non-parametric regression.

The regression results are shown in Table 1.4. In the regression results presented in this table, I use Yelp ratings as quality measures, both the Yelp ratings as of November 2016 and the “predetermined” ratings. I also use Google ratings as a quality measure to conduct the exercise, and given that the results are very similar, I leave them to the Appendix.

In Table 1.4, the left panel displays coefficient estimates for the key parameters of interest (i.e. \(\theta_a, \theta_{ar}, \theta_{ac}\) and \(\theta_{acr}\)) and those of the main control variables; the right panel presents the composite coefficient \(\theta_a + \theta_{ar} R_j\), which, as mentioned before, captures the slopes of the revenue-age profiles for each quality class. Both panels include two columns showing results from using November 2016 Yelp ratings and the “predetermined" Yelp rating from restaurants’ first 10 reviews respectively. As can be seen from the first column of the left panel in Table 1.4, the coefficient for the interaction term \(\log(\text{age}_{jt}) \times R_j\) for independent restaurants, \(\theta_{ar}\), is 0.0517, positive and significant, indicating that the slopes of the revenue-age profiles increase as restaurant qualities increase. The coefficient, \(\theta_a\), represents the slope associated with a rating of “0" if there is such a class. Its estimate is −0.172, negative and significant, suggesting that when the quality rating is very low, the revenue-age profile of an independent restaurant indeed slopes downwards. These individual coefficients are translated into the composite ones in the first column of the right panel. As shown, the slopes of the revenue-age profiles for lower rated restaurant on Yelp (2 and 3 stars) are negative, while those for 4 and 5 stars restaurants are positive. This bi-furcating pattern of the revenue-age profiles is consistent
Table 1.4: Revenue-Age Profiles by Quality

<table>
<thead>
<tr>
<th></th>
<th>(1) Log Revenue</th>
<th>(2) Log Revenue</th>
<th>Slope Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Independent, (\theta_a))</td>
<td>(Independent, (\theta_{ar}))</td>
<td>(Chain, (\theta_{ac}))</td>
</tr>
<tr>
<td>Log Age (Independent, (\theta_a))</td>
<td>-0.172*** (0.0559)</td>
<td>-0.333*** (0.0698)</td>
<td>2</td>
</tr>
<tr>
<td>Log Age (\times) Rating (Independent, (\theta_{ar}))</td>
<td>0.0517*** (0.0155)</td>
<td>0.0996*** (0.0179)</td>
<td>4</td>
</tr>
<tr>
<td>Log Age (Chain, (\theta_{ac}))</td>
<td>-0.101*** (0.0260)</td>
<td>-0.0860 (0.0724)</td>
<td>5</td>
</tr>
<tr>
<td>Log Age (\times) Rating (Chain, (\theta_{acr}))</td>
<td>0.00742 (0.00827)</td>
<td>0.00645 (0.0215)</td>
<td>6</td>
</tr>
<tr>
<td>Traffic (thousands)</td>
<td>0.0188 (0.0727)</td>
<td>0.0421 (0.101)</td>
<td>7</td>
</tr>
<tr>
<td>Log Total Population</td>
<td>0.128 (0.232)</td>
<td>-0.211 (0.424)</td>
<td>8</td>
</tr>
<tr>
<td>Log Visitor Spending</td>
<td>0.216 (0.140)</td>
<td>-0.0238 (0.238)</td>
<td>9</td>
</tr>
<tr>
<td>Population Density</td>
<td>-55.05 (139.1)</td>
<td>26.31 (196.1)</td>
<td>10</td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>0.331 (1.098)</td>
<td>-1.039 (0.855)</td>
<td>11</td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>2.402* (1.346)</td>
<td>1.853 (1.473)</td>
<td>12</td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>0.0658 (1.340)</td>
<td>-0.515 (1.542)</td>
<td>13</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>-0.349 (0.344)</td>
<td>0.0672 (0.315)</td>
<td>14</td>
</tr>
<tr>
<td>Share of White population</td>
<td>0.154 (0.176)</td>
<td>0.268** (0.125)</td>
<td>15</td>
</tr>
<tr>
<td>Share of Black population</td>
<td>0.716 (0.507)</td>
<td>0.367 (0.309)</td>
<td>16</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>-1.062** (0.492)</td>
<td>-0.692* (0.303)</td>
<td>17</td>
</tr>
<tr>
<td>Number of Independent Rivals</td>
<td>-0.00589*** (0.00235)</td>
<td>-0.00116 (0.00227)</td>
<td>18</td>
</tr>
<tr>
<td>Number of Chain Rivals</td>
<td>-0.0127*** (0.00447)</td>
<td>-0.00433 (0.00825)</td>
<td>19</td>
</tr>
<tr>
<td>Cuisine Dummy (\times) Demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>20</td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
<td>21</td>
</tr>
<tr>
<td>Time FE(\times)Chain Dummy</td>
<td>Yes</td>
<td>Yes</td>
<td>22</td>
</tr>
<tr>
<td>Restaurant FE</td>
<td>Yes</td>
<td>Yes</td>
<td>23</td>
</tr>
<tr>
<td>Quality Measure</td>
<td>Yelp November 2016 Rating</td>
<td>Yelp Rating at First 10 reviews</td>
<td>24</td>
</tr>
<tr>
<td>N</td>
<td>257,933</td>
<td>138,794</td>
<td>25</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* p < 0.10, ** p < 0.05, *** p < 0.01
with Implication 1.

The coefficients associated with chain restaurants’ revenue-age profiles are also interesting. The coefficient of the interaction term $\log(\text{age}_{jt}) \times R_j$, $\theta_{acr}$ is 0.00742, very small and insignificant. This further confirms Implication 4 that the slopes of the revenue-age profiles of chain restaurants should not change with respect to quality classes. The coefficient for $\log(\text{age}_{jt})$, $\theta_{ac}$, however, is $-0.101$, negative and significant, indicating the revenue-age profiles for chain restaurants of all quality classes slope downwards. This does not contradict Implication 4, which states flat chain restaurants’ profiles. This downward slope simply reflects the initial spike in the revenue-age profiles of chain restaurants as shown in both Figures 6 and 7.

As a robustness check to results from the first column, I also use the “predetermined” ratings as a quality measure. As mentioned before, the “predetermined” ratings are more exogenous quality measures though they come with the caveat that they are associated with a smaller sample size. The endogeneity concern of using the November 2016 Yelp average ratings as a quality measure is that they measure quality at the end of the sample period and might have picked up random demand shocks to restaurants’ revenues during the sample period. For example, some restaurants might have drawn some large and positive quality shocks from their quality distribution, and these shocks would push the restaurants into a higher quality class based on the average of all past signals (in this case Yelp ratings). Therefore, the ex-post average ratings might have simply captured the positive quality shocks to a restaurant, and that could contribute to an upward sloping curve of the revenue-age profile of higher-quality restaurants, especially if these quality shocks are persistent overtime.

To eliminate this type of endogeneity, I use the “predetermined” ratings as a quality measure and run the regression. The “predetermined” rating I use here is the average of restaurants’ first 10 reviews on Yelp, and the sample associated with it are restaurants’ revenue streams after they had received their first 10 reviews. This “predetermined” rating would not pick up any quality shocks in the later periods. The regression results are shown in the second column of the left panel of Table 1.4. As can be seen, the coefficients associated with the independent restaurants’ age profiles, $\theta_a$ and $\theta_{ar}$, are significant and share the same signs as those from the previous regression. This implies that the bi-furcating pattern of the revenue-age profiles for independent restaurants still holds. Furthermore, the coefficients for chain restaurants $\theta_{ac}$ and $\theta_{acr}$ ($-0.0805$ and $0.00645$ respectively) are both small and insignificant, reflecting two features of the revenue-age profiles for chains: (1) the slopes of the revenue-age profiles of chain restaurants are not very different across quality classes; (2) the revenue-age profiles for chains are flat with a slope that is not significantly different from 0. These findings for chains are again consistent with Implication 4. Note that these results are slightly different from those in the first regression as shown in column 1. The coefficient $\theta_{ac}$ ($-0.0805$) for chains is not significant in this regression, while the one from the first regression ($-0.101$) is. This is because the sample used in the second regression includes revenues after restaurants first received 10 reviews on Yelp, and therefore the initial spike in the revenue-age profiles for chain restaurants, as shown in both Figures 6 and 7, do not appear here in this regression.
Another interesting aspect between the results from these two regressions with different quality measures is that the slopes of revenue-age profiles shown in the second regression are steeper than those shown in the first. This is especially discernable from comparing the two columns of the slope coefficients from these two regressions in the right panel of Table 1.4. As can be seen, the slope coefficients from the second regression with the "predetermine" ratings, as shown by the second column, all bigger in magnitude compared to those from the first column. This difference in magnitude reflects the difference in the sample used: the second regression uses restaurant revenues after the penetration of Yelp, while the first uses restaurant revenues during all periods in the sample starting from January 1995. If Yelp helps consumers learn about restaurant qualities, then the revenue-age profiles of restaurants that opened after the penetration of Yelp would have steeper slopes compared to those restaurants that entered earlier. Therefore, if Yelp indeed helps consumers learn, the slopes of the revenue-age profiles from the second regression sample should be larger in magnitude than those of the first regression sample, which includes the earlier years of many old restaurants. In this aspect, the difference between the slopes of the revenue-age profiles from these two samples seem to suggest that Yelp helps consumers learn.

To conclude, overall, the shapes of the revenue-age profiles for chain and independent restaurants are consistent with Implications 1 and 4 of the consumer learning model. In the next section, I examine how online review platforms help consumers learn by examining their effects on restaurant revenues by age.

1.4.2 Bifurcating Effect of Exposure to Online Review Platforms

To investigate the effects of online review platforms on consumer learning, I test if Implications 2 and 3 of the learning model hold. These implications state that online reviews platforms should have a negative effect on low-quality restaurants’ revenues and a positive effect on high-quality ones, and that their effects are stronger at the earlier ages of a restaurant than later ages. Essentially, greater the exposure to online platforms should increase the slopes of the revenue-age profiles of restaurants. I call the differential effects of online review platforms on high- and low-quality restaurants’ revenues the bi-furcating effect of online review platforms.

To test these implications, I first focus on the effect of Yelp, since it is the dominant platform. Then I extend the econometric model to include other online review platforms. The subsection that follows discusses the overarching empirical strategy for testing the bi-furcating effects of Yelp, and Subsection 1.4.2.2 demonstrates the results for the effects of Yelp on both restaurants’ revenue and

---

16One may argue that these patterns of the revenue-age profiles are also consistent with the theory of passive Bayesian learning on the firm side as demonstrated in Jovanovic (1982). The shape of the revenue-age profiles is showing firms’ learning not consumers’ learning. Here, I argue here that these two theories are consistent with each other. Firms learn about their quality from revenues as signals per period and update their beliefs of their own qualities, and then they decide the optimal output per period. Firms’ learning source, the revenues, come from demand as consumers gradually discover the true qualities of restaurants. From this perspective, firms’ learning really comes from consumer learning. Furthermore, in the restaurant industry, the output is driven on an order by order basis, unlike in many other industries, where outputs are determined by firms ex-ante. Therefore, Jovanovic’s (1982) theory, where firms deciding output ex-ante, is not necessarily a very suitable model in my context.
survival probability. A number of robustness checks for the endogeneity of the measure of Yelp exposure are also explored here. Subsection 1.4.2.3 investigates the effects of both Yelp and other online platforms on learning.

1.4.2.1 Empirical Strategy

Testing Implications 2 and 3 requires that the exposure to Yelp be interacted with both quality and age. Interacting Yelp exposure with each individual age is practically infeasible for two reasons: (1) first, it leaves very little power of estimate for the coefficient associated with each interaction term; (2) second, including these interaction terms prevent us from incorporating restaurant fixed effects because a fixed effect estimation relies on the change in each restaurant’s revenues to identify the effects of Yelp; however, there is only one observation associated with each age for each restaurant. There is no way to identify Yelp’s effect at each age by including restaurant fixed effect, unless we assume a functional form for age. As shown in the analysis of the revenue-age profiles, the functional form log(age) seem to approximate well the effect of age. Interacting Yelp exposure with log(age) could be an option. However, it is not a good option because Figure 1 from the numerical example in the model section shows that Yelp’s effect diminishes with age. An interaction term of Yelp exposure and log(age) would imply an increasing effect of Yelp as restaurant ages. Therefore, the log(age) functional form is not suitable in this analysis either.

Due to the difficulty in choosing an appropriate functional form for age in the reduced-form analysis, I choose a simpler approach and leave the estimation of Yelp’s effect on revenue at each age to the structural analysis. Since the purpose of the reduced-form is to see if the data show patterns that support the various assumptions in the structural model, a simple approach that divides restaurants into young and old groups and examine differential effects of Yelp on these groups is entirely appropriate. I divide restaurants into young and old groups based on their entry dates. More specifically, those that entered two years before Yelp’s penetration in March, 2005 are classified as old restaurants and those that entered after are defined as young. Based on this division of groups and the consumer learning theory, Yelp should have little effects on the revenue of the old group and significant effect on that of the young group. I interact the measure Yelp exposure with quality and run regressions separately for the young and old groups of restaurants. The econometric specification of the regression is shown below:

\[
\log(\text{Rev}_{jt}) = \theta_y + \theta_{yt}R_j \times \log(\text{Yelp}_{mt}) \times (1 - D_C^C) + \theta_{yc} + \theta_{ycr} \times R_j \times \log(\text{Yelp}_{mt}) \times D_C^C + \sum_{i} X_{jt} \theta_x + \theta_{tn}n_{mt}^t + \theta_{nc}n_{mt}^c + \theta_t + \theta_{tc}D_i^C + \theta_j + \xi_{jt} \]

(1.4.3)

where \(\log(\text{Yelp}_{mt})\) is the logarithm of the Google Trends measure of Yelp exposure in market \(m\) at time \(t\). All other variables and parameters are the same as previously defined.

\(^{17}\)Here I can also use the average number of reviews per restaurant for each metro region as a measure for Yelp exposure. However, as presented by Figure in the data section, these two measures are essentially the same, and the regression results from using the other measure do not change much. Therefore, I present only results from using the Google Trends data as the measure for Yelp exposure.
In this regression, the effect of Yelp is easy to interpret: for a given quality level, the coefficient \( \theta_y + \theta_{yr} R_j \) (or \( \theta_{yc} + \theta_{ycr} R_j \) for chains) is the percentage change in restaurant \( j \)'s revenue when Yelp exposure increases by 100%, and this is also the average treatment effect on all restaurants of the quality class \( R_j \) across all ages. It should be noted that the Yelp exposure measure and age are conditionally independent after controlling for the various FEs, and therefore, the coefficient from this regression reflects the average treatment effect of Yelp. By running this regression separately for old and young restaurants, we can see the differential effects of Yelp exposure on the revenue of different age cohorts.

The identification of Yelp's effect in this regression comes from the difference in the change of restaurant revenues across various regions that are subject to various levels of Yelp exposure. By including the restaurant fixed effects (\( \theta_j \)) and the calendar time fixed effects (\( \theta_t \) and \( \theta_{tc} \)), and by interacting Yelp exposure with quality levels, the main identification strategy can be seen as a triple difference-in-differences (3-DD) approach; that is, within a given quality group, the difference in the change of restaurant revenue teases out Yelp's effect, and these effects are compared across different quality groups. In addition to the comparison in the quality dimension, I also compare restaurants by chain affiliations (chain v.s. independent) by including a chain dummy, and across age groups (young v.s. old) by running the regression separately for each age group.

The key underlying assumption for identification is that by controlling for these fixed effects, Yelp exposure is exogenous; that is, during the process of Yelp penetration, there were no other unobservable factors that would affect revenues the same way as Yelp exposure did, with opposite effects on high- and low-quality young independent restaurants and very little effects on chain or old independent restaurants.

Nonetheless, there remain two key concerns of endogeneity for the measure for Yelp exposure. The first one is that Yelp exposure could be correlated with the inherent ability or willingness to learn in each region. For example, the people in Austin might be different from those in Houston in that Austinites may be much eager to learn about restaurants’ quality than Houstonites. Therefore, even without Yelp, learning would have still been faster in Austin than in Houston. This source of endogeneity can be detected from the trends in the change of restaurant revenues across regions before the penetration of Yelp, the “pre-trend.” A test that examines if the “pre-trend” is correlated with the measure of Yelp exposure across geographic regions will be sufficient to detect how severe this source of endogeneity is. To do so, I conduct a “pre-trend” analysis or a Placebo test: I move Yelp’s penetration and Yelp exposure data to earlier dates by 10 years; that is, instead of starting in March 2005, I move Yelp’s penetration to March 1995. Then I ran the regression (1.4.3) for the sample of restaurant revenues before Yelp’s penetration. If Yelp exposure is indeed correlated with the innate ability or willingness to learn across regions, then the “placebo" Yelp exposure will pick up the effect of these across-region differences. This placebo test is shown in Subsection 1.4 of the Appendix. This test result demonstrates that this type of endogeneity is not of serious concern.

The second source of endogeneity regarding the measure of Yelp’s penetration is that although Yelp was the dominant online review platform for restaurants during the sample period, Yelp was
not the only platform. TripAdvisor and Google played an important role. As consumers rely more and more on information from the Internet, Yelp’s penetration is likely to be correlated with the penetration of other platforms. In other words, without controlling for other platforms, the estimated effects of Yelp may also pick up the effects of other platforms. To test this source of endogeneity, I control for the penetration of other platforms in the revenue regression for Yelp listed restaurants, such that we can separate out the marginal effect of Yelp on restaurants listed on Yelp in addition to the effect of other platforms. Furthermore, I also compare restaurants listed on Yelp with those that are not listed on Yelp. The comparison between restaurants listed on Yelp and off Yelp is particularly revealing: if Yelp’s penetration measure is confounded with the penetration of other platforms, then the revenue of restaurants that are not listed on Yelp would appear to respond to Yelp’s penetration. My analysis shows that this is not the case. The penetration of other platforms do not affect the revenue of Yelp listed restaurants, and Yelp does not affect the revenues of restaurants that are not listed on Yelp once the penetration of other platforms are controlled for. This result implies that Yelp’s effect on restaurants listed on Yelp is not significantly affected by other platforms’ presence. This analysis is shown in Section 1.4.2.3.

It is important to note that this analysis not only examines the exogeneity of the measure of Yelp exposure, but also investigates the effect of the penetration of other platforms on restaurant revenues across platforms. It captures the broader effect of online review platforms on restaurants. This analysis forms the foundation for the structural analysis in Section 1.5. Below are a demonstration of results.

1.4.2.2 Empirical Results of Yelp’s Bi-furcating Effects

Using the sample of restaurants listed on Yelp, I ran regression 1.4.3 separately for old and young restaurants. For quality, I use a variety of measures: overall average Google Review ratings as of November 2016, overall average Yelp ratings in November 2016, and the “pre-determined” ratings for when restaurants receive their first 10 reviews. Note that using the average of the first 10 reviews as the pre-determined ratings is without the loss of generality. The results for using the average of first 15 or 20 reviews are very similar. The results are shown in Table 1.5.

The first two columns represent results for old and young restaurants by using the Google rating as quality measure. The middle two columns are results from using the Yelp ratings, and the last two are those from using the first 10 ratings. For the first two quality measures, the regressions are run on the entire sample of Yelp listed restaurants, while for the last quality measure, the sample is restaurant revenues after restaurants received their first 10 reviews on Yelp. In this table, the first 4 rows of coefficients are the key parameters of interest. They represent the effect of Yelp exposure on restaurant revenues by quality class. There are 3 notable features in these regression results: first, it is uniformly true across all regression results that the coefficients associated with chain restaurants are insignificant, implying that Yelp exposure does not have an effect on chain restaurants. This is consistent with Implication 5 that online review platforms should not affect chain restaurants through learning.
Table 1.5: Bifurcating Effects of Yelp Exposure on Revenues

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Yelp (θ_y)</td>
<td>-0.272</td>
<td>-0.575***</td>
<td>-0.0246</td>
<td>-0.357***</td>
<td>-0.118</td>
<td>-0.474***</td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td>(0.143)</td>
<td>(0.137)</td>
<td>(0.0999)</td>
<td>(0.115)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>Log Yelp × Rating (θ_yr)</td>
<td>0.0323</td>
<td>0.150***</td>
<td>-0.00661</td>
<td>0.108***</td>
<td>0.0467</td>
<td>0.130***</td>
</tr>
<tr>
<td></td>
<td>(0.0333)</td>
<td>(0.0303)</td>
<td>(0.0444)</td>
<td>(0.0281)</td>
<td>(0.0326)</td>
<td>(0.0417)</td>
</tr>
<tr>
<td>Log Yelp × Chain Dummy (θ_yc)</td>
<td>-0.251</td>
<td>-0.0547</td>
<td>-0.0143</td>
<td>0.00516</td>
<td>0.251</td>
<td>-0.151</td>
</tr>
<tr>
<td></td>
<td>(0.170)</td>
<td>(0.0641)</td>
<td>(0.0919)</td>
<td>(0.0418)</td>
<td>(0.286)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>Log Yelp × Chain Dummy × Rating (θ_ycr)</td>
<td>0.0625</td>
<td>0.0225</td>
<td>-0.00414</td>
<td>0.00714</td>
<td>-0.0504</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td>(0.0413)</td>
<td>(0.0174)</td>
<td>(0.0294)</td>
<td>(0.0135)</td>
<td>(0.0700)</td>
<td>(0.0277)</td>
</tr>
<tr>
<td>Traffic (thousands)</td>
<td>-0.129</td>
<td>-0.00069</td>
<td>-0.0296</td>
<td>-0.0037</td>
<td>-0.154</td>
<td>0.261**</td>
</tr>
<tr>
<td></td>
<td>(0.191)</td>
<td>(0.0864)</td>
<td>(0.102)</td>
<td>(0.0876)</td>
<td>(0.209)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Log Total Population</td>
<td>0.404</td>
<td>-0.174</td>
<td>0.257</td>
<td>-0.149</td>
<td>0.665</td>
<td>-0.453</td>
</tr>
<tr>
<td></td>
<td>(0.302)</td>
<td>(0.190)</td>
<td>(0.238)</td>
<td>(0.221)</td>
<td>(0.206)</td>
<td>(0.203)</td>
</tr>
<tr>
<td>Income (millions)</td>
<td>7.75^*</td>
<td>0.719</td>
<td>3.41</td>
<td>1.54</td>
<td>-0.394</td>
<td>1.82</td>
</tr>
<tr>
<td></td>
<td>(4.09)</td>
<td>(1.77)</td>
<td>(3.89)</td>
<td>(2.27)</td>
<td>(1.73)</td>
<td>(3.66)</td>
</tr>
<tr>
<td>Log Visitor Spending</td>
<td>-0.107</td>
<td>0.308***</td>
<td>-0.171</td>
<td>0.355***</td>
<td>-0.0282</td>
<td>0.0883</td>
</tr>
<tr>
<td></td>
<td>(0.171)</td>
<td>(0.118)</td>
<td>(0.161)</td>
<td>(0.131)</td>
<td>(0.240)</td>
<td>(0.250)</td>
</tr>
<tr>
<td>Population Density</td>
<td>-277.5</td>
<td>-124.9</td>
<td>-7.998</td>
<td>-72.52</td>
<td>-138.9</td>
<td>-231.1</td>
</tr>
<tr>
<td></td>
<td>(218.4)</td>
<td>(145.4)</td>
<td>(205.8)</td>
<td>(148.6)</td>
<td>(281.0)</td>
<td>(235.1)</td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>3.254</td>
<td>0.891</td>
<td>3.058^*</td>
<td>-0.0143</td>
<td>-0.813</td>
<td>-2.747</td>
</tr>
<tr>
<td></td>
<td>(2.352)</td>
<td>(1.094)</td>
<td>(1.845)</td>
<td>(1.134)</td>
<td>(1.73)</td>
<td>(2.980)</td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>1.750</td>
<td>1.207</td>
<td>1.407</td>
<td>2.735</td>
<td>0.923</td>
<td>3.357</td>
</tr>
<tr>
<td></td>
<td>(2.452)</td>
<td>(0.944)</td>
<td>(2.424)</td>
<td>(1.908)</td>
<td>(0.821)</td>
<td>(2.920)</td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>3.281**</td>
<td>-1.451</td>
<td>2.648</td>
<td>-1.268</td>
<td>-0.410</td>
<td>-3.174</td>
</tr>
<tr>
<td></td>
<td>(1.766)</td>
<td>(1.445)</td>
<td>(1.630)</td>
<td>(1.413)</td>
<td>(2.46)</td>
<td>(3.587)</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>-0.135</td>
<td>-0.00043</td>
<td>-0.002</td>
<td>-0.0506</td>
<td>-0.280</td>
<td>-0.618</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.037)</td>
<td>(0.004)</td>
<td>(0.027)</td>
<td>(0.056)</td>
<td>(0.458)</td>
</tr>
<tr>
<td>Share of White population</td>
<td>0.801**</td>
<td>-0.0181</td>
<td>0.822**</td>
<td>-0.130</td>
<td>0.429^*</td>
<td>0.134</td>
</tr>
<tr>
<td></td>
<td>(0.330)</td>
<td>(0.150)</td>
<td>(0.330)</td>
<td>(0.158)</td>
<td>(0.235)</td>
<td>(0.241)</td>
</tr>
<tr>
<td>Share of Black population</td>
<td>1.319</td>
<td>0.928^*</td>
<td>1.365^*</td>
<td>0.083</td>
<td>-1.400</td>
<td>1.015***</td>
</tr>
<tr>
<td></td>
<td>(0.841)</td>
<td>(0.549)</td>
<td>(0.826)</td>
<td>(0.543)</td>
<td>(0.948)</td>
<td>(0.621)</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>0.880</td>
<td>-1.149**</td>
<td>1.499</td>
<td>-1.125**</td>
<td>-0.300</td>
<td>-0.888</td>
</tr>
<tr>
<td></td>
<td>(1.002)</td>
<td>(0.432)</td>
<td>(0.002)</td>
<td>(0.430)</td>
<td>(0.060)</td>
<td>(0.520)</td>
</tr>
<tr>
<td>Number of Independent Rivals</td>
<td>-0.00501*</td>
<td>-0.00074</td>
<td>-0.00527**</td>
<td>-0.00464**</td>
<td>-0.00289</td>
<td>-0.00298</td>
</tr>
<tr>
<td></td>
<td>(0.00234)</td>
<td>(0.00025)</td>
<td>(0.00030)</td>
<td>(0.00033)</td>
<td>(0.00022)</td>
<td>(0.00031)</td>
</tr>
<tr>
<td>Number of Chain Rivals</td>
<td>-0.01021**</td>
<td>-0.01036*</td>
<td>-0.07833**</td>
<td>-0.006433</td>
<td>0.00460</td>
<td>0.00200</td>
</tr>
<tr>
<td></td>
<td>(0.000482)</td>
<td>(0.000484)</td>
<td>(0.000768)</td>
<td>(0.00097)</td>
<td>(0.00172)</td>
<td>(0.00100)</td>
</tr>
</tbody>
</table>

Cuisine Dummy × Demographics
- Yes
- Yes
- Yes
- Yes
- Yes
- Yes

Time FE
- Yes
- Yes
- Yes
- Yes
- Yes
- Yes

Time FE × Chain Dummy
- Yes
- Yes
- Yes
- Yes
- Yes
- Yes

Restaurant FE
- Yes
- Yes
- Yes
- Yes
- Yes
- Yes

Sample Entry 2 years before Yelp penetration Entry after Yelp penetration Entry 2 years before Yelp penetration Entry after Yelp penetration Entry 2 years before Yelp penetration Entry after Yelp penetration


N 121,083 165,221 145,004 180,278 59,220 65,219

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01
Second, the top four coefficients associated with Yelp exposure in Columns 1, 3 and 5 are also insignificant. These results are for the sample of old restaurants that entered long before Yelp’s penetration; they suggest that Yelp exposure does not have an effect on old restaurants. This pattern is consistent with Implication $3$ that the effect of online review platforms on restaurant revenue should diminish with age.

Third, the top two coefficients in the 2nd, 4th and 6th columns are all significant at the 99% confidence level. These coefficients reflect the effect of Yelp exposure on young independent restaurants that entered after Yelp’s penetration. The intercept parameter $\theta_y$ is negative in all 3 columns, indicating that when a restaurant’s quality is low, Yelp exposure has a negative effect on the revenue. The slope parameter $\theta_{yr}$ is positive in all 3 regressions, suggesting that the effect of Yelp exposure on revenue increases with the quality level. A negative intercept and positive slope together, however, illustrates a bi-furcating effect of Yelp on revenue. That is, when a restaurant’s quality is very low, Yelp has a large and negative effect on the revenue, and when the quality is very high, Yelp’s effect is still big but positive. For restaurants in the average range of quality, Yelp’s effect is small and may be insignificant. This insight is better illustrated once we translate $\theta_y$ and $\theta_{yr}$ into the composite effect of Yelp by quality level, $\theta_y + \theta_{yr}R_j$. Table 1.6 shows these composite parameters.$^{18}$

<table>
<thead>
<tr>
<th>Star Rating</th>
<th>Google</th>
<th>Yelp</th>
<th>Pre-determined Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-0.266***</td>
<td>-0.141***</td>
<td>-0.197**</td>
</tr>
<tr>
<td></td>
<td>(0.0715)</td>
<td>(0.0468)</td>
<td>(0.0784)</td>
</tr>
<tr>
<td>3</td>
<td>-0.111***</td>
<td>-0.0325</td>
<td>-0.0581</td>
</tr>
<tr>
<td></td>
<td>(0.0380)</td>
<td>(0.0259)</td>
<td>(0.0473)</td>
</tr>
<tr>
<td>4</td>
<td>0.0433**</td>
<td>0.0757***</td>
<td>0.0806**</td>
</tr>
<tr>
<td></td>
<td>(0.0202)</td>
<td>(0.0270)</td>
<td>(0.0425)</td>
</tr>
<tr>
<td>5</td>
<td>0.198***</td>
<td>0.184***</td>
<td>0.219***</td>
</tr>
<tr>
<td></td>
<td>(0.0447)</td>
<td>(0.0487)</td>
<td>(0.0697)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

As shown in the table, when Yelp exposure increases by 100%, the revenue of a 2 star restaurant based on on Google ratings would decline by 26.6%, while that of a 5 star restaurant would increase by 19.8%. The effect on a 2 star rating restaurant is more subdued when we use Yelp ratings. For a 2 star restaurant rated on Yelp, the revenue drops by 14.1% as Yelp exposure doubles, instead of 26.6%. At first glance, this appears a big difference. However, a careful examination of the Google and Yelp ratings as shown in Figure $5$ tells us that Google ratings have a much higher average than Yelp ratings due to the inclusion of many low-quality restaurants.

$^{18}$Since the sample of new restaurants do not include restaurants with a 1-star rating, I do not report the Yelp effect for that quality class.
Yelp ratings. In particular, a restaurant rated 2 stars on Google has an exceptionally low quality, and that explains why its revenue would decrease so much when Yelp exposure doubles. The 2 star rated restaurant on Yelp, on the other hand, may be equivalent to a 2.5 to 3 star restaurant on Google, and therefore, its revenue declines much less dramatically. At the higher end, the effects of Yelp on top rated 5 star restaurants based on both ratings are very similar (19.8% v.s. 18.4%). This is not surprising since 5 stars is the top tier, and restaurants rated 5 stars on Yelp are very likely rated also 5 stars on Google. The results from using the pre-determined ratings show a similar pattern to those based on the other two rating systems.

To summarize, these three key features of the results confirm that Implications 2, 3 and 5 of the consumer learning model hold, and they hold regardless of the quality measure that we use. As additional robustness checks to these results, I also use Heckman’s correction to control for endogenous exit and time\times metro region fixed effects to account for other factors that may be correlated with the measure of Yelp exposure. The analyses from these robustness checks give a very similar set of results, and they are left to Section 3 of the Appendix.

1.4.2.2.1 Effect on Survival Rates Given Yelp’s effects on revenue, it should also affect restaurants’ survival rates. In particular, it should help the market weed out the lower quality restaurants by pushing them to exit faster, and help the higher-quality restaurants stay in the market longer. As this may be one of the best functions of online review platforms, I here briefly present an analysis of Yelp’s effect on restaurants’ survival rates.

To examine this aspect, I run a linear probability model by using exactly the same econometric structure as that shown in equation 1.4.3 only that I replace the dependent variable (log(\text{Rev}_{jt})) with \text{action}_{jt} = 1 \text{restaurant } j \text{ is active at time } t \text{ and I add in } w_{mt}, \text{ wage for food service workers, as an additional explanatory variable.} \text{ Here I use only Yelp ratings as a quality measure, since the samples associated with both Google ratings and predetermined ratings do not have enough observations in restaurants’ exit. The regression results are shown in Table 25 of Section 8 of the Appendix. Here I present only the composite parameters that demonstrate Yelp’s effect on the survival rate of restaurants in Table 1.7.}

\footnote{It is worth noting that these results are different from the results regarding Bayesian learning in Luca (2016). Here I use ratings as a noisy measure of quality, and then examine how Yelp exposure affects differently the revenues of low-quality and high-quality independent restaurants. Luca, on the other hand, treats rating as a cause of fluctuations in revenues, and examines if this causal effect is greater when the number of reviews is higher. The interaction term of the number of reviews with average rating in Luca’s specification (on page 12 of his paper) is analogous to the interaction term of Yelp exposure and quality in my paper. However, based on his specification and results, no matter what quality level a restaurant has, restaurant revenues always increase with the number of reviews when the average rating is held fixed. This result contradicts the Bayesian learning theory. That is, if the average rating of a restaurant is low, a greater precision of this signal (i.e. a higher number of reviews) should damage a restaurant’s revenue, not increases it. The problem with Luca’s specification is that the number of reviews is an endogenous measure of precision because the number of reviews is driven by the number of consumers who visited the restaurant and thereby driven by revenues. Contrary to Luca’s findings, I show that for bad quality restaurants, a greater exposure to Yelp (or greater precision of signals) can harm revenues, a result that is consistent with the Bayesian learning theory. Another difference between Luca’s work and mine in demonstrating learning is that I relate Yelp’s effect to age. I show that when learning ends, rating changes on Yelp no longer have an effect on revenues. This provides additional evidence for Bayesian learning.}
Table 1.7: Effects of Yelp Exposure on Survival Rate (Yelp Ratings)

<table>
<thead>
<tr>
<th>Rating</th>
<th>Effect on Survival Rate</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Star</td>
<td>-0.00179*</td>
<td>(0.00102)</td>
</tr>
<tr>
<td>3 Star</td>
<td>-0.000556</td>
<td>(0.000631)</td>
</tr>
<tr>
<td>4 Star</td>
<td>0.000678</td>
<td>(0.000618)</td>
</tr>
<tr>
<td>5 Star</td>
<td>0.00191*</td>
<td>(0.000992)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

As can be seen, when Yelp exposure doubles, the survival rate of a 5-star new independent restaurant would increase by 19.1 basis points, and that of a 2-star restaurant would decrease by 17.9 basis points. For the two other rating classes, the effects are insignificant. This result confirms that Yelp helps the market weed out bad quality restaurants by pushing them to exit faster.

1.4.2.3 Yelp and Other Platforms

In this section, I extend my analysis to other online platforms and their effects on restaurant revenues. This analysis serves two purposes: one is to deal with the concern of endogeneity in the Yelp exposure measure; the other is to test if other platforms also have a bifurcating effect on restaurant revenue through consumer learning. This analysis ultimately allows me to capture the effects of all the main online platforms on consumer learning and welfare in the structural analysis.

As mentioned previously, a main concern of the endogeneity of the measure for Yelp exposure is that it may be correlated with the propagation of other platforms. To test if this source of endogeneity is important, I compare restaurants listed on Yelp to those that are not, and I control for the penetration of other platforms at the same time. The idea is that once I control for consumers’ exposure to other platforms, Yelp should have no effect on restaurants that are not listed on Yelp. In other words, Yelp’s effect should be limited to only those that are listed on Yelp.

For this analysis, I use the sample of all restaurants that are listed on Google. As mentioned in the data section, Google covers over 9,000 restaurants and has a rating for every restaurant on it. This rating provides a uniform quality measure for restaurants listed on Yelp vs. those that are not. This uniform quality measure is one of the main advantages of using the sample of all restaurants listed on Google for this analysis. Another advantage of this sample is that it includes a wide distribution of restaurants listed across platforms. According to Figure 2, there are four types of restaurants based on which platform they are listed on: (1) those listed only on Yelp, (2) those listed on both Yelp and TripAdvisor, (3) those listed only on TripAdvisor, and (4) those listed neither on Yelp nor on TripAdvisor. There are at least 1,100 restaurants in each category, providing power for identifying the effects of various online platforms. A comparison of the effects of Yelp exposure on these four types of restaurants while controlling for the penetration of other platforms provides an assessment of whether the correlation between Yelp’s penetration and that of other platforms significantly biased the results based on regression 1.4.3.

\[^{20}\]Since the sample of new restaurants do not include restaurants with a 1-star rating, I do not report the Yelp effect for that quality class.
To control for the penetration of other platforms, I use the Google Trends’ measure of the search interest on TripAdvisor website. This measure captures not only the penetration of TripAdvisor but also partly the penetration of other platforms, especially Google. In the same vein, the measure for Yelp exposure also represents partly consumers’ exposure to other platforms including Google. Ideally, we want to have a measure for the exposure to Google for its restaurant review service specifically, but that data is difficult to obtain. Since restaurant review service is only a fraction of Google’s total function, Google Trends data on the search term “Google” is contaminated with many other search purposes besides writing or reading restaurant reviews. Given this limitation, the comparison of Yelp’s effects on restaurants listed on Yelp vs. those listed off Yelp becomes an important source of identification: if our Yelp exposure measure has picked up important variations in Google’s penetration, then given that all restaurants in the sample are listed on Google, Yelp exposure should have significant effects on restaurants that are not listed on Yelp as Google should affect all restaurants listed on it. And that is the key reason behind this research design. In addition, once we control for the penetration of another major platform TripAdvisor, we can tease out further the marginal effect of Yelp exposure on restaurants listed on Yelp.

To conduct this analysis, I use the following specification:

\[
\log(\text{Rev}_{jt}) = \sum_{p=1}^{4} D_j^p \left( \left[ \theta_p^y + \theta_p^g R_j \right] \times \log(\text{Yelp}_{mt}) \times (1 - D_j^C) + \left[ \theta_p^y + \theta_p^{ycr} \times R_j \right] \times \log(\text{Yelp}_{mt}) \times D_j^C \right) \\
\sum_{p=1}^{4} D_j^p \left( \left[ \theta_p^y + \theta_p^g R_j \right] \times \log(\text{TripAdvisor}_{mt}) \times (1 - D_j^C) + \left[ \theta_p^y + \theta_p^{ycr} \times R_j \right] \times \log(\text{TripAdvisor}_{mt}) \times D_j^C \right) \\
+ X_j \theta_p + \theta_{n_{mt}} n_{mt} + \theta_{nc} n_{mt} + \theta_t + \theta_{o_c} D_j^C + \theta_j + \xi_{jt} \quad (1.4.4)
\]

where \(D_j^p\) is a dummy for what type restaurant \(j\) is, with \(p = 1\) denoting that restaurant \(j\) is listed only on Yelp, 2 listed on both Yelp and TripAdvisor, 3 TripAdvisor only, and 4 on neither Yelp nor TripAdvisor. \(\log(\text{TripAdvisor}_{mt})\) is the Google Trends measure of search interest for TripAdvisor website, similarly defined as \(\log(\text{Yelp}_{mt})\). Parameters \(\theta_p^y, \theta_p^g, \theta_p^{ycr}\) capture the effect of TripAdvisor on the revenues of type \(p\) restaurants. The parameters associated with the Yelp effect are also indexed with \(p\) to differentiate Yelp’s effect on each type of restaurant.

This specification is very similar to regression 1.4.3 only that I allow Yelp and TripAdvisor to have different effects on the revenues of each type of restaurant. I run the regression for only young restaurants that entered after Yelp’s penetration, and the results are shown in Table 1.9. The four columns at top of the table show the parameters associated with the effects of Yelp and TripAdvisor respectively for each type of restaurant. The bottom half illustrate the common parameters associated with controls. As can be seen from columns (1) and (2), even after controlling for the penetration of TripAdvisor, the parameters associated with Yelp’s effect is still significant only for independent restaurants, not for chain restaurants. Furthermore, the parameters associated

---

21 Another way to collect information on the penetration of Google’s restaurant review service is to collect the rating history for each restaurant listed on Google. That activity is also met with challenges: first, Google is very sophisticated at detecting any type of automatic web-scraping; second, the review for a restaurant on Google does not have a precise date, but rather labeled as how many months ago for recent reviews and years ago for older reviews. This information gives too coarse a measure for the penetration of Google’s review service.
with the effect of TripAdvisor are not significant on Yelp listed restaurants. Note that the estimates of Yelp’s effect on restaurants that are listed only on Yelp (shown in column 1) are much higher in magnitude compared to those that are listed both on TripAdvisor and on Yelp, with the intercept parameters at -0.912 vs. -0.425 and slopes parameters at 0.231 vs. 0.124. These parameters imply that Yelp’s bi-furcating effects are much stronger for restaurants that are listed only on Yelp. This difference is striking, but very reasonable in two aspects: (1) first, restaurants that are listed only on Yelp are likely to be more sensitive to Yelp’s penetration compared to those restaurants that are listed also on TripAdvisor, which provides additional information on restaurants. (2) second, there may be a selection effect: although restaurants can be listed by anyone on either TripAdvisor or Yelp, the people who create restaurant profiles may be different for TripAdvisor from those for Yelp. Users of TripAdvisor are more likely to be travellers, while those of Yelp can be both locals and tourists. Therefore, it is possible that restaurants that are listed on TripAdvisor are those that are more visible and accessible to travellers, and they do not need to rely on Yelp as much for information exposure.

It is also interesting to see from columns (1) and (2) that TripAdvisor’s penetration has very little effect on the revenues of Yelp listed restaurants. This result is very reasonable for restaurants that are listed only on Yelp, but surprising for those that are also listed on TripAdvisor. Why would TripAdvisor have no effect on these restaurants? This phenomenon is even more puzzling for restaurants that are listed only on TripAdvisor, as shown in column 3. The parameters associated TripAdvisor’s penetration are not significant, except for the interaction term between rating and chain, which has a confidence level at 90%. The weak effects of TripAdvisor estimated in this regression may be explained by the measure of quality we use in this regression. Recall from Section 1.3.2 that TripAdvisor’s ratings are correlated with those from Yelp and Google only at about 58%, while the correlation between the other two platforms’ ratings are around 72%. The quality measure used in this regression is the Google ratings, which might have lumped together both good and bad restaurants as rated on TripAdvisor into one quality class, given the low correlation between the two rating systems. As a result, TripAdvisor’s penetration shows very little effect on the revenues of restaurants for each quality class as defined by Google.

Columns (3) and (4) of the table show the effects of Yelp and other platforms on restaurants that are not listed on Yelp. As can be seen, none of these coefficients associated with the Yelp exposure is significant. However, interestingly, those coefficients associated with the TripAdvisor exposure are significant for those independent restaurants that are listed on neither Yelp nor TripAdvisor. This result is puzzling: TripAdvisor seems to affect restaurants that are not listed on it, while having no effect on those that are listed only on TripAdvisor. This seemingly odd result may be explained by that the measure for TripAdvisor’s penetration could be highly correlated with the penetration of Google’s restaurant review service and that of other smaller platforms. As a result, TripAdvisor seems to affect independent restaurants that are listed neither on Yelp nor on TripAdvisor (i.e. only on Google). Note that the coefficients associated with TripAdvisor for those restaurants listed only on Google imply a bi-furcating effect on restaurant revenues. The intercept coefficient $\theta_g$ is
−0.318, negative and significant, while the slope parameter $\theta_{gp}$ is 0.0825, positive and significant. These parameter suggest that for restaurants with a below 4 star rating, TripAdvisor’s penetration affects revenue negatively, whereas for those with a 4 star or higher rating, the effect is positive. The precise effects for each rating are shown in Table 1.8. This bi-furcating effect again indicates that reviews from other platforms help consumers learn about restaurant qualities.

Table 1.8: Effect of Other Platforms on Revenue (Google Rating)

<table>
<thead>
<tr>
<th>Effect on Revenue</th>
<th>2 Star</th>
<th>3 Star</th>
<th>4 Star</th>
<th>5 Star</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.153**</td>
<td>-0.0705</td>
<td>0.012</td>
<td>0.0945*</td>
</tr>
<tr>
<td></td>
<td>(0.0754)</td>
<td>(0.0589)</td>
<td>(0.0226)</td>
<td>(0.0556)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

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

Based on these results, we should interpret the measure for TripAdvisor exposure as a measure for other platforms, especially Google. These result tell a number of interesting stories: (1) Firstly, Yelp exposure affects restaurants listed only on Yelp, not restaurants listed only on other platforms, eliminating the concern of the endogeneity issue associated with Yelp’s exposure measure; (2) Secondly, in addition to Yelp, other platforms also help consumers learn about restaurant qualities; in particular, when the exposure to other platforms (as measured by TripAdvisor’s search interest on Google) doubles, the revenues of 2-star restaurants as rated by Google would decrease by 15.3%, whereas those of 5 star restaurants increase by nearly 9.5%. (3) Finally, other platforms have negligible effects on restaurants listed on Yelp, confirming that Yelp is the most dominate online review platform in the restaurant industry, at least during the sample period of my study.

To summarize, the empirical analysis in this section provides evidence that supports that online review platforms speed up consumers’ learning about restaurant qualities. In particular, through learning, Yelp affects both restaurants’ revenues and survival probabilities. Specifically, doubling Yelp exposure increases the revenue of a high-quality new independent restaurant by 18-27% and decreases that of a low-quality restaurant by about the same amount. Doubling Yelp exposure also raises the survival rate of a new high-quality independent restaurant by 18 basis points and reduces that of a low-quality restaurant by a similar magnitude. In contrast, Yelp does not affect chain or old independent restaurants. Other platforms have a similar effect on those restaurants that are listed not Yelp but on other platforms. These results are consistent with Implications 2, 3 and 5 of the learning model.

Although there may be other channels through which Yelp can have an effect on revenue besides consumer learning, such as managerial learning and horizontal sorting of consumers into their favourite restaurants, I argue in detail that those alternative channels do not play an important role in Section 4.1 of the Appendix. In particular, to argue that managerial learning, where managers can improve their restaurants’ qualities over time by responding to reviews, is not a salient factor for Yelp to affect revenue, I examine the rating history of each Yelp listed restaurants. If managerial learning is important, then restaurants’ ratings should go up over time. The results from the analysis over ratings show that restaurants’ ratings do not show systematically upward
Table 1.9: Effects of Yelp Exposure on Restaurants Listed on Vs. off Yelp

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log Revenue</td>
<td>Log Revenue</td>
<td>Log Revenue</td>
<td>Log Revenue</td>
</tr>
<tr>
<td>Log Yelp ($\theta_y$)</td>
<td>-0.912**</td>
<td>-0.425***</td>
<td>-0.231</td>
<td>-0.183</td>
</tr>
<tr>
<td></td>
<td>(0.3850)</td>
<td>(0.1500)</td>
<td>(0.2030)</td>
<td>(0.1700)</td>
</tr>
<tr>
<td>Log Yelp×Rating ($\theta_{yr}$)</td>
<td>0.231**</td>
<td>0.124***</td>
<td>0.0681</td>
<td>0.0412</td>
</tr>
<tr>
<td></td>
<td>(0.0088)</td>
<td>(0.0371)</td>
<td>(0.0524)</td>
<td>(0.0449)</td>
</tr>
<tr>
<td>Log Yelp×Chain Dummy ($\theta_{yc}$)</td>
<td>-0.0307</td>
<td>-0.0655</td>
<td>0.172</td>
<td>0.335</td>
</tr>
<tr>
<td></td>
<td>(0.0063)</td>
<td>(0.0876)</td>
<td>(0.1440)</td>
<td>(0.3580)</td>
</tr>
<tr>
<td>Log Yelp×Chain Dummy×Rating ($\theta_{ycr}$)</td>
<td>0.0127</td>
<td>0.0284</td>
<td>-0.0496</td>
<td>-0.0886</td>
</tr>
<tr>
<td></td>
<td>(0.0264)</td>
<td>(0.0243)</td>
<td>(0.0353)</td>
<td>(0.1020)</td>
</tr>
<tr>
<td>Log TripAd ($\theta_g$)</td>
<td>0.421</td>
<td>0.101</td>
<td>0.0761</td>
<td>-0.318**</td>
</tr>
<tr>
<td></td>
<td>(0.4970)</td>
<td>(0.1680)</td>
<td>(0.1840)</td>
<td>(0.1270)</td>
</tr>
<tr>
<td>Log TripAd×Rating ($\theta_{gr}$)</td>
<td>-0.107</td>
<td>-0.0255</td>
<td>-0.0073</td>
<td>0.0825**</td>
</tr>
<tr>
<td></td>
<td>(0.1280)</td>
<td>(0.0413)</td>
<td>(0.0449)</td>
<td>(0.0322)</td>
</tr>
<tr>
<td>Log TripAd×Chain Dummy ($\theta_{gc}$)</td>
<td>-0.0293</td>
<td>0.00093</td>
<td>-0.287</td>
<td>-0.282</td>
</tr>
<tr>
<td></td>
<td>(0.1120)</td>
<td>(0.0576)</td>
<td>(0.2770)</td>
<td>(0.2680)</td>
</tr>
<tr>
<td>Log TripAd×Chain Dummy×Rating ($\theta_{gcr}$)</td>
<td>-0.00045</td>
<td>-0.000861</td>
<td>0.143*</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>(0.0290)</td>
<td>(0.0156)</td>
<td>(0.0359)</td>
<td>(0.0715)</td>
</tr>
</tbody>
</table>

Traffic (thousands)  

Log Total Population  

Income (millions)  

Log Visitor Spending  

Population Density  

Share of population (age 15-34)  

Share of population (age 35-64)  

Share of population (age 65 and up)  

Share of Hispanic population  

Share of White population  

Share of Black population  

Share of Asian population  

Number of Independent Rivals  

Number of Chain Rivals  

Cuisine Dummy × Demographics Yes  

Time FE Yes  

Time FE×Chain Dummy Yes  

Restaurant FE Yes  

Sample Period After Yelp’s penetration  

Sample Platform Presence Yelp not TripAd Yelp & TripAd not Yelp but TripAd neither Yelp nor TripAd  

Rating Google Rating  

N 239,104  

Standard errors in parentheses  

* p < 0.10, ** p < 0.05, *** p < 0.01
or downward trends. This finding also motivates the assumption in my structural analysis that restaurants’ qualities stay constant over time. In the following Section 1.5, I investigate the effect of online review platforms on restaurants’ life-time revenue, market shares, and consumer welfare through a structural analysis.

1.5 Structural Estimation

Structural model and estimation have several advantages compared to reduced-form analysis: the most well known advantage is that the structural parameters are policy invariant and thereby, structural models are easier to use in policy experiments. For example, in the context of my study, the effects of Yelp and other online platforms in the reduced-form section are an average effect across restaurants of all ages. Depending on the makeup of the age groups, the policy experiment of removing the existence of Yelp and other online platforms will produce different results. The reduced-form parameters will change as a result. The structural model, on the other hand, can pin point these age-specific effects by estimating the learning parameters. Furthermore, a structural model allows us to comment on how online review platforms affect the life-time revenue of a restaurant, consumers’ speed of learning, dynamics of market shares, and consumer welfare. These are the key questions under examination in the counterfactual analysis in this section.

The demand model with consumer learning introduced in Section 3.2 lends itself well to structural estimation, and I make additional extensions to the model to accommodate some features of the data. The extensions and estimating equations of the structural model are described in Section 1.5.1. One important extension is adding serial correlation into the aggregate restaurant-specific demand shocks. It increases the challenge in estimation because one cannot apply the traditional Kalman filter to form the likelihood function. To overcome this challenge, I derived the closed-form representation of the log-likelihood function, such that the maximum likelihood estimation approach can be used. These issues related to identification and estimation of the model are discussed in Section 1.5.2. The parameter estimates and their interpretations are described in Section 1.5.3. Once I have the structural parameter estimates in hand, I conduct counterfactual analyses to examine the effects of online review platforms on restaurants’ life-time revenues, dynamics in market shares, consumers’ speed of learning and their welfare. The results of the counterfactual analysis are shown in Section 1.5.4.

1.5.1 Structural Model

As mentioned before, I make a few extensions to the demand model with learning discussed in Section 3.2 to accommodate the features of my data. These extensions include:

1. The initial prior mean (denoted by $A_j^0$ hereafter) of a restaurant’s quality is modelled to vary across quality class, instead of being constant across all restaurants as in the previous model. In particular, it is a linear function of the quality measure: $A_j^0 = \theta^j R_j$. In addition to its flexibility, this setup also controls for unobserved time-invariant characteristics of a restaurant.
that are correlated with our quality measure but do not require learning.

2. Instead of identifying the true mean quality for each individual restaurant, $A_j$, I model $A_j$ as a linear function of the quality measure: $A_j = \theta_0 + \theta_r R_j$, similar to that of the prior. This formulation keeps the structural model parsimonious. Note that the linear function of $A_j$ includes an intercept, but that of $A_j^p$ does not. This is because only one intercept can be identified in addition to the constant.

3. The aggregate demand shock $\xi_{jt}$ is assumed to be serially correlated with an autocorrelation $\rho$. This assumption of $\xi_{jt}$ being serially correlated is important in the structural estimation of a learning model because both consumers’ expectations of quality $\eta_{jt}$ and $\xi_{jt}$ are unobserved, and $\eta_{jt}$ is serially dependent. Without the flexibility in controlling for serial correlation in $\xi_{jt}$, the importance of learning can be overestimated.

4. Restaurants listed on various platforms are given their own learning parameters; that is, I expand the $\kappa$ and $\lambda$ parameters to accommodate the different learning speed for the four types of restaurants listed on various platforms.

With these modifications, the estimating equation can be written as

$$y_{jt} = X_{jt} \theta - a \log(p_{jt}) + \eta_{jt} + \sum_{k=1}^{5} \theta_k \tau^k + \theta_c + \xi_{jt}$$

$$\xi_{jt} = \rho \xi_{jt-1} + \varepsilon_{jt}, \varepsilon_{jt} \sim i.i.d. N(0, \sigma_{\varepsilon}^2)$$

(1.5.1)

where $y_{jt} \equiv \log(Rev_{jt}) - \log(Rev_{0mt})$ is the log ratio of restaurant $j$’s revenue at time $t$ to that of the outside option; $\tau$ is the calendar time in month, and $\sum_{k=1}^{5} \theta_k \tau^k$ is a fifth order polynomial that controls the time trend in the changes of consumers’ taste; $\theta_c$ is a constant; $\eta_{jt}$ is consumers’ expected true mean quality for restaurant $j$. It follows a non-stationary Markov process:

$$\eta_{jt} = (1 - \beta_{jt-1}) \eta_{jt-1} + \beta_{jt-1} A_{jt-1} + \nu_{jt}$$

(1.5.2)

where $A_j = \theta_0 + \theta_r R_j$; $\nu_{jt} = \beta_{jt-1} (\bar{A}_{jt-1} - A_j)$, and

$$\beta_{jt-1} = \frac{1}{\sigma_\Delta^2} \frac{P_{jt-1} q_{jt-1}}{1 + \sum_{l=1}^{4} \frac{P_{jl} q_{jl}}{\sigma_\Delta^2}}$$

(1.5.3)

where $P_{jt} = \sum_{p=1}^{4} D_p^j (\kappa_p + \lambda_p y Elp_{pmt} + \lambda_p g TripAd_{mt})$; $\lambda_p y$ is the learning parameter associated with the Yelp exposure measure and $\lambda_p g$ the one for TripAdvisor. Again, $D_p^j$ is a dummy for what type restaurant $j$ is in terms of which platform it is listed. Here, each type of restaurant has their own learning parameters $\kappa_p$, $\lambda_p y$ and $\lambda_p g$. Furthermore, as shown in Section 3.2, $\bar{A}_{jt-1} \sim N(A_j, \sigma^2_{\Delta})$, and therefore,

$$\nu_{jt} \sim N(0, \frac{\beta_{jt-1}^2 \sigma^2_{\Delta}}{P_{jt} q_{jt-1}})$$

(1.5.4)
is an error term with mean 0 and a time-varying variance $22$.

The system of equations 1.5.1 to 1.5.4 are the estimating equations in the structural analysis. The sample I use in the structural estimation includes all restaurants in my dataset. However, given that online review platforms seem to have little effect on chain restaurants and older independent restaurants, I use only the new independent restaurants that are listed on Google (i.e. those restaurants included in Section 1.4.2.3) to estimate the structural parameters, especially the learning parameters. All the other restaurants are included in the model for the purpose of computing market shares.

For measures of Yelp$_{mt}$ and TripAd$_{mt}$, I use the natural log of the Google Trends Yelp search interest data for each platform. The quality measure here is Google rating. The revenue of the outside good Rev$_{0mt}$ is constructed by subtracting the sum of revenues of all restaurants in market $m$ at time $t$ from the total consumer expenditure spent on eating out in that market for that month. The total consumer expenditure on eating out is calculated by multiplying the total income of all population in market $m$ at time $t$ by the share of total income spent on eating out at time $23$. The following section discusses the identification and estimation of this structural model.

1.5.2 Identification, Estimation Strategy and Structural Estimates

The identification of most structural parameters in equations 1.5.1 to 1.5.4 is straightforward except for the learning parameters. One important issue is how we differentiate $\eta_{jt}$, learning, from $\xi_{jt}$ since both are serially correlated. The key source of identification between the two elements comes from the fact that the serial correlation in $\eta_{jt}$ depends on age, while that in $\xi_{jt}$ does not. In particular, once restaurants’ quality has been mostly learned, $\eta_{jt}$ stays constant, i.e. with an autocorrelation equal to 1, whereas $\xi_{jt}$ still follows a stationary autoregressive process with a constant autocorrelation at $\rho$. Most of the learning parameters can easily be shown to be identifiable, except for $\sigma^2$, which cannot be separately identified from the $\kappa$ and $\lambda$ parameters, and is thereby normalized to 1. A detailed discussion regarding the identification of each learning parameter is provided in Section 1.5 of the Appendix.

To estimate equations 1.5.1 to 1.5.4 I use a quasi-maximum likelihood approach. The key issue of estimation of this set of equations is that $\eta_{jt}$ is unobservable to the econometrician, and needs to be integrated out. Without the serial correlation in $\xi_{jt}$ ($\rho = 0$), integrating out $\eta_{jt}$ to form the likelihood function is relatively easy because the structural model with $\rho = 0$ is in fact a Gaussian hidden Markov model (HHM), where the Kalman filter can be applied. However, with $\rho \neq 0$, the Kalman filter no longer applies. To overcome this challenge, I derive the closed-form solution for the likelihood function, and use a quasi-maximum likelihood approach to estimate the structural parameters. To the best of my knowledge, this paper is the first that uses this technique to estimate a Bayesian learning model. Most papers shy away from the autocorrelation assumption in the

$22$ Although I do not have quantity data for each restaurant, I use restaurant revenues divided by the midpoint of the Yelp price range to approximate $q_{jt}$ in the Bayesian updating equations.

$23$ As mentioned in Section 1.4, the data for the share of total income spent on eating out is collected from the Consumer Expenditure Survey held by the U.S. Bureau of Labor Statistics for the South Region of the U.S.
demand shocks $\xi_{jt}$ due to technical difficulties. The detail of this estimation approach is presented in Section 6 of the Appendix.

An important issue in the estimation of this model is the endogenous price. To correct for endogeneity, I use the standard BLP instruments as an IV for prices ((Berry et al. (1995))). More specifically, the IVs are the average ratings of the independent and chain competitors in a market respectively. The assumption associated with these IVs is that the entry of the competing restaurants is exogenous in the quality dimension. That is, at the time of entry, competing restaurants choose a quality level that is unrelated to $\eta_{jt}$ and the demand shock $\xi_{jt}$ of restaurant $j$. Since entry is sticky, once entered, competing restaurants’ mean quality is fixed, and thereby uncorrelated with any future demand shocks for restaurant $j$, but it affects the prices that restaurant $j$ can charge. Under this assumption, these instruments are valid. I carry out this IV approach in two steps. In the first step, I regress restaurant prices on these instruments and obtain a fitted value. The estimating equation in the first step is shown below:

$$\log(p_{jt}) = \theta_p I_{AR}^I_{jt} + \theta_{pC} I_{AR}^{ch}_{jt} + \tilde{\epsilon}_{jt}$$

where $AR^I_{jt}$ and $AR^{ch}_{jt}$ are the average ratings of restaurant $j$’s independent and chain rivals at time $t$ respectively.

In the second step, I plug the fitted log price into equation (1.5.1) and conduct the estimation using the quasi-maximum likelihood approach. To adjust the standard errors from the second step to incorporate the errors from the first step estimation, I implement the method introduced in Murphy and Topel (1985).²⁴

### 1.5.3 Structural Estimates

The estimates of the key structural parameters of interest are shown in Table 1.10; they include the learning parameters and the price coefficients. For the estimates of control variables, please refer to the online Appendix. In this table, I present two sets of results from two different models: one without serial correlation in $\xi_{jt}$ (i.e. $\rho = 0$) and the other with serial correlation (i.e. $\rho \neq 0$). Comparing these models in terms of the parameter estimates, log likelihood values and Bayesian information criteria (BICs) give us an idea of how important including the serial correlation in the aggregate demand shocks $\xi_{jt}$ is for the consumer learning process we examine.

As can be seen from the table, the model without serial correlation performs worse than the one with, both in terms of the log likelihood values and the BIC criteria than the model that accounts for serial correlation. Specifically, the log likelihood for the model with $\rho = 0$ is $-79,232$ much lower than that from the model with serial correlation, $-61,051$. Consequently, the BIC for the model without serial correlation is also much higher than that of the model with at 159,591 v.s. 123,242. These statistics imply that the model that accounts for serial correlation is a better model.

In terms of the estimates for the key learning parameters, however, these models yield fairly

²⁴Specifically, page 374, equation 15.
Table 1.10: Structural Estimates: Key Parameters

<table>
<thead>
<tr>
<th></th>
<th>Model with $\rho = 0$</th>
<th>Model with $\rho \neq 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log Revenue Ratio</td>
<td>Log Revenue Ratio</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>4.01*** (0.531)</td>
<td>3.88*** (0.692)</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>1.24e-4*** (4.38e-6)</td>
<td>8.18e-5*** (8.068e-6)</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>2.39e-5*** (1.71e-6)</td>
<td>5.03e-5*** (6.81e-6)</td>
</tr>
<tr>
<td>$\kappa_3$</td>
<td>4.09e-5*** (4.07e-6)</td>
<td>1.44e-4*** (1.67e-5)</td>
</tr>
<tr>
<td>$\kappa_4$</td>
<td>4.96e-5*** (1.44e-5)</td>
<td>6.90e-5 (2.7e-4)</td>
</tr>
<tr>
<td>$\lambda_{1y}$</td>
<td>4.49e-8 (1.92e-7)</td>
<td>4.68e-8 (1.68e-7)</td>
</tr>
<tr>
<td>$\lambda_{2y}$</td>
<td>2.09e-5*** (1.19e-6)</td>
<td>3.10e-5*** (3.39e-6)</td>
</tr>
<tr>
<td>$\lambda_{3y}$</td>
<td>4.15e-8 (1.24e-7)</td>
<td>4.24e-8 (2.47e-7)</td>
</tr>
<tr>
<td>$\lambda_{4y}$</td>
<td>8.13e-6** (3.78e-6)</td>
<td>1.58e-7 (2.27e-6)</td>
</tr>
<tr>
<td>$\lambda_{1g}$</td>
<td>2.10e-7 (3.14e-6)</td>
<td>7.25e-8 (4.49e-7)</td>
</tr>
<tr>
<td>$\lambda_{2g}$</td>
<td>4.25e-8 (1.35e-7)</td>
<td>4.15e-8 (2.44e-7)</td>
</tr>
<tr>
<td>$\lambda_{3g}$</td>
<td>4.22e-8 (1.32e-7)</td>
<td>4.90e-8 (3.56e-7)</td>
</tr>
<tr>
<td>$\lambda_{4g}$</td>
<td>2.34e-5*** (6.07e-6)</td>
<td>5.22e-5* (2.75e-5)</td>
</tr>
<tr>
<td>$\sigma_A^2$</td>
<td>2.31*** (0.0269)</td>
<td>0.954*** (0.0179)</td>
</tr>
<tr>
<td>$\theta_r^p$</td>
<td>-0.0826*** (0.00155)</td>
<td>-0.0954*** (0.00245)</td>
</tr>
<tr>
<td>$\theta_0$</td>
<td>-1.67*** (0.0781)</td>
<td>-1.56*** (0.0763)</td>
</tr>
<tr>
<td>$\theta_r$</td>
<td>0.380*** (0.020)</td>
<td>0.368*** (0.0193)</td>
</tr>
<tr>
<td>$\sigma_e^2$</td>
<td>0.105*** (0.000465)</td>
<td>0.107*** (0.000475)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.494*** (0.00340)</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>162,875</td>
<td>162,875</td>
</tr>
<tr>
<td>log likelihood</td>
<td>-79,232</td>
<td>-61,051</td>
</tr>
<tr>
<td>BIC</td>
<td>159,591</td>
<td>123,242</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01
similar results. The estimates of \( \kappa \) and \( \lambda \) have fairly similar size, only that the significance level drops once we add in the serial correlation. In particular, \( \lambda_{4y} \), the coefficient of the Yelp exposure for the group of restaurants that are listed only on Google, not on Yelp or TripAdvisor, is significant at the 95% level based on the model without serial correlation, but insignificant and substantially smaller in the model that accounts for serial correlation. This result from the model with serial correlation is consistent with that from the reduced-form analysis, as shown in Table 1.9 that Yelp does not affect restaurants that are not listed on Yelp.

Nonetheless, there is a notable difference between the two models in terms of parameter estimates; that is, the variance of the initial prior \( \sigma^2_{A} \) is estimated to be much smaller (0.954) in the model with serial correlation than the one in the model without (2.31). The smaller estimate of \( \sigma^2_{A} \) implies that consumers’ initial prior beliefs are more precise in the model with serial correlation than in the one without, and therefore, the belief updating process puts more weight on the prior and less weight on new signals, leading to a slower convergence to the true mean quality. In other words, the model that accounts for serial correlation suggests a slower learning process than that one that accounts for serial correlation.

The comparison of these models shows that the model with serial correlation is a better fit to the data. Therefore, here I take a closer look at the estimates of a few key structural parameters from that model and interpret their meaning. The estimate for the log price coefficient \( \alpha \) is 3.88. It can be used to construct the price elasticity of demand, which has the form \( 1 + \alpha(1 - s_{jt}) \), where \( s_{jt} \) is the revenue share of restaurant \( j \) at time \( t \). Since \( s_{jt} \) is usually very small for an individual restaurant (99% of restaurants in the sample have a market share less than 0.085.), the price elasticity of demand for a restaurant can be approximated by \( (1 + \alpha) \), i.e. 4.88. This elasticity implies that if a restaurant’s price declines by 10%, its demand will increase by 48.8%.

The estimates of the key learning parameters \( \kappa \) and \( \lambda \) for each group illustrate that almost all \( \kappa \) coefficients are significant, except for the fourth group of restaurants, and the coefficients for Yelp exposure are significant only for the second group of restaurants, those listed on both Yelp and TripAdvisor; furthermore, the coefficients for TripAdvisor exposure (which can be interpreted also as exposure to other online platforms) is significant for only the fourth group of restaurants, those listed only on Google, but on neither Yelp or TripAdvisor. These results are consistent with the reduced-form analysis in that Yelp affects only restaurants listed on Yelp, and the TripAdvisor exposure measure picks up the exposure to other platforms (especially Google) and thereby is shown to affect restaurants listed only on Google. What is surprising, however, is that the coefficient for Yelp exposure is small and insignificant for the first group of restaurants, those that are listed only on Yelp but not on TripAdvisor. This contradicts the reduced-form analysis results from Table 1.9 which demonstrates that the first group of restaurants’ revenues are most sensitive to Yelp exposure. At the moment, I do not have an explanation for this contradiction.

It should be noted that the estimates of \( \kappa \)’s and \( \lambda \)’s are very telling. They represent what information source is the most important in the consumer learning process. The ratio \( \frac{\lambda_{4y}\text{Yelp}_{\text{mt}}}{(\kappa_{4p} + \lambda_{4y}\text{Yelp}_{\text{mt}} + \lambda_{4y}\text{TripAdvisor}_{\text{mt}})} \) can be interpreted as what percent age of information that consumers use to learn about restaurants
in a particular month comes from Yelp for each group. Similarly, \( \frac{\lambda_{pg,\text{TripAdvisor}}}{\kappa_p + \lambda_{pg,\text{Yelp}} \text{Yelp}_{\text{mt}} + \lambda_{pg,\text{TripAdvisor}} \text{TripAdvisor}_{\text{mt}}} \) and \( \frac{\kappa_p}{\kappa_p + \lambda_{pg,\text{Yelp}} \text{Yelp}_{\text{mt}} + \lambda_{pg,\text{TripAdvisor}} \text{TripAdvisor}_{\text{mt}}} \) can be interpreted as how much information comes from other platforms and other information sources respectively. Plugging in the average of the log measure of Yelp exposure and that of TripAdvisor exposure from each group, we can calculate these percentages for each group. Table 1.11 presents the results.

Table 1.11: Distribution of Information Source

<table>
<thead>
<tr>
<th>Type</th>
<th>Yelp</th>
<th>Other Online Platforms</th>
<th>Other Info Source</th>
<th>No. of Restaurants</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) On Yelp but not TripAdvisor</td>
<td>0.16%</td>
<td>0.27%</td>
<td>99.6%</td>
<td>750</td>
</tr>
<tr>
<td>(2) On Yelp and on TripAdvisor</td>
<td>62.9%</td>
<td>0.1%</td>
<td>37.0%</td>
<td>1,438</td>
</tr>
<tr>
<td>(3) Not on Yelp but on TripAdvisor</td>
<td>0.08%</td>
<td>0.1%</td>
<td>99.8%</td>
<td>450</td>
</tr>
<tr>
<td>(4) Neither on Yelp nor on TripAdvisor</td>
<td>0.17%</td>
<td>69.3%</td>
<td>30.3%</td>
<td>884</td>
</tr>
</tbody>
</table>

As shown in the table, for restaurants that are listed on both Yelp and TripAdvisor, about 63% of the information that consumers use to learn about restaurants comes from Yelp, and about 37% are from other information sources. Other online platforms contribute very little. For restaurants that are listed neither on Yelp nor TripAdvisor, it seems that other online platforms (as represented by TripAdvisor’s penetration) account for about 69% of the information about restaurants, and other information sources contribute about 31%, and the Yelp exposure measure picks up barely any information source for learning.

The coefficients related to consumers’ prior beliefs and restaurants’ true qualities (\( \theta_p, \theta_0 \) and \( \theta_r \)) are also very informative; they tell us how much learning changes consumers’ knowledge of restaurants’ quality. As mentioned before, consumers’ prior for each restaurant is modelled as being linear to a restaurant’s rating. The coefficient associated with the rating in the prior is \( \theta_p \), and it is estimated to be \(-0.0954\), negative and significant at the 99% confidence level. The negative sign implies that consumers’ prior beliefs are somehow negatively correlated with their Google rating. This could be due to that there are unobserved time-invariant restaurant characteristics that are negatively correlated with restaurants’ ratings on Google, but do not require learning. For example, a lower-quality restaurant might have a fancy exterior, is at a great location and has a big parking lot, and thereby without learning, consumers prefer this restaurant to a higher-quality restaurant.

After learning, however, consumers seem to prefer more highly rated restaurants. This can be seen from the estimate of the slope coefficient associated with rating in the true mean quality (\( \theta_r \)), which is a positive 0.368. The difference between the true mean quality and the prior \( \theta_0 + (\theta_r - \theta_p^2) \ast R_j \) is very meaningful: it captures how much consumers have updated their beliefs during the learning process, and whether the posterior true mean quality is higher or lower than the prior for each quality class. A negative number indicates that consumers’ prior beliefs are higher than the true mean quality of the restaurants, and a positive one means that the prior beliefs are lower than the posterior. It is easy to calculate this gap for each rating class. It is shown in Table 1.12. As can be seen, for 2 and 3 star restaurants, consumers lower their beliefs through the learning process, and
for 4 and 5 star restaurants, consumers increase their expectations through learning. This pattern indeed resembles the bi-furcating result in the reduced-form analysis.

Table 1.12: Gap Between True Mean Quality and Prior

<table>
<thead>
<tr>
<th></th>
<th>2 Star</th>
<th>3 Star</th>
<th>4 Star</th>
<th>5 Star</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta^* + (\theta^r - \theta^p) \cdot R_j$</td>
<td>-0.633</td>
<td>-0.170</td>
<td>0.294</td>
<td>0.757</td>
</tr>
</tbody>
</table>

1.5.4 Counterfactual Analysis

With the structural parameter estimates in hand, I conduct a counterfactual analysis to answer three specific questions: (1) How much faster do Yelp and other online review platforms help consumers learn? (2) What is their effect on the lifetime revenue of a typical restaurant? (3) What is the effect of online review platforms on consumer welfare and market shares of restaurants? The answers to the two questions are discussed in Section 1.5.4.1 and that for the welfare question is explored in Section 1.5.4.2.

1.5.4.1 How Fast Do Online Platforms Help Consumers Learn?

To assess how fast online platforms help consumers learn and how they affect restaurants’ lifetime revenues, I simulate the expected mean quality of $\eta_{jt}$ and revenue for for a typical restaurant in an average market for four major metropolitan regions in Texas: Austin, Houston, Dallas-Fort Worth and San Antonio. I compare restaurants’ revenues and the characteristics of consumers’ beliefs of their quality under the scenarios with and without online review platforms. In this simulation exercise, I assume that the typical restaurant survives for the entire duration of the sample period since Yelp’s penetration.\footnote{Here I use August 2007 as the date for Yelp’s penetration because this is the time when Yelp started to pick up momentum in Texas, according to Figure 4. I simulate the paths of $\eta_{jt}$ and revenues for the typical restaurant from this point onwards till the end of the sample period, December 2015. Since $\eta_{jt}$ has a random component due to the uncertainty of experience signals released every period, and it is path dependent, I simulate the path of $\eta_{jt}$ 1,000 times and then take the average of all paths.} As shown very distinctly in the structural estimation, there are two groups of restaurants whose revenues are affected the most by online review platforms: those listed on all three platforms: Yelp, TripAdvisor and Google, and those listed only on Google not on Yelp or TripAdvisor. The first group is affected primarily by Yelp, and the second by other platforms. I conduct this analysis for each group separately, and the comparison between the effects of Yelp and other online platforms on both the speed of learning and restaurants’ lifetime revenues gives us an idea of their relative impact in different regions during the sample period.

Tables 1.13 and 1.14 illustrate the characteristics of a typical restaurant in each metro region and the counterfactual results regarding how much online platforms reduce the learning period and their effects on restaurant revenues. Table 1.13 demonstrates these results for Yelp’s effect on the group of restaurants that are listed on all three platforms, and Table 1.14 shows that for other platforms’ effect on the group of restaurants that are listed only on Google. As can be seen from
the tables, a typical restaurant in both groups and across all major regions has a rating of 4 stars on Google and is moderately priced with an average menu price of $20, and their cuisine types are mostly bars (and traditional American), except that in Houston, a typical restaurant in the first group is a Tex-Mex restaurant. In terms of monthly revenues, those from the first group are substantially bigger than those from the second group, with a median revenue at least $11,000 from the first group vs. $7,600 from the second group. This pattern indicates that large restaurants are listed on all three platforms, and Google picks up the smaller restaurants that are not captured by other platforms.

To assess the speed of learning, I use a rule of thumb based on the variance of consumers’ posterior beliefs (or the perception precision). During the learning process, the variance of consumers’ updated beliefs will become smaller and smaller, such that consumers’ beliefs are more and more precise. Usually, the variance reduces quickly at the beginning of the updating process, and then slows down in later periods. Therefore, an absolute convergence to the true beliefs with variance equal to 0 can take a long time. Instead, I use a rough measure of the period that it takes for the variance to reduce to half of its original size (i.e. the variance of the initial prior), and use that as a measure of the period of learning for this exercise.

As mentioned before, I simulate the revenue and consumers’ expected mean quality for a typical restaurant in each region under the scenario with and without online platforms for both groups of restaurants. How much that online platforms shorten the learning periods and improve a typical restaurant’s revenue are shown in the second half (lines 5-8) of Tables 1.13 and 1.14. As can be gleaned from the table, for the first group of restaurants, Yelp reduces the learning period by 8 to 22 months and increases a typical restaurant’s lifetime revenue by 1.7% to 3.4%; for the second group of restaurants, other online platforms, especially Google, shorten the learning period by 15 to 31 months and improve a restaurant’s revenue by about 1.5% to 2.5%. The moderate increase in restaurants’ revenues is mostly due to the fact that the quality level of a typical restaurant is 4 stars, the average of ratings on Google. According to the structural parameter estimates as shown in Table 1.12 the gap between consumers’ prior belief and the true mean quality of a 4 star restaurant is not substantially high, and thereby, the increase in restaurant revenues should be moderate.

Comparing Yelp effects on the first group and other platforms’ effect on the second group, we can see that Yelp’s effects seem to be higher in magnitude than other platforms’ effects both in terms of the speed of learning and increase in revenues. This once again is consistent with the general observation that Yelp is the dominant review platform in the restaurant industry.

1.5.4.2 Welfare Effect of Online Review Platforms

In this counterfactual analysis, I assess the effects of Yelp and other platforms on consumer welfare by holding the supply side constant; that is, I do not consider online platforms’ effects on restaurants’ prices and their exit behaviors. Holding prices constant is an innocuous assumption in this industry and for my sample. It is easy to derive that to optimize prices, restaurants’ optimal pricing strategy is

$$p_{jt} = c_{jt}(1 + \frac{1}{\alpha(1-s_{jt})}),$$

where $c_{jt}$ is the marginal cost at restaurant $j$ at time $t$. Since the
Table 1.13: Yelp’s Effect on Learning and Life-Time Revenue of a Typical Restaurant in Selected Regions

<table>
<thead>
<tr>
<th></th>
<th>Austin</th>
<th>Houston</th>
<th>Dallas-Fort Worth</th>
<th>San Antonio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Price ($)</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Median Google Rating</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Mode Cuisine Category</td>
<td>Bar</td>
<td>Tex-Mex</td>
<td>Bar</td>
<td>Bar</td>
</tr>
<tr>
<td>Median Monthly Revenue ($)</td>
<td>21,000</td>
<td>15,000</td>
<td>11,000</td>
<td>19,000</td>
</tr>
<tr>
<td>Learning Period With Yelp (months)</td>
<td>9</td>
<td>22</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>Learning Period Without Yelp (months)</td>
<td>21</td>
<td>37</td>
<td>40</td>
<td>23</td>
</tr>
<tr>
<td>Difference in Learning Periods (months)</td>
<td>12</td>
<td>15</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>Increase in Revenue</td>
<td>3.38%</td>
<td>2.24%</td>
<td>3.10%</td>
<td>1.69%</td>
</tr>
</tbody>
</table>

Note: Monthly revenues are in December 2000 dollars.

Table 1.14: Other Platforms’ Effect on Learning and Life-Time Revenue of a Typical Restaurant in Selected Regions

<table>
<thead>
<tr>
<th></th>
<th>Austin</th>
<th>Houston</th>
<th>Dallas-Fort Worth</th>
<th>San Antonio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Price ($)</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Median Rating</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Mode Cuisine Category</td>
<td>Bar</td>
<td>Bar</td>
<td>Bar</td>
<td>Bar</td>
</tr>
<tr>
<td>Median Monthly Revenue ($)</td>
<td>13,000</td>
<td>8,700</td>
<td>7,600</td>
<td>16,000</td>
</tr>
<tr>
<td>Learning Period With Other Platforms (months)</td>
<td>9</td>
<td>14</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Learning Period Without Other Platforms (months)</td>
<td>24</td>
<td>44</td>
<td>45</td>
<td>23</td>
</tr>
<tr>
<td>Difference in Learning Periods (months)</td>
<td>15</td>
<td>30</td>
<td>31</td>
<td>15</td>
</tr>
<tr>
<td>Increase in Revenue</td>
<td>1.53%</td>
<td>2.45%</td>
<td>2.52%</td>
<td>1.52%</td>
</tr>
</tbody>
</table>

Note: Monthly revenues are in December 2000 dollars.

The restaurant industry is highly competitive, the revenue share $s_{jt}$ is usually very small for an individual restaurant. As mentioned previously, 99% of restaurants in the sample have a market share less than 0.085 with the median market share at 0.004. Therefore, even though online review platforms may affect restaurants’ revenue shares, the change in prices for these restaurants would not be big. As for endogenous exit behaviors of firms due to exposure on online platforms, although my previous analysis regarding restaurant exits in Section 1.4.2.2 demonstrates that Yelp affects the exit probabilities of restaurants, it cannot not pin down the effect at each restaurant’s age. Given the heterogenous effects of online review platforms on restaurants’ revenue across age, their effects on the exit rates of restaurants should also depend on restaurants’ age. To account for the age specific effect, a structural model that combines consumer learning on the demand side and dynamic oligopoly games with entry and exit is needed, and that is beyond the scope of the current paper. Therefore, in this welfare analysis, I abstract from the welfare effects that arise from the exits of restaurants, with the acknowledgement that without accounting for that aspect, my welfare effect estimates here should be seen as the lower bound.

To compare consumer welfare in the world with the existence of online review platforms and that without, I transform the welfare calculations into dollar amount by using the concept of compensating variation. In particular, I calculate how much consumers’ income would increase or decrease

\[ \frac{1}{1-s_{jt}} \] will not change much.

\[ \text{Since I use a constant expenditure model, where prices come into the indirect utility in log form, the consumer} \]

Because
such that their utilities in the counterfactual world without online platforms would be brought up to the same level as those in the real world. Let $\psi$ denote the percentage change in consumers’ income in the counterfactual world to maintain the same level of utility. I derive $\psi$ and it has the following expression:

$$\psi = \exp \left( \frac{\sum_{t=1}^{T} \sum_{m=1}^{N_m} (CS_{mt} - CS^c_{mt}) \gamma}{\alpha T \cdot N_m} \right) - 1$$  \hspace{1cm} (1.5.6)$$

where $CS^c_{mt}$ and $CS_{mt}$ are the expected consumer surplus at time $t$ in market $m$ in the counterfactual world and the real world respectively. $T$ is the number of periods in the data, and $N_m$ is the total number of geographic markets.\(^{28}\)

This derivation of this formula is based on the concept of inclusive values, which is as a measure for consumer welfare, as in the standard logit demand models. However, there is an important difference in my demand model with consumer learning: in the standard logit demand model, consumers make choices of restaurants based on their true beliefs of restaurants’ quality, and thereby their choices are optimal. In a model with consumer learning, on the other hand, consumers are making choices based on their prior (often biased) beliefs of restaurants’ quality, and thereby not optimal at least at the beginning of the learning period. As they discover the true mean quality of restaurants, they adjust their choices to become close to the optimal ones. This is where the gain in consumer welfare comes from in this analysis. Since online review platforms help consumers learn, consumers are able to make optimal choices of restaurants faster than before. Therefore, their welfare would increase as a result of their exposure to online review platforms.

Let $\Delta_{jt}$ denote the expected utility that consumers base their choices on and $\Delta^T_{jt}$ the actual utility that consumers receive when visiting restaurant $j$ at time $t$. Then based on the analysis done by Lewis and Zervas (2016), the expected consumer utility in market $m$ at time $t$ is

$$CS_{mt} = \log \left( \exp(\Delta_{0mt}) + \sum_{j \in M_{mt}} \exp(\Delta_{jt}) \right) + C + \sum_{j \in M_{mt}} s_{jt}(\Delta^T_{jt} - \Delta_{jt})$$

$$= \log \left( 1 + \sum_{j \in M_{mt}} \exp(\Delta_{jt} - \Delta_{0mt}) \right) + C + \sum_{j \in M_{mt}} s_{jt}(\Delta^T_{jt} - \Delta_{jt}) + \alpha \gamma^{-1} \log(\text{Income}_{mt})$$  \hspace{1cm} (1.5.7)$$

where $\Delta_{0mt} = \alpha \gamma^{-1} \log(y_{mt})$ is the mean utility of the outside option; $M_{mt}$ denotes the set of restaurants in market $m$ at time $t$; $C$ is the Euler constant.

Similarly, in the counterfactual world without online review platforms, the expected consumer utility cannot be interpreted in dollar amount as in the case of the standard logit demand model. However, the presence of $\alpha \gamma^{-1} \log(y_{mt})$ in the indirect utility function (as shown in equation 1.2.7 in Section 3.2) makes it feasible to translate changes in consumers’ utility into dollar amounts by using the compensating variation principle.\(^{28}\)

If a geographic market has no restaurants at time $t$, then the expected utility for that market at time $t$ is 0.
utility can be written as

$$\text{CS}_{mt}^c = \log\left(1 + \sum_{j \in M_{mt}} \exp(\Delta_{jt}^c - \Delta_{0mt})\right) + C + \sum_{j \in M_{mt}} s_{jt}(\Delta_{jt}^T - \Delta_{jt}^c) + \alpha \gamma^{-1} \log(\text{Income}_{mt})$$

(1.5.8)

where $\Delta_{jt}^c$ denote the expected utility that consumers base their choices on in the counterfactual world.

The total expected utility in the real world over the sample period is then $\sum_{t=1}^{T} \sum_{m=1}^{N_m} \text{CS}_{mt}$, and that for the counterfactual world is $\sum_{t=1}^{T} \sum_{m=1}^{N_m} \text{CS}_{mt}^c$. To bring consumers’ expected utility in the counterfactual world to the same level as in the real world, we can multiply the income in the counterfactual world by $(\psi + 1)$. Then we have

$$\sum_{t=1}^{T} \sum_{m=1}^{N_m} \text{CS}_{mt} = \sum_{t=1}^{T} \sum_{m=1}^{N_m} (\text{CS}_{mt}^c - \alpha \gamma^{-1} \log(y_{mt}) + \alpha \gamma^{-1} \log((\psi + 1)y_{mt}))$$

$$= \sum_{t=1}^{T} \sum_{m=1}^{N_m} (\text{CS}_{mt}^c + \alpha \gamma^{-1} \log(\psi + 1))$$

(1.5.9)

Rearranging this equation gives us the expression for $\psi$ as shown in equation [1.5.6]. The detailed technical issues related to calculating $\text{CS}_{mt}$, $\text{CS}_{mt}^c$ and the counterfactual revenue shares can be found in Section 7 of the Appendix.

The results for the welfare analysis are shown in Table 1.15 below. As can be seen, the estimated welfare effect of online review platforms in terms of percentage increase in consumer monthly income ($\psi$) is about 0.23%. This number means that consumers are willing to give up about only 0.23% of their monthly income to pay for having online review platforms. However, since eating out accounts for only 4% of consumers’ total income, this percentage increase is more meaningful in the context of consumers’ total budget on eating out. Dividing this percentage by $\gamma$, the budget share for eating out, gives 5.4%, indicating that consumers are willing to use 5.4% of their monthly budget on restaurants to pay for having online review platforms. These percentage changes translate into about $15.2 million in consumers’ income for over 400 zipcode markets, where there were restaurants listed on online review platforms, during the sample period.

Table 1.15: Welfare Analysis

<table>
<thead>
<tr>
<th>Baseline Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi$</td>
</tr>
<tr>
<td>$\psi/\gamma$</td>
</tr>
<tr>
<td>Dollar Amount ($ million, in Aug 2017 dollars)</td>
</tr>
</tbody>
</table>

To assess how online platforms affect total restaurant revenues in the sample, I calculated the total change in the sum of their revenues from the counterfactual world without online reviews to the real world with online reviews. This change in total revenues is $23.2 million, representing about
5.9% increase in industry revenues. However, not all restaurants are winners. As can be seen from Table 1.12 for lower quality restaurants with 2-3 star ratings, consumers’ beliefs of their quality decline through learning, whereas the opposite is true for higher quality restaurants with 4-5 star ratings. This belief updating process implies that online review platforms redistribute demand from lower-quality restaurants to higher-quality restaurants during the sample period. An analysis of the change in revenue shares of restaurants at various quality levels from the counterfactual world to the real world shows that the revenue shares of lower-quality restaurants indeed declined, and those of higher-quality restaurants increased. These changes are shown in Table 1.16.

Table 1.16: Percentage Change in Revenue Shares of a Restaurant

<table>
<thead>
<tr>
<th>Quality</th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-star</td>
<td>-43.5%</td>
<td>-84.4%</td>
<td>0.128%</td>
</tr>
<tr>
<td>3-star</td>
<td>-6.20%</td>
<td>-64.9%</td>
<td>21.0%</td>
</tr>
<tr>
<td>4-star</td>
<td>16.3%</td>
<td>-0.770%</td>
<td>38.4%</td>
</tr>
<tr>
<td>5-star</td>
<td>22.5%</td>
<td>-0.332%</td>
<td>45.1%</td>
</tr>
</tbody>
</table>

As can be seen, a low quality 2-star restaurant’s revenue share can decline by as high as 84.4% from a world without online review platforms to one with. For a high quality 5-star restaurant, the revenue share can increase by as much as 45.1%. The signs of these percentage changes in revenue shares are consistent with the reduced-form analysis on the bifurcating effect of online review platforms. These numbers show that online review platforms help consumers make better choices by directing them to higher-quality restaurants.

In conclusion, the counterfactual analysis shows that online review platforms’ effect on consumer welfare is equivalent to increasing consumers’ budget on eating out by 5.4%, which translates into about $15.2 million for over 400 zipcode markets in Texas for the period of 2005-2015. Online platforms also increases the total industry revenues by 5.9%, which is equivalent to about $23.2 million, during the same period.

1.6 Conclusion

In this paper, I study the effect of online review platforms on consumer learning and develop a structural demand model to quantify the effects on consumer learning, dynamics of market share, and consumer welfare. Using a novel dataset containing full-service restaurants in Texas, consumer search interest on Yelp and TripAdvisor, and online review information from Google, Yelp and TripAdvisor, I show that consumer learning is important in the restaurant industry and that online review platforms speed up the learning process. Their effects on learning show up in restaurant revenues and survival probabilities. Specifically, doubling consumers’ exposure to Yelp, the dominant platform, increases the revenue of a high-quality new independent restaurant by 8-20% and decreases that of a low-quality restaurant by about the same amount. Doubling Yelp exposure also raises the survival rate of a new high-quality independent restaurant by 7-19 basis points and
reduces that of a low-quality restaurant by a similar level. Other platforms, especially Google, have similar effects but with smaller magnitudes. In contrast, online platforms do not affect the revenues or survival rates of chains and old independent restaurants. Counterfactual analysis based on a structural demand model of differentiated restaurants with consumer learning shows that online review platforms speed up the learning process by 0.5 to 2.5 years, increase consumer welfare by 5.4% and the total industry revenue by 5.9% during the period of 2005-2015.

There are a few limitations in my current study. First, consumer learning is inferred from the bi-furcating pattern in the revenue-age profiles, but not directly observed. This limitation lies mostly in the data available for this study. The data used here is the aggregate revenue at the restaurant level. An ideal dataset for learning is a time series of individual consumer’s expenditure at a restaurant and their search activities on online review platforms. Unfortunately that type of granular data is difficult to obtain. The bi-furcating pattern of the revenue-age profiles and the bi-furcating effects of online platforms on revenues imply that consumer learning exists and that online reviews affect consumer learning, but that mechanism could be confounded with other forces at work. The most plausible one is the managerial effort story: once managers find out that their restaurants are of high quality, they continuously invest in quality, but if they discover that their restaurants are lower quality, they disinvest in quality and only reduce price. I have tried to use the Yelp rating history data to detect this mechanism, but due to the serial correlation found in ratings, I cannot find a systematic pattern of ratings declining with age for low-quality restaurants and increasing for higher quality restaurants to validate the existence of that mechanism. For the lack of evidence supporting the managerial effort story, I abstract from that mechanism in my current study, but will explore it further in future research. Second, there is another limitation in the current welfare analysis: it is a lower bound of welfare estimate because I have not accounted for the effect of online review platforms on the entry and exit behaviors of restaurants. Once that aspect is incorporated, the welfare effect of online review platforms can be even more substantial. I intend to conduct this analysis in future research which models the entry and exit behaviors of restaurants in a dynamic oligopoly game. Despite these limitations, I believe that this paper tells a compelling story that online review platforms have an important impact on the restaurant industry, and by providing information on products, they are welfare improving.
Bibliography


The Economist, May 6, 2017, The world’s most valuable resource is no longer oil, but data.


.1 Indirect Utility in Constant Expenditure Model

In this section, I derive the indirect utility function in a constant expenditure model. This derivation is based on Hendel (1999) and Dubé (2004). Let $C_0$ denote the quantity of goods outside of the product category we are interested in, and let the price of those goods be normalized to 1. In the context of my study, $C_0$ would include any consumption outside of eating out at full-service restaurants. Within the full-service restaurant category, there are $J$ restaurants, and they come into a consumer’s utility function in the following way:

$$\max_{C_0, X_j} U = S_c C_0^{1-\gamma} \left( \sum_{j=1}^{J} W_j X_j \right)^{\gamma} \exp(\varepsilon)$$

s.t. $C_0 + \sum_{j=1}^{J} P_j X_j = Y$ \hspace{1cm} (.1.1)

where $W_j$ is product characteristics of alternative $j$, $\varepsilon$ follows an extreme value type I distribution. $S_c$ is consumer characteristics. Since $\left( \sum_{j=1}^{J} W_j X_j \right)^{\gamma}$ specifies that all alternatives are perfect substitutes to each other, consumers should choose only one alternative. Assume that $\varepsilon$ is a idiosyncratic shock associated with each choice alternative, then we can solve the above maximization problem backwards. That is, for a given product $j$, we derive the indirect utility function $U_j^*$ from utility maximization and then compare the indirect utility for each product in the choice set and choose $U_k^* = \max(U_1^*, \ldots, U_J^*)$.

The derivation of the indirect utility function is shown as follows:

$$\max_{C_0, X_j} U = S_c C_0^{1-\gamma} (W_j X_j)^{\gamma} \exp(\varepsilon_j)$$

s.t. $C_0 + P_j X_j = Y$ \hspace{1cm} (.1.3)

The standard Cobb-Douglas utility function maximization gives:

$$C_0 = (1-\gamma)Y$$ \hspace{1cm} (.1.4)

$$X_j = \frac{\gamma Y}{P_j}$$ \hspace{1cm} (.1.5)

The indirect utility $U_j^*$ is then

$$U_j^* = S_c W_j^{\gamma} \exp(\varepsilon_j) (1-\gamma)^{1-\gamma} Y^{1-\gamma} \left( \frac{\gamma Y}{P_j} \right)^{\gamma}$$

$$= S_c W_j^{\gamma} \exp(\varepsilon_j) (1-\gamma)^{1-\gamma} (\gamma Y P_j^{-\gamma})$$ \hspace{1cm} (.1.6)

$$= S_c W_j^{\gamma} \exp(\varepsilon_j) (1-\gamma)^{1-\gamma} Y P_j^{-\gamma}$$ \hspace{1cm} (.1.7)
Taking the log on both sides gives

\[ \ln(U^*_j) = \ln(S_c) + \gamma \ln(W_j) + \theta_c + \ln(Y) - \gamma \ln(P_j) + \varepsilon_j \]  

where \( \theta_c \) is a constant.

The indirect utility function used in this paper is the log utility above times a multiplier \( \alpha/\gamma \), i.e. \( \alpha/\gamma \ln(U^*_j) \). The log form of the indirect utility function is convenient because it linearizes the different component that enters into consumers’ utility. For example, \( \ln(S_c) \) and \( \ln(W_j) \) correspond to the \( X_{jt} \theta_x \) component in equation 1.2.1. The unknown quality of a restaurant \( \tilde{A}_{ijt} \) in equation 1.2.1 is part of \( \ln(W_j) \) in equation 1.8.

\section{Data Related to Market Demands and Costs}

For demographics and income, my data come mostly from the 1990-2010 decennial censuses. To control for the census geographic boundary definition changes during the three decennial census periods, I use GeoLytics’ harmonized census dataset, the Neighborhood Change Database (NCDB) Tract Data from 1970-2010. This database adjusts earlier censuses to 2010 census geography, making feasible the intertemporal comparison of demographic changes in a given area. The geography in this database can be organized into census tracts, zip code tabulation areas, or counties. In this study, I define markets at the zip code tabulation area level. For the intercensal and postcensal periods, I use US Census Bureau’s intercensal estimates and American Community Survey. Since the intercensal estimates and some American Community Survey data are available only at the county level, I distributed them to the zip code tabulation area level based on historical trends from the census. Given that restaurants receive more revenues in tourist peak seasons, I control for demands from tourists by collecting Texas annual Visitor Spending data at the county and city levels from 1995 to 2015 from Texas Economic Development & Tourism. In addition, I also collected traffic volume data and Texas road network GIS shapefiles from the Department of Transportation of Texas. For full-service restaurants, traffic volumes are often one of the most direct demand indicators. The traffic volume data includes annual average daily traveler counts for about 31,400 traffic monitoring stations in Texas and covers the years from 1999 to 2014.

For the wage data in the food service industry, I use the Occupational Employment Statistics provided by the Bureau of Labour Statistics. This data cover 26 metropolitan areas in Texas and have an annual frequency and date back to 1997. I interpolate the data to a monthly level to match my market definition. The consumer expenditure data come from the Consumer Expenditure Survey, which is also provided by the Bureau of Labour Statistics. This data cover the South Region of the United States, including Texas. It has an annual frequency and dates back to 1995. Again, I interpolate this data to a monthly level.
.3 Heckman’s Correction and Other Robustness Checks

It is important to note that the results shown previously do not correct for selection due to endogenous exit. To deal with selection, I again use Heckman’s correction. To control for restaurant fixed effects in the first-stage probit regression, I use the method introduced in Wooldridge (1995), which includes the whole history of control variables as regressors. The specification for the first-stage probit regression is very similar to equation 1.4.3, only that the dependent variable is $action_{jt}$ instead of $log(Rev_{jt})$ and that I add wage $w_{mt}$ and the history of other independent variables as additional controls. Given the limited number of exits observed for Yelp listed restaurants (about 660 out of 5930 restaurants), I run regression using only Google and Yelp November 2016 ratings as the quality measures as the sample associated with “predetermined” ratings does not have many exits. In addition, as a robustness check, I include region $\times$ time FEs to capture any type of time varying factors that are specific for each region. The raw regression results for using Google and Yelp ratings are shown in Tables 19 and 18 respectively. The main results do not change substantially from the ones that do not control for selection. Below I present the translated results for the effects of Yelp on revenues for given rating classes in Table 17. As can be seen in the table, once we control for selection and time $\times$ metro FEs, the results do not change much, only that once we add time $\times$ metro FEs, the significance levels of the estimated effects reduce. Nonetheless, the magnitude of the effects and the significance levels for very high and very low quality restaurants stay almost the same.

.4 Placebo Test

As mentioned in Section 1.4.2.1 an important concern regarding the results is the endogeneity of the Yelp exposure measure even after controlling for the rich set of fixed effects and selection. In particular, Yelp exposure could be correlated with the inherent ability or willingness to learn about restaurant qualities in each region. To examine this potential source of endogeneity, I conduct a placebo test to see if Yelp exposure has an effect on restaurant revenues before Yelp’s penetration. If Yelp exposure is correlated with people’s inherent interest in learning about restaurant qualities across geographic regions, then before Yelp’s penetration, learning should still be faster in those regions where Yelp became much more popular later on. Therefore, if we move the treatment, i.e. the Yelp exposure, to the pre-treatment period, we should expect to see that the treatment has an effect on restaurant revenues in that period.

Based on this logic, I date Yelp exposure data to 10 years earlier to a start date of March 1995, and then run the regression 1.4.3 using the sample of restaurant revenues for the period between March 1995 to March 2005. I again use both Google ratings and Yelp ratings as quality measures.

---

29The Wooldridge method works well only if the number of observations is much larger than the number of time periods in the panel data; otherwise, the number of independent variables could exceed or become very close to the number of observations, rendering the probit regression invalid. In my application, some restaurants have a very long time horizon. To solve the problem, I cut my sample into many 12-month periods and include the history of controls during the 12 months.
Chapter 61

Table 17: Effects of Yelp Exposure on Revenue by Quality Level With Heckman Correction

<table>
<thead>
<tr>
<th>Star Rating</th>
<th>Google</th>
<th>Google</th>
<th>Yelp</th>
<th>Yelp</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-.260***</td>
<td>-.233**</td>
<td>-.155***</td>
<td>-.174***</td>
</tr>
<tr>
<td></td>
<td>(.0705)</td>
<td>(.103)</td>
<td>(.0468)</td>
<td>(.0688)</td>
</tr>
<tr>
<td>3</td>
<td>-.109***</td>
<td>-.0665</td>
<td>-.0436**</td>
<td>-.0502</td>
</tr>
<tr>
<td></td>
<td>(.0377)</td>
<td>(.0831)</td>
<td>(.0251)</td>
<td>(.0548)</td>
</tr>
<tr>
<td>4</td>
<td>.0413**</td>
<td>.100</td>
<td>.0682**</td>
<td>.0731</td>
</tr>
<tr>
<td></td>
<td>(.0203)</td>
<td>(.0753)</td>
<td>(.0273)</td>
<td>(.0542)</td>
</tr>
<tr>
<td>5</td>
<td>.184***</td>
<td>.267***</td>
<td>.180***</td>
<td>.196***</td>
</tr>
<tr>
<td></td>
<td>(.0490)</td>
<td>(.0827)</td>
<td>(.0504)</td>
<td>(.0672)</td>
</tr>
</tbody>
</table>

Time FE  | Yes  | Yes  | Yes  | Yes  |
Time FE×Chain Dummy | Yes  | Yes  | Yes  | Yes  |
Time ×Metro FE  | No   | Yes  | No   | Yes  |
Restaurant FE | Yes  | Yes  | Yes  | Yes  |

Sample Entry after Yelp penetration

\[ N \]
131,138 131,024 158,817 158,724

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

The pre-determined ratings cannot be used here because the sample associated with it does not exist before March 2005. The regression results are shown in Table 20. As can be seen, none of the coefficients associated with the Yelp exposure is significant. These results indicate that this source of endogeneity for the Yelp exposure measure is not of great concern.

4.1 Other Channels of Yelp Effect Besides Learning

The results from Section 3.6 are consistent with the implications of a model of consumer learning. However, consumer learning may not be the only channel through which Yelp has an effect on revenues. In this section, I discuss three other potential channels of Yelp effect, including managerial learning, consumer sorting, herding and online reputation. I argue that none of these three other channels are important.

For the role of managerial learning, I focus on the aspect that managers might be able to receive more feedback by reading consumers’ reviews on Yelp and thereby improve their services and food qualities over time. As shown by Gin and Leslie (2003), greater information exposure incentivizes restaurants to improve their hygiene levels. If managers improve quality in response to Yelp reviews, we should see the average of new ratings that a restaurant receives per month goes up over time. I test this hypothesis by regressing a restaurant ratings on its age, and I use the following specification:

\[
\log(ANR_{jt}) = \theta_a \log(a_{jt}) + \theta_{ac} \log(a_{jt}) \times D_j + \theta_y \log(Yelp_{mt}) + \theta_{yc} \log(Yelp_{mt}) \times D_j + X_{mt} \theta_x + \theta_{nt} n_{mt} + \theta_{nc} n_{mt}^C + \theta_t + \theta_{tc} D_t + \theta_j + \xi_{jt}
\]  

(4.1)
Table 18: Effects of Yelp Exposure on Revenue With Heckman Correction (Google Ratings)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Yelp ($\theta_y$)</td>
<td>-0.166</td>
<td>-0.560***</td>
<td>-0.567***</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
<td>(0.140)</td>
<td>(0.158)</td>
</tr>
<tr>
<td>Log Yelp×Rating ($\theta_{yr}$)</td>
<td>0.0187*</td>
<td>0.150***</td>
<td>0.167***</td>
</tr>
<tr>
<td></td>
<td>(0.0259)</td>
<td>(0.0357)</td>
<td>(0.0346)</td>
</tr>
<tr>
<td>Log Yelp×Chain Dummy ($\theta_{yc}$)</td>
<td>-0.0807</td>
<td>-0.0319</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>(0.0542)</td>
<td>(0.0655)</td>
<td>NA</td>
</tr>
<tr>
<td>Log Yelp×Chain Dummy×Rating ($\theta_{ycr}$)</td>
<td>0.0392</td>
<td>0.0195</td>
<td>0.0393</td>
</tr>
<tr>
<td></td>
<td>(0.0149)</td>
<td>(0.0174)</td>
<td>(0.0185)</td>
</tr>
<tr>
<td>Traffic (thousands)</td>
<td>-0.0265</td>
<td>-0.0454</td>
<td>-0.00892</td>
</tr>
<tr>
<td></td>
<td>(0.0872)</td>
<td>(0.102)</td>
<td>(0.108)</td>
</tr>
<tr>
<td>Log Total Population</td>
<td>0.310*</td>
<td>-0.269</td>
<td>-0.326</td>
</tr>
<tr>
<td></td>
<td>(0.170)</td>
<td>(0.242)</td>
<td>(0.259)</td>
</tr>
<tr>
<td>Income (millions)</td>
<td>-1.43</td>
<td>1.69</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>(1.68)</td>
<td>(1.91)</td>
<td>(1.97)</td>
</tr>
<tr>
<td>Log Visitor Spending</td>
<td>0.473***</td>
<td>0.381***</td>
<td>0.0180</td>
</tr>
<tr>
<td></td>
<td>(0.157)</td>
<td>(0.140)</td>
<td>(0.194)</td>
</tr>
<tr>
<td>Population Density</td>
<td>-158.7</td>
<td>-114.2</td>
<td>-61.32</td>
</tr>
<tr>
<td></td>
<td>(107.4)</td>
<td>(151.7)</td>
<td>(154.1)</td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>-0.312</td>
<td>0.769</td>
<td>0.691</td>
</tr>
<tr>
<td></td>
<td>(0.743)</td>
<td>(1.176)</td>
<td>(1.184)</td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>-0.618</td>
<td>0.888</td>
<td>1.362</td>
</tr>
<tr>
<td></td>
<td>(0.794)</td>
<td>(1.040)</td>
<td>(1.052)</td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>0.228</td>
<td>-1.166</td>
<td>-1.655</td>
</tr>
<tr>
<td></td>
<td>(1.195)</td>
<td>(1.771)</td>
<td>(1.868)</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>-0.0993</td>
<td>0.0658</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>(0.312)</td>
<td>(0.312)</td>
<td>(0.333)</td>
</tr>
<tr>
<td>Share of White population</td>
<td>0.250</td>
<td>0.0116</td>
<td>0.0927</td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
<td>(0.162)</td>
<td>(0.178)</td>
</tr>
<tr>
<td>Share of Black population</td>
<td>-1.008**</td>
<td>0.950</td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.618)</td>
<td>(0.630)</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>-0.331</td>
<td>-1.287***</td>
<td>-1.089**</td>
</tr>
<tr>
<td></td>
<td>(0.698)</td>
<td>(0.483)</td>
<td>(0.498)</td>
</tr>
<tr>
<td>Number of Independent Rivals</td>
<td>-0.00243</td>
<td>-0.00224</td>
<td>-0.00229</td>
</tr>
<tr>
<td></td>
<td>(0.00215)</td>
<td>(0.00240)</td>
<td>(0.00241)</td>
</tr>
<tr>
<td>Number of Chain Rivals</td>
<td>-0.00112</td>
<td>-0.000901*</td>
<td>-0.00872</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
<td>(0.00563)</td>
<td>(0.00592)</td>
</tr>
<tr>
<td>Inverse Mills Ratio</td>
<td>-0.177</td>
<td>-0.0891</td>
<td>-0.0793</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.115)</td>
<td>(0.118)</td>
</tr>
<tr>
<td>Cuisine Dummy × Demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE×Chain Dummy</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time ×Metro FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Restaurant FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample</td>
<td>Entry 2 years before Yelp penetration</td>
<td>Entry after Yelp penetration</td>
<td>Entry after Yelp penetration</td>
</tr>
<tr>
<td>N</td>
<td>116,831</td>
<td>131,138</td>
<td>131,024</td>
</tr>
</tbody>
</table>

Standard errors in parentheses are clustered by restaurant.

*p < 0.10, ** p < 0.05, *** p < 0.01
Table 19: Effects of Yelp Exposure on Revenue With Heckman Correction (Yelp Ratings)

<table>
<thead>
<tr>
<th></th>
<th>(1) Log Revenue</th>
<th>(2) Log Revenue</th>
<th>(3) Log Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Yelp ($\theta_y$)</td>
<td>-0.0376</td>
<td>-0.379***</td>
<td>-0.420***</td>
</tr>
<tr>
<td></td>
<td>(0.138)</td>
<td>(0.102)</td>
<td>(0.115)</td>
</tr>
<tr>
<td>Log Yelp x Rating ($\theta_{yr}$)</td>
<td>-0.000344</td>
<td>0.122***</td>
<td>0.123***</td>
</tr>
<tr>
<td></td>
<td>(0.0446)</td>
<td>(0.0290)</td>
<td>(0.0287)</td>
</tr>
<tr>
<td>Log Yelp x Chain Dummy ($\theta_{yc}$)</td>
<td>-0.00818</td>
<td>0.00535</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>(0.0940)</td>
<td>(0.0428)</td>
<td>NA</td>
</tr>
<tr>
<td>Log Yelp x Chain Dummy x Rating ($\theta_{ycr}$)</td>
<td>-0.00558</td>
<td>0.00779</td>
<td>0.00838</td>
</tr>
<tr>
<td></td>
<td>(0.0301)</td>
<td>(0.0139)</td>
<td>(0.0146)</td>
</tr>
<tr>
<td>Traffic (thousands)</td>
<td>-0.0323</td>
<td>0.0247</td>
<td>0.00651</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.0062)</td>
<td>(0.0029)</td>
</tr>
<tr>
<td>Log Total Population</td>
<td>0.267</td>
<td>-0.247</td>
<td>-0.344</td>
</tr>
<tr>
<td></td>
<td>(0.248)</td>
<td>(0.236)</td>
<td>(0.258)</td>
</tr>
<tr>
<td>Income (millions)</td>
<td>3.35</td>
<td>-1.79</td>
<td>-2.26</td>
</tr>
<tr>
<td></td>
<td>(3.90)</td>
<td>(2.54)</td>
<td>(2.71)</td>
</tr>
<tr>
<td>Log Visitor Spending</td>
<td>-0.136</td>
<td>0.340**</td>
<td>-0.00144</td>
</tr>
<tr>
<td></td>
<td>(0.162)</td>
<td>(0.149)</td>
<td>(0.188)</td>
</tr>
<tr>
<td>Population Density</td>
<td>-8.602</td>
<td>1.187</td>
<td>55.38</td>
</tr>
<tr>
<td></td>
<td>(215.7)</td>
<td>(157.7)</td>
<td>(164.6)</td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>3.172*</td>
<td>0.0382</td>
<td>-0.0687</td>
</tr>
<tr>
<td></td>
<td>(1.871)</td>
<td>(1.51)</td>
<td>(1.189)</td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>1.518</td>
<td>2.562</td>
<td>2.717</td>
</tr>
<tr>
<td></td>
<td>(2.478)</td>
<td>(2.050)</td>
<td>(1.904)</td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>2.962*</td>
<td>-1.626</td>
<td>-1.866</td>
</tr>
<tr>
<td></td>
<td>(1.659)</td>
<td>(1.451)</td>
<td>(1.596)</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>-1.025</td>
<td>0.0830</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>(0.660)</td>
<td>(0.304)</td>
<td>(0.320)</td>
</tr>
<tr>
<td>Share of White population</td>
<td>0.826**</td>
<td>-0.143</td>
<td>-0.0292</td>
</tr>
<tr>
<td></td>
<td>(0.237)</td>
<td>(0.170)</td>
<td>(0.178)</td>
</tr>
<tr>
<td>Share of Black population</td>
<td>1.425*</td>
<td>0.573</td>
<td>0.340</td>
</tr>
<tr>
<td></td>
<td>(0.840)</td>
<td>(0.601)</td>
<td>(0.608)</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>1.519</td>
<td>-1.009**</td>
<td>-0.703</td>
</tr>
<tr>
<td></td>
<td>(1.024)</td>
<td>(0.496)</td>
<td>(0.526)</td>
</tr>
<tr>
<td>Number of Independent Rivals</td>
<td>-0.00520*</td>
<td>-0.00506***</td>
<td>-0.00469**</td>
</tr>
<tr>
<td></td>
<td>(0.00288)</td>
<td>(0.00218)</td>
<td>(0.00236)</td>
</tr>
<tr>
<td>Number of Chain Rivals</td>
<td>-0.0181**</td>
<td>-0.00500</td>
<td>-0.00437</td>
</tr>
<tr>
<td></td>
<td>(0.00791)</td>
<td>(0.00516)</td>
<td>(0.00550)</td>
</tr>
<tr>
<td>Inverse Mills Ratio</td>
<td>-0.674</td>
<td>-0.558</td>
<td>-0.548</td>
</tr>
<tr>
<td></td>
<td>(0.553)</td>
<td>(0.484)</td>
<td>(0.493)</td>
</tr>
<tr>
<td>Cuisine Dummy x Demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE x Chain Dummy</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time x Metro FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Restaurant FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Sample: Entry 2 years before Yelp penetration | Entry after Yelp penetration | Entry after Yelp penetration

N: 139,311 | 158,817 | 158,724

Standard errors in parentheses are clustered by restaurant.

* p<0.10, ** p<0.05, *** p<0.01
### Table 20: Placebo Test Results

<table>
<thead>
<tr>
<th></th>
<th>(1) Log Revenue</th>
<th>(2) Log Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Yelp ($\theta_y$)</td>
<td>0.344</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
<td>(0.175)</td>
</tr>
<tr>
<td>Log Yelp×Rating ($\theta_{yr}$)</td>
<td>-0.0747</td>
<td>-0.0254</td>
</tr>
<tr>
<td></td>
<td>(0.0563)</td>
<td>(0.0483)</td>
</tr>
<tr>
<td>Log Yelp×Chain Dummy ($\theta_{yc}$)</td>
<td>-0.0385</td>
<td>0.0707</td>
</tr>
<tr>
<td></td>
<td>(0.0977)</td>
<td>(0.0692)</td>
</tr>
<tr>
<td>Log Yelp×Chain Dummy×Rating ($\theta_{ycr}$)</td>
<td>0.0224</td>
<td>-0.0103</td>
</tr>
<tr>
<td></td>
<td>(0.0274)</td>
<td>(0.0175)</td>
</tr>
<tr>
<td>Traffic (thousands)</td>
<td>-0.149</td>
<td>-0.197</td>
</tr>
<tr>
<td></td>
<td>(0.190)</td>
<td>(0.210)</td>
</tr>
<tr>
<td>Log Total Population</td>
<td>0.0404</td>
<td>0.305</td>
</tr>
<tr>
<td></td>
<td>(0.217)</td>
<td>(0.391)</td>
</tr>
<tr>
<td>Income (millions)</td>
<td>10.4</td>
<td>7.54</td>
</tr>
<tr>
<td></td>
<td>(7.58)</td>
<td>(8.02)</td>
</tr>
<tr>
<td>Log Visitor Spending</td>
<td>-0.152</td>
<td>-0.326</td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
<td>(0.294)</td>
</tr>
<tr>
<td>Population Density</td>
<td>-55.62</td>
<td>-282.9</td>
</tr>
<tr>
<td></td>
<td>(254.1)</td>
<td>(363.7)</td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>3.684</td>
<td>3.555</td>
</tr>
<tr>
<td></td>
<td>(2.673)</td>
<td>(3.541)</td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>5.212*</td>
<td>1.991</td>
</tr>
<tr>
<td></td>
<td>(2.681)</td>
<td>(3.592)</td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>0.939</td>
<td>1.816</td>
</tr>
<tr>
<td></td>
<td>(2.455)</td>
<td>(3.390)</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>0.695</td>
<td>0.0114</td>
</tr>
<tr>
<td></td>
<td>(1.011)</td>
<td>(1.281)</td>
</tr>
<tr>
<td>Share of White population</td>
<td>1.025**</td>
<td>1.286*</td>
</tr>
<tr>
<td></td>
<td>(0.486)</td>
<td>(0.657)</td>
</tr>
<tr>
<td>Share of Black population</td>
<td>0.681</td>
<td>0.494</td>
</tr>
<tr>
<td></td>
<td>(1.077)</td>
<td>(1.327)</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>2.194</td>
<td>1.617</td>
</tr>
<tr>
<td></td>
<td>(1.653)</td>
<td>(2.335)</td>
</tr>
<tr>
<td>Number of Independent Rivals</td>
<td>-0.00540</td>
<td>-0.00704</td>
</tr>
<tr>
<td></td>
<td>(0.00554)</td>
<td>(0.00642)</td>
</tr>
<tr>
<td>Number of Chain Rivals</td>
<td>-0.0260***</td>
<td>-0.00509</td>
</tr>
<tr>
<td></td>
<td>(0.00821)</td>
<td>(0.0101)</td>
</tr>
<tr>
<td>Cuisine Dummy × Demographics</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE×Chain Dummy</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Restaurant FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>Google Rating</td>
<td>Yelp Nov 2016 Rating</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>91,963</td>
<td>56,147</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01
where $ANR_{jt}$ is the average of new ratings that restaurant $j$ receives in month $t$. The rest of the parameters and variables are defined in the same way as those in equation 1.4.3.

In this regression, $\theta_a$ and $\theta_{ac}$ are the parameters of interest. Them being positive and significant implies that there is quality improvement and thereby managerial learning. Here I include both calendar time fixed effects and their interactions with the chain dummy to control for aggregate shifts in consumer tastes that affect how Yelp reviewers rate restaurants. Furthermore, I also include restaurant fixed effects to account for restaurant-specific time-invariant characteristics that make them survive longer or exit faster. Failure to include restaurant fixed-effects can create bias in the estimation of the age coefficients $\theta_a$ and $\theta_{ac}$ since the identification of these coefficients for older ages comes from restaurants who have survived longer and tend to have higher qualities or ratings. The regression results are shown in Table 21. Columns 1 to 3 of the table are regression results without including market characteristics. Column 1 are the results for the entire sample including all restaurants listed on Yelp. Columns 2 and 3 are the results for the sample before and after Yelp penetration respectively. Columns 4 and 5 are regression results by including all market controls variables, and they are for the full sample and the sample with only new restaurants respectively.

As can be seen, none of the coefficients associated with age in these results are significant. This implies that there is no significant improvement in restaurants’ services or qualities as a result of managerial learning.

Consumer sorting is another channel through which Yelp can affect revenues. Due to taste differentials among consumers, it is possible that consumers rank the same set of restaurants differently. Information displayed on Yelp could help consumers with horizontal matching. Consumers will therefore sort themselves into their own favourite restaurants overtime, resulting in that only customers who love a restaurant would visit that restaurant. This type of sorting behavior should also cause a restaurant’s rating to increase overtime. As shown in Table 21, there is no increasing or decreasing trend in a restaurant’s ratings. This empirical finding therefore rules out the channel of consumer sorting due to horizontal taste differentials.

To investigate the channels of online reputation and herding, I use an argument: if online reputation and herding are important, then changes in Yelp rating should affect both young and old independent restaurants. Theoretically, both online reputation and herding imply that rating changes on Yelp directly affects consumers’ preferences. As a platform for reputation, Yelp would encourage consumers to go to higher rated restaurants, and consumer herding behaviors indicate that more popular restaurants attract more consumers; by providing indicators of popularity as ratings, Yelp would direct consumers also to higher rated restaurants. Therefore, if these channels are influential, ratings on Yelp will affect both young and old independent restaurants. The empirical result from the RDD analysis, as shown in Table ?? in the Appendix, shows that it is not the case. Changes in Yelp ratings affect only young independent restaurants and has no effect on old established restaurants. This evidence, however, is consistent with the theory of consumer learning.

In short, other channels of managerial learning, consumer sorting, reputation and herding do not seem to play an important role in the effects of Yelp on restaurant revenues. The empirical
Table 21: Relationship Between Average New Rating Per Month and Age

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log Average New Rating Per Month</td>
<td>Log Average New Rating Per Month</td>
<td>Log Average New Rating Per Month</td>
<td>Log Average New Rating Per Month</td>
<td>Log Average New Rating Per Month</td>
</tr>
<tr>
<td>Log Age</td>
<td>-0.00291</td>
<td>-0.0375</td>
<td>-0.00493</td>
<td>-0.00671</td>
<td>-0.00516</td>
</tr>
<tr>
<td></td>
<td>(0.00320)</td>
<td>(0.0401)</td>
<td>(0.00370)</td>
<td>(0.00321)</td>
<td>(0.00368)</td>
</tr>
<tr>
<td>Log Age×Chain Dummy</td>
<td>-0.00595</td>
<td>-0.0308</td>
<td>-0.00423</td>
<td>-0.00536</td>
<td>-0.00346</td>
</tr>
<tr>
<td></td>
<td>(0.00673)</td>
<td>(0.0769)</td>
<td>(0.00820)</td>
<td>(0.00669)</td>
<td>(0.00815)</td>
</tr>
<tr>
<td>Log Yelp</td>
<td>-0.0121</td>
<td>-0.00970</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00751)</td>
<td>(0.0102)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Yelp×Chain Dummy</td>
<td>-0.000707</td>
<td>-0.0227</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01360)</td>
<td>(0.0207)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traffic (thousands)</td>
<td>0.0648*</td>
<td>0.0918*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0357)</td>
<td>(0.0496)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Total Population</td>
<td>0.0237</td>
<td>0.0479</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00667)</td>
<td>(0.0988)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income (millions)</td>
<td>0.383</td>
<td>1.59***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.413)</td>
<td>(0.597)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Visitor Spending (millions)</td>
<td>-0.00403</td>
<td>-0.0116</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00392)</td>
<td>(0.0331)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(31.78)</td>
<td>(47.23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>-0.114</td>
<td>0.0839</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.154)</td>
<td>(0.264)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>-0.137</td>
<td>-0.0346</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.197)</td>
<td>(0.392)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>-0.257**</td>
<td>-0.065*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.364)</td>
<td>(0.517)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>-0.0003</td>
<td>-0.0830</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0738)</td>
<td>(0.112)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of White population</td>
<td>0.00059</td>
<td>0.0136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0064)</td>
<td>(0.0570)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Black population</td>
<td>-0.123</td>
<td>0.0343</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.128)</td>
<td>(0.188)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>0.215*</td>
<td>0.169</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.126)</td>
<td>(0.162)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Independent Rivals</td>
<td>-0.000491</td>
<td>0.000129</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000513)</td>
<td>(0.000766)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Chain Rivals</td>
<td>-0.00150</td>
<td>0.00175</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00255)</td>
<td>(0.00285)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE×Chain Dummy</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Restaurant FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample Period</td>
<td>All</td>
<td>Before Yelp Penetration</td>
<td>After Yelp Penetration</td>
<td>All</td>
<td>After Yelp Penetration</td>
</tr>
<tr>
<td>N</td>
<td>144,659</td>
<td>70,899</td>
<td>84,754</td>
<td>144,659</td>
<td>84,754</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* p < 0.10, ** p < 0.05, *** p < 0.01
evidence in this section points to consumer learning as the main mechanism for Yelp’s effect.

.5 Identification of the Structural Model

Here I discuss in greater detail about the identification of the learning parameters. For simplicity of presentation, I discuss the simpler model without including restaurant type $p$ in terms of the platforms they are listed and without incorporating the TripAdvisor penetration. The identification issues are the same between the simpler and the more complex versions of the model. First, as shown in equations 1.5.3 to 1.5.4, $\sigma_\delta$ cannot be separately identified from $\kappa$ and $\lambda$. Therefore, it is normalized to 1. Second, the identification of other learning parameters $\omega$, the initial prior mean $A_p^j$, the initial prior variance $\sigma^2_A$, $\kappa$ and $\lambda$ can be demonstrated through the following analysis. $\eta_{jt}$ can be written as follows

$$\eta_{jt} = \frac{A_p^j + \sigma^2_A N_{jt-1} A_j + \sigma^2_A (\sum_{t=1}^{N_{jt-1}} \delta_l)}{1 + \sigma^2_A N_{jt-1}} \tag{.5.1}$$

where $N_{jt-1} \equiv \sum_{t=1}^{i-1}(\kappa + \lambda Yelp_{ml})q_{jt}$, i.e. the total number of experience signals received by consumers; $\delta_l$ is the random component of the experience signals. It is i.i.d. and follows a standard normal distribution since $\sigma_\delta$ has been normalized to 1.

Then $\eta_{jt}$ can be decomposed into three terms:

$$\eta_{jt} = \frac{A_p^j}{1 + \sigma^2_A N_{jt-1}} + \frac{\sigma^2_A N_{jt-1} A_j}{1 + \sigma^2_A N_{jt-1}} + \frac{\sigma^2_A (\sum_{t=1}^{N_{jt-1}} \delta_l)}{1 + \sigma^2_A N_{jt-1}} \tag{.5.2}$$

It is easier to think about the identification of the parameters for a given rating class first. Suppose we take all restaurants with a rating of 3 in our estimation, then $A_p^j$ and $A_j$ are the prior mean and true mean quality of these restaurants. We can focus on how we can identify $A_p^j$ and $A_j$ and other learning parameters for this rating class. The first and second terms of equation .5.2 is a weighted average of the initial prior mean $A_p^j$ and the true mean quality $A_j$. Together they determine the evolution of the mean revenues with respect to age $t$ after all other demand variables have been controlled for. In particular, at age 0, the revenues start with $A_p^j$, and when a restaurant reaches a certain age and $N_{t-1}$ is large enough, the revenues will converge to $A_j$. Therefore, restaurants’ revenues at age $t = 0$ can identify $A_p^j$, and the revenues at a much later age identify $A_j$. How fast the mean revenues converge from $A_p^j$ to $A_j$ identifies $\sigma^2_A N_{t-1}$. It is important to know if $\sigma^2_A$ can be separately identified from $N_{t-1}$. If so, the variation of Yelp in $N_{t-1}$ will identify both $\kappa$ and $\lambda$.

The variance of the third term in equation .5.2 identifies $\sigma^2_A$. To see this, we can write the variance as

$$Var = \frac{\sigma^4_A}{(1 + \sigma^2_A N_{t-1})^2} N_{t-1} = \frac{\sigma^2_A^2}{(1 + \sigma^2_A N_{t-1})^2} \sigma^2_A N_{t-1} \tag{.5.3}$$
Since $\sigma_A^2 N_{t-1}$ can be identified from the speed of convergence of the revenues, the variance of the third term therefore identifies $\sigma_A^2$.

Once $A_j^p$ and $A_j$ have been identified for each rating class, we can collect these $A_j^p$ and $A_j$ for all rating classes from 1 star to 5 stars and fit them as a linear function of these star ratings. Fitting either $A_j^p$ or $A_j$ as a linear function of rating $R_j$ will result in an intercept and a slope. However, it can be easily shown that only the difference between the intercepts for the linear function of $A_j^p$ and that of $A_j$ can be identified because the common part of the intercepts for both equations can be taken out of the learning component $\eta_{jt}$, and once taken out, it will be absorbed by the market fixed effects. Therefore, I formulate the linear function of the initial prior $A_j^p = \theta_j^p R_j$ without an intercept, but the function of the true mean quality $A_j = \theta_0 + \theta_r R_j$ does have an intercept. Once $A_j^p$ and $A_j$ are identified for all rating classes, the intercept and slope parameters $\theta_0$, $\theta_j^p$, and $\theta_r$ can be easily identified.

.6 Derivation of the Likelihood Function

In this section, I derive the closed form of the log-likelihood function. For cleaner presentation, I rewrite the model by excluding all control variables:

\[
y_{jt} = \eta_{jt} + \xi_{jt} \quad (.6.1)
\]
\[
\eta_{jt} = (1 - \beta_{jt-1})\eta_{jt-1} + \beta_{jt-1} A + \nu_{jt}, \nu_{jt} \sim N(0, \sigma_{\nu_{jt}}^2) \quad (.6.2)
\]
\[
\xi_{jt} = \rho \xi_{jt-1} + \varepsilon_{jt}, \varepsilon_{jt} \sim N(0, \sigma_{\varepsilon_{jt}}^2) \quad (.6.3)
\]

where $y_{jt}$ can simply be seen as the residual by excluding all the effects from the linear controls. Rearrange the above variables to eliminate $\xi_{jt}$, we have

\[
y_{jt} = \rho y_{jt-1} + \eta_{jt} - \rho \eta_{jt-1} + \varepsilon_{jt} \quad (.6.4)
\]
\[
\eta_{jt} = (1 - \beta_{jt-1})\eta_{jt-1} + \beta_{jt-1} A + \nu_{jt} \quad (.6.5)
\]

Let $y_j = (y_{j0}, y_{j1}, \cdots, y_{jT_j})$ denote the time series revenue data for restaurant $j$ and $p(y_j|x, \theta)$ be the likelihood for observing $y_j$ given the structural parameters $\theta$ and independent variables $x$. Then the log likelihood function of observing the revenue streams for all restaurants is

\[
L = \sum_{j=1}^{n} \log(p(y_j|x, \theta)) \quad (.6.6)
\]

where $n$ is the total number of restaurants in the sample.

To obtain this log likelihood function, we need to compute $\log(p(y_j|x, \theta))$ for each restaurant. Note that due to the serial correlation in $\xi_{jt}$, we will not be able to count the likelihood associated with the first observation $y_{j0}$. Given that $\eta_{jt}$ is unobserved, we need to integrate it out. The most straightforward approach is to integrate out $\eta_{jt}$ over the joint distribution of $(\eta_{j1}, \cdots, \eta_{jT_j})$;
This equation implies an iterative process, where the likelihood of observations, $p(y_j | x_0, \theta)$, can be obtained recursively. Once we know the first period’s likelihood, $p(y_1 | x_0, \theta, y_0)$, we can obtain $p(y_1 | x_0, \theta, y_0)$ from equation (6.8). With that we can calculate $p(y_j | x_0, \theta, y_{j-1})$ using the last line in equation (6.7). Once we continue to do this iteratively, we can obtain $p(y_j | x_0, \theta, y_{j-1})$ for all $t$. Given that all errors in this model are Gaussian, a closed-form of these probability functions exist. All of the above conditional probabilities come from normal
distributions, and their mean and variances can be written as follows:

\[
\hat{\eta}_{t|t-1} = (1 - \beta_{t-1})\hat{\eta}_{t-1|t-1} + \beta_{t-1} A
\]

(6.9)

\[
\hat{\sigma}^2_{\eta_{t|t-1}} = (1 - \beta_{t-1})^2 \hat{\sigma}_{\eta_{t-1|t-1}}^2 + \sigma_{\nu_t}^2
\]

(6.10)

\[
\hat{y}_t = y_t - \rho y_{t-1} - \hat{\eta}_{t|t-1} + \rho \hat{\delta}_{t-1|t-1}
\]

(6.11)

\[
S_t = \sigma^2_x + \hat{\sigma}^2_{\eta_{t|t-1}} + \rho (\rho - 2(1 - \beta_{t-1})) \hat{\sigma}_{\eta_{t-1|t-1}}^2
\]

(6.12)

\[
\hat{\eta}_{t|t} = \hat{\eta}_{t|t-1} + \left( \hat{\sigma}_{\eta_{t|t-1}}^2 - \rho (1 - \beta_{t-1}) \hat{\sigma}_{\eta_{t-1|t-1}}^2 \right) S_t^{-1} \hat{y}_t
\]

(6.13)

\[
\hat{\sigma}^2_{\eta_{t|t}} = \left[ \sigma^2_x \hat{\sigma}_{\eta_{t|t-1}}^2 + \rho^2 \hat{\sigma}_{\eta_{t-1|t-1}}^2 \hat{\sigma}^2_{\eta_{t|t-1}} - \rho^2 (1 - \beta_{t-1})^2 \hat{\sigma}_{\eta_{t-1|t-1}}^4 \right] S_t^{-1}
\]

(6.14)

\[
\hat{y}_{t|t-1} = \rho y_{t-1} + \hat{\eta}_{t|t-1} - \rho \hat{\delta}_{t-1|t-1}
\]

(6.15)

\[
\hat{\delta}_{\eta_{t|t-1}}^2 = S_t
\]

(6.16)

where \( \hat{\eta}_{j|t-1} \) is the mean of \( \eta_{jt} \) conditional on past data \( y_{j0:t-1} \). \( \hat{\eta}_{j|t} \) is similarly defined. \( \hat{\sigma}^2_{\eta_{j|t-1}} \) is the variance of \( \eta_{jt} \) given past data \( y_{j0:t-1} \) and \( \hat{\sigma}^2_{\eta_{j|t|t}} \) is also similarly defined. \( \hat{y}_{jt} \) is the residual. \( S_{jt} \) is an intermediate factor that facilitates the updating of \( \hat{\eta}_{j|t|t} \) and \( \hat{\sigma}^2_{\eta_{j|t|t}} \). \( \hat{y}_{jt:t-1} \) is the mean of \( y_{jt} \) conditional on past observations \( y_{j0:t-1} \), and \( \hat{\delta}_{y_{jt:t-1}} \) is the variance.

Once we have \( p(y_j|x, \theta) \) for each restaurant, we can obtain the total log likelihood for all restaurants \( L \) as shown in equation (6.6) and the structural parameters can be estimated by maximizing \( L \):

\[
\theta^* = \arg \max_{\theta} \sum_{j=1}^{n} \log(p(y_j|x, \theta))
\]

.7 Technical Issues Regarding Welfare Calculation

It is easy to derive that the difference between \( CS_{mt} \) and \( CS_{mt}^c \) is

\[
CS_{mt} - CS_{mt}^c = \log(Rev_{0|mt}) - \log(Rev_{0|mt})
\]

\[
+ \sum_{j \in J_{mt}} \left[ (s_{jt} - s_{jt}^c) (\Delta_{jt}^T - \Delta_{0mt}) - s_{jt} (\Delta_{jt} - \Delta_{0mt}) + s_{jt}^c (\Delta_{jt}^c - \Delta_{0mt}) \right]
\]

(7.1)

where \( J_{mt} \) is the set of new independent restaurants that are listed on Google in market \( m \) at time \( t \). Note that there are restaurants that are not listed on Google during the sample period. For those restaurants, I simply assume that consumers’ true mean utility and expected utility are the same, and thereby \( \Delta_{jt}^T - \Delta_{jt} = 0 \). Those restaurants include chain restaurants, older established independent restaurants listed on Google, and restaurants that are not listed on Google.

Based on the above equation, to calculate the difference, we need to know the true mean utility difference \( \Delta_{jt}^T - \Delta_{0mt} \) for new independent restaurants listed on Google. This difference can be predicted given that we have uncovered all the structural parameters and each restaurant’s true
where $\xi_{jt}$ represents the sum of the estimated aggregate demand shocks $\xi_{jt}$ and experience signal shocks $\delta_{t}$.

To extract $\hat{\xi}_{jt}$, I substitute equation (5.2) into equation (1.5.1) and rewrite the revenue function as follows:

$$y_{jt} = X_{mt} \hat{\theta}_{x} - \hat{\alpha} \log(p_{jt}) + \sum_{k=1}^{5} \theta_k r^k + \hat{\theta}_{c}$$

Recall from equation (5.2) that $\eta_{jt}$ has a deterministic part $\frac{\hat{A}_{j}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}} + \frac{\hat{\sigma}_{A}^{2}N_{jt-1}\hat{A}_{j}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}}$ and a noisy part $\frac{\hat{\sigma}_{A}^{2}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}}(\sum_{t=1}^{N_{jt-1}} \delta_{t})$. This noisy part can be lumped into the aggregate demand shocks $\hat{\xi}_{jt}$ to form the aggregate residual $\hat{\xi}_{jt}:

$$\hat{\xi}_{jt} = \frac{\hat{\sigma}_{A}^{2}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}}(\sum_{t=1}^{N_{jt-1}} \delta_{t}) + \hat{\xi}_{jt}$$

$$= y_{jt} - X_{mt} \hat{\theta}_{x} - \hat{\alpha} \log(p_{jt}) - \sum_{k=1}^{5} \theta_k r^k - \hat{\theta}_{c}$$

$$- \frac{\hat{A}_{j}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}} - \frac{\hat{\sigma}_{A}^{2}N_{jt-1}\hat{A}_{j}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}}$$

(7.4)

To construct $\Delta_{jt}^c - \Delta_{0t}$, I use a similar equation as equation (7.2)

$$\Delta_{jt}^c - \Delta_{0mt} = X_{mt} \hat{\theta}_{x} - \hat{\alpha} \log(p_{jt}) + \sum_{k=1}^{5} \theta_k r^k + \hat{\theta}_{c}$$

$$+ \frac{\hat{A}_{j}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}^c} + \frac{\hat{\sigma}_{A}^{2}N_{jt-1}^c\hat{A}_{j}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}^c} + \hat{\xi}_{jt}$$

(7.5)

where $N_{jt-1}^c \equiv \sum_{t=1}^{t-1}(\kappa + \lambda Y_{elp_{mt}})\eta_{jt}^{30}$ and $\frac{\hat{A}_{j}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}^c} + \frac{\hat{\sigma}_{A}^{2}N_{jt-1}^c\hat{A}_{j}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}^c}$ represents the deterministic part of the expected quality $\eta_{jt}^{c}$ in the counterfactual. The noisy part of $\eta_{jt}^{c}$ should be $\frac{\hat{\sigma}_{A}^{2}}{1+\hat{\sigma}_{A}^{2}N_{jt-1}^c}(\sum_{t=1}^{N_{jt-1}^c} \delta_{t})$, which differs from the noisy part of $\eta_{jt}$ in terms of the total revenue stream $N_{jt-1}$. To retain the shocks in terms of experience signals that firms encounter, I use the estimate

$^{30}$Here again, for notational simplicity, I am using a simpler notation where $Y_{elp_{mt}}$ represents the penetration of all platforms.
of this component from the real world to approximate its value in the counterfactual\(^{31}\) that is, I set this component in the counterfactual to \(\frac{\hat{\sigma}^2}{1+\sigma^2 N_{jt-1}}(\sum_{l=1}^{N_{jt-1}} \delta_l)\). Furthermore, I also retain the aggregate demand shocks from the factual world \(\hat{\xi}_{jt}\). Therefore, the sum of these two types of shocks \(\hat{\xi}_{jt}\) are added to the equation for determining \(\Delta_{jt} - \Delta_{0mt}\).

To obtain \(N_{jt}^c\), we need counterfactual revenue streams \(Rev_{jt}^c\), which could be translated into quantities once divided by price. To get \(Rev_{jt}^c\), we simply need to calculate the revenue share \(s_{jt}^c\), and then use that to multiply the market size. \(s_{jt}^c\) can be attained using the following equation:

\[
s_{jt}^c = \frac{\exp(\Delta_{jt}^c - \Delta_{0mt}^c)}{1 + \sum_{k \in J_{mt}} \exp(\Delta_{kt}^c - \Delta_{0mt}^c) + \sum_{l \in Oth_{mt}} \exp(\log(Rev_{lt}) - \log(Rev_{0mt}))} = \frac{\exp(\Delta_{jt}^c - \Delta_{0mt}^c)}{1 + \sum_{k \in J_{mt}} \exp(\Delta_{jt}^c - \Delta_{0mt}^c) + \sum_{l \in Oth_{mt}} Rev_{lt}}
\]

(7.6)

where \(Oth_{mt}\) is the set of all other independent restaurants in market \(m\) at time \(t\) in addition to new independent restaurants listed on Google. Note that their mean utility can be approximated by \(\exp(\log(Rev_{lt}) - \log(Rev_{0mt}))\). Once the shares are computed, the counterfactual revenue stream of a restaurant can be easily obtained by applying equation \(1.2.13\) and the revenue of the outside option in the counterfactual \(Rev_{0mt}^c\) can be computed in the same way.

Now we have attained \(Rev_{0mt}^c\), \(\Delta_{jt} - \Delta_{0mt}, \Delta_{jt}^c - \Delta_{0mt}^c\), and \(s_{jt}^c\), and we know \(\Delta_{jt} - \Delta_{0mt}\) and \(s_{jt}\) from the data; we can calculate \(CS_{mt} - CS_{0mt}^c\) for all markets.

### 8 Tables

#### Table 22: Yelp Restaurant Category Classification

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>breakfast, brunch, sandwiches, cafe, and vegan</td>
</tr>
<tr>
<td>2</td>
<td>Asian, including Chinese, Japanese, Korean and Indian</td>
</tr>
<tr>
<td>3</td>
<td>bars, gastropubs, traditional and new American</td>
</tr>
<tr>
<td>4</td>
<td>Tex-Mex, cajun, Caribbean and other southern cuisines</td>
</tr>
<tr>
<td>5</td>
<td>European, including French, Spanish, German, Greek, and other Mediterranean cuisines</td>
</tr>
</tbody>
</table>

\(^{31}\)This approximation is innocuous since a closer examination of this component shows that it is the average of all experience signals up to time \(t - 1\). This component in the counterfactual differs from that in the factual world only in terms of the precision or variance, but not in terms of signs. Therefore, firms that encountered good shocks in terms of experience signals in the real world will also experience good shocks in the counterfactual.
Table 23: Yelp Restaurant Prices

<table>
<thead>
<tr>
<th>Category</th>
<th>Dollar Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>$10</td>
</tr>
<tr>
<td>$$</td>
<td>$20</td>
</tr>
<tr>
<td>$$</td>
<td>$45</td>
</tr>
<tr>
<td>$$$$</td>
<td>$70</td>
</tr>
</tbody>
</table>
Table 24: Revenue-Age Profiles by Quality

<table>
<thead>
<tr>
<th>Model</th>
<th>Log Revenue</th>
<th>Google Rating</th>
<th>Slope Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Independent, $\theta_a$)</td>
<td>2</td>
<td>-0.0727***</td>
</tr>
<tr>
<td></td>
<td>(Chain, $\theta_{ac}$)</td>
<td>4</td>
<td>0.0073</td>
</tr>
<tr>
<td></td>
<td>$\times$ Rating (Independent, $\theta_{ar}$)</td>
<td>3</td>
<td>-0.0327*</td>
</tr>
<tr>
<td></td>
<td>$\times$ Rating (Chain, $\theta_{acr}$)</td>
<td>5</td>
<td>0.0473***</td>
</tr>
<tr>
<td>Traffic (thousands)</td>
<td>-0.0817</td>
<td></td>
<td>(0.0949)</td>
</tr>
<tr>
<td>Log Total Population</td>
<td>0.174</td>
<td></td>
<td>(0.258)</td>
</tr>
<tr>
<td>Log Visitor Spending</td>
<td>0.222*</td>
<td></td>
<td>(0.131)</td>
</tr>
<tr>
<td>Population Density</td>
<td>-134.8</td>
<td></td>
<td>(155.7)</td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>0.653</td>
<td></td>
<td>(1.130)</td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>1.244</td>
<td></td>
<td>(0.873)</td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>-0.780</td>
<td></td>
<td>(1.682)</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>-0.237</td>
<td></td>
<td>(0.326)</td>
</tr>
<tr>
<td>Share of White population</td>
<td>0.345**</td>
<td></td>
<td>(0.168)</td>
</tr>
<tr>
<td>Share of Black population</td>
<td>0.539</td>
<td></td>
<td>(0.546)</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>-0.690</td>
<td></td>
<td>(0.604)</td>
</tr>
<tr>
<td>Number of Independent Rivals</td>
<td>-0.00635***</td>
<td></td>
<td>(0.00223)</td>
</tr>
<tr>
<td>Number of Chain Rivals</td>
<td>-0.0151***</td>
<td></td>
<td>(0.00436)</td>
</tr>
<tr>
<td>Cuisine Dummy $\times$ Demographics</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time FE $\times$ Chain Dummy</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restaurant FE</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 289,139

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

Standard errors in parentheses
Table 25: Effects of Yelp Exposure on Survival Rates

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) action</th>
<th>(2) action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Yelp ($\theta_y$)</td>
<td>-0.000726</td>
<td>-0.00426**</td>
</tr>
<tr>
<td></td>
<td>(0.00115)</td>
<td>(0.00204)</td>
</tr>
<tr>
<td>Log Yelp $\times$ Rating ($\theta_{yr}$)</td>
<td>-0.0000190</td>
<td>0.00123**</td>
</tr>
<tr>
<td></td>
<td>(0.000336)</td>
<td>(0.000557)</td>
</tr>
<tr>
<td>Log Yelp $\times$ Chain Dummy ($\theta_{yc}$)</td>
<td>-0.00231</td>
<td>0.000342</td>
</tr>
<tr>
<td></td>
<td>(0.00158)</td>
<td>(0.000956)</td>
</tr>
<tr>
<td>Log Yelp $\times$ Chain Dummy $\times$ Rating ($\theta_{ycr}$)</td>
<td>0.000625</td>
<td>-0.000251</td>
</tr>
<tr>
<td></td>
<td>(0.000441)</td>
<td>(0.000312)</td>
</tr>
<tr>
<td>Traffic (thousands)</td>
<td>0.00106</td>
<td>-0.00327</td>
</tr>
<tr>
<td></td>
<td>(0.000856)</td>
<td>(0.00319)</td>
</tr>
<tr>
<td>Log Total Population</td>
<td>-0.000261</td>
<td>0.00227</td>
</tr>
<tr>
<td></td>
<td>(0.000683)</td>
<td>(0.00391)</td>
</tr>
<tr>
<td>Income (millions)</td>
<td>-0.0120</td>
<td>0.0225</td>
</tr>
<tr>
<td></td>
<td>(0.0112)</td>
<td>(0.0399)</td>
</tr>
<tr>
<td>Log Visitor Spending</td>
<td>0.000605</td>
<td>0.00225</td>
</tr>
<tr>
<td></td>
<td>(0.000511)</td>
<td>(0.00287)</td>
</tr>
<tr>
<td>Population Density</td>
<td>-0.118</td>
<td>1.669</td>
</tr>
<tr>
<td></td>
<td>(0.862)</td>
<td>(3.378)</td>
</tr>
<tr>
<td>Share of population (age 15-34)</td>
<td>-0.00177</td>
<td>0.00140</td>
</tr>
<tr>
<td></td>
<td>(0.00831)</td>
<td>(0.0233)</td>
</tr>
<tr>
<td>Share of population (age 35-64)</td>
<td>-0.00407</td>
<td>0.0110</td>
</tr>
<tr>
<td></td>
<td>(0.00829)</td>
<td>(0.0304)</td>
</tr>
<tr>
<td>Share of population (age 65 and up)</td>
<td>0.00341</td>
<td>0.0749**</td>
</tr>
<tr>
<td></td>
<td>(0.00933)</td>
<td>(0.0314)</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>0.00373*</td>
<td>0.00231</td>
</tr>
<tr>
<td></td>
<td>(0.00200)</td>
<td>(0.00720)</td>
</tr>
<tr>
<td>Share of White population</td>
<td>0.000879</td>
<td>-0.00140</td>
</tr>
<tr>
<td></td>
<td>(0.00189)</td>
<td>(0.00362)</td>
</tr>
<tr>
<td>Share of Black population</td>
<td>0.00170</td>
<td>-0.0182</td>
</tr>
<tr>
<td></td>
<td>(0.00238)</td>
<td>(0.0126)</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>0.0132**</td>
<td>-0.00928</td>
</tr>
<tr>
<td></td>
<td>(0.00609)</td>
<td>(0.0139)</td>
</tr>
<tr>
<td>Number of Independent Rivals</td>
<td>-0.00000204</td>
<td>0.0000846</td>
</tr>
<tr>
<td></td>
<td>(0.0000173)</td>
<td>(0.0000589)</td>
</tr>
<tr>
<td>Number of Chain Rivals</td>
<td>-0.00000290</td>
<td>0.0000365</td>
</tr>
<tr>
<td></td>
<td>(0.0000329)</td>
<td>(0.0000137)</td>
</tr>
<tr>
<td>Wage</td>
<td>-0.000119</td>
<td>0.0104***</td>
</tr>
<tr>
<td></td>
<td>(0.000211)</td>
<td>(0.00237)</td>
</tr>
</tbody>
</table>

Cuisine Dummy $\times$ Demographics: Yes, Yes
Time FE: Yes, Yes
Time FE$\times$Chain Dummy: Yes, Yes
Restaurant FE: Yes, Yes

<table>
<thead>
<tr>
<th>Sample Period</th>
<th>Yelp November 2016 Rating</th>
<th>Yelp November 2016 Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entry 2 years before March, 2005</td>
<td>131,079</td>
<td>177,373</td>
</tr>
<tr>
<td>Entry after March, 2005</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01
Chapter 2

Measuring Preemption in Dynamic Oligopoly Games

2.1 Introduction

Preemption is an excessive entry behavior of an incumbent firm aiming to prevent the entry of a potential competitor. Preemption can be spatial or non-spatial. In a spatial context, preemption relates to a multi-product incumbent firm aggressively filling out all market “niches” in order to leave no room for potential entrants. In a non-spatial context, preemption refers to a single-product incumbent firm entering a market early so as to occupy the market before its rivals. In both contexts, preemption pertains to an aggressive entry strategy with the ultimate goal of keeping potential competitors out of the market.

Given its anti-competitive nature, preemption has been of keen interest to anti-trust authorities. In the late 1970s, for example, the U.S. Federal Trade Commission (FTC) charged breakfast cereal companies Kellogg’s, General Mills, and General Foods with preempting the market by introducing about 150 new brands during the period of 1950-1970 (Viscusi et al., 2005). More recently in March 2014, the Israel parliament passed the Law for Encouragement of Competitiveness in the Food Sector, which prevents a large wholesaler from opening a second big store in a pre-defined “competitive geographic area” for preemptive purposes (Library of Congress, 2014). Recent empirical research also shows that preemptive motives can play an important role in technological innovation and in the organizational structure of franchised firms (Igami, 2015; Nishida and Yang, 2016). Despite its anti-competitive nature, preemption could be beneficial to welfare given that it motivates firms to enter a market and serve consumers early. This intriguing aspect of preemption has inspired empirical researchers to explore the welfare effects of preemption.

Preemption is an inherently dynamic phenomenon. As a result, dynamic oligopoly models with entry and exit under the Ericson and Pakes (1995) framework (EP henceforth) are great tools for examining preemption.\(^1\) However, measuring the effects of preemption is still a challenge. A key

\(^1\)This framework models forward-looking behaviors of firms and endogenizes firms’ entry and exit decisions. Recent
challenge is that preemption is the indirect strategic effect of competition, not the direct effect of competition. While conceptually different, these two types of effects are very closely entangled in a dynamic oligopoly game. It is hard to separate the two. To add to the difficulty, firms may enter a market excessively for long-term investment purposes other than preemption. For example, McDonald’s has been reported to receive more profits from the real estate investment associated with its stores than the immediate profit from selling burgers (Love, 1995). Separating out firms’ preemptive motives from other dynamic considerations is also a challenge. Due to these obstacles, defining preemption and capturing the effect of preemption in the EP framework have been difficult.

This paper aims to tackle these challenges by isolating out preemptive motives from all other entry motives of firms in the equilibrium conditions of a dynamic oligopoly game. In doing so, it proposes a definition of preemption and a counterfactual that eliminates the effect of preemption. To define preemptive motives, I first decompose a firm’s entry motives into the marginal benefit of entry in the current period and that in future periods. The marginal benefit of entry in the future can be further broken down into three parts: (1) the return on investment through savings in entry costs or growth in scrap values; (2) the return on entry through sales to consumers due to a firm’s own entry and exit behaviors in the market; (3) the return on entry through sales to consumers due to the entry-deterrence effect of a firm’s incumbency status on its rival. The third component is identified as the preemptive motives.

By forcing firms to ignore their preemptive motives and re-optimize their values, I obtain a counterfactual of preemption that eliminates only the effect of preemption. Specifically, I replace a firm’s future benefit-of-entry, say firm A’s, with a counterfactual benefit that imposes the restriction that the rival’s future entry-exit behavior is the same regardless firm A is active or not in the market, keeping all the structural parameters constant. I calculate a new (counterfactual) equilibrium under this restriction. This replacement eliminates the gain from preemption in the future marginal benefit of entry. Because the marginal benefit of entry in the current period remains intact, this counterfactual retains the direct competitive effect. Compared to other approaches in the literature, this counterfactual has the advantage of preserving the direct competitive effect and the dynamic features of the game while removing the indirect strategic effect of competition. By contrasting firms’ behaviors in the counterfactual with those in the equilibrium, I attain a measure of preemption.

I apply this counterfactual and measure of preemption in two studies: one numerical and the other empirical. The numerical study examines the influence of entry and exit costs on pre-emption. Contrary to the conventional wisdom which derives from Judd’s (1985) seminal work, I find that firms can successfully pre-empt entry without large exit costs. The empirical study investigates how

\footnote{The direct effect is the impact of the incumbent’s action on its own profits by taking its rival’s actions as given. This effect is often considered welfare improving because it undermines the incumbent’s ability to charge a monopoly price. The strategic effect, on the other hand, is the influence of the incumbent’s current action on its rival’s future behaviors. The ultimate goal of this effect is to induce rivals’ exit and reduce competition in the long run (Tirole, 1988; Besanko et al., 2004). It is usually seen as anti-competitive.}

\footnote{For example, McDonald’s return on real estate investment will be part of this return.}
the effects of preemption vary across market size in the Canadian burger industry, I find that the effects are much stronger in medium-sized markets than small or large markets.

Preemption has been studied extensively both theoretically and empirically. Early theoretical papers emphasize the importance of strategic commitment, which is the ability to commit to staying in the market after entry. Seminal papers by Schmalensee (1978) and Eaton and Lipsey (1979) use models with irreversible entry, which enable firms to fully commit to staying in the market after entry. Under this setup, preemption is always successful. Both papers show that the incumbent will preempt entry by crowding out the product space. However, their findings were challenged by Judd (1985) who made an important contribution to the field by pointing out that if entry is not irreversible, the incumbent cannot preempt entry without a large exit cost. A substantial part of the subsequent theoretical literature after Judd (1985) focuses on proving that the exit cost for the incumbent can be large. In the empirical literature, many studies use static reduced-form models to examine preemption and have found evidence suggesting preemption. However, since preemption is an inherently dynamic phenomenon, the empirical results based on static models are mostly indicative and cannot establish a direct causal link.

All of the models in both the theoretical and empirical studies mentioned above provide limited entry and exit opportunities. In the real world, firms interact with each other repeatedly over a long period of time. A realistic model should incorporate this feature. The dynamic oligopoly games under the EP framework is one such model. A number of recent studies have examined preemption using the EP framework. Examples include Aguirregabiria and Ho (2012), Igami (2015), Igami and Yang (2016), Zheng (2016), and Hünermund et al. (2014). All studies employ some type of counterfactual of preemption. However, none of the counterfactuals remove only the effect of

---

4This non-monotonic relationship between preemption and market size resonates with those results in Ellison and Ellison (2011), who study preemptive investment through a stylized three-stage game and find that non-monotonicity of investment levels is an indicator of preemption. It should be noted that their findings are very different from the non-monotonicity results in the empirical application in this paper. Their non-monotonicity test applies to the overall equilibrium investment levels, while my non-monotonicity results come from the difference between firms' equilibrium behaviors and their counterfactual ones. In the context of my model, firms' equilibrium behaviors with preemptive motives are often monotonic, and therefore, non-monotonicity cannot provide a test for preemptive motives.

5Schmalensee (1978) illustrates this in the context of the breakfast cereal industry, and Eaton and Lipsey (1979) prove this theory for a growing market. Other examples of theoretical papers on entry deterrence include Gilbert and Newberry (1982), Gilbert and Harris (1984), Fudenberg and Tirole (1985) and Bonanno (1987).

6This is because the incumbent's multi-products are substitutes, and fierce competition at one product location will reduce the incumbent's profits at other product locations. If a multi-product incumbent is challenged by the entrant at one product location, intense competition at that location will bring down the incumbent's profits at all locations. Without a large exit cost, the incumbent would prefer exiting the challenged location in order to alleviate competition in the entire market. Knowing that, it will not have preempted the market by product proliferation in the first place.

7For example, Hadfield (1991) shows that the renegotiation of franchise contract can be very costly for the incumbent if it decides to close a store. Choi and Scarpa (1992) illustrate that withdrawing a product can damage a brand's reputation and reduce sales of other products under the same brand, and therefore, the incumbent will not withdraw a product without careful contemplation.


9Recent empirical work that use this type of dynamic model has investigated the general issue of predation; for example, Schmidt-Dengler (2006) and Chicu (2012). However, these studies focus on firms' competitive strategies through choice of time or capital investment, not firms' entry and exit activities, which are the focus of preemption in this paper.
These counterfactuals can be grouped into three types: (1) The first type limits preemption by changing the structural parameters. For example, Aguirregabiria and Ho (2012), who examine preemption in the airline industry, use a counterfactual that removes the complementarity between products in order to reduce the incumbent firm’s ability to commit. However, as mentioned earlier, preemption is an endogenous phenomenon, by changing exogenous structural parameters, this type of counterfactual eliminates not only preemption but also other aspects of the equilibrium outcome. (2) The second type eliminates preemptive motives by removing the competition faced by one firm. For example, Igami and Yang (2016), who examine competition in the fast food industry, propose the counterfactual where the potential entrants ignore McDonald’s presence in the market. This setup erases the competitive effect of McDonald’s on other firms’ profits all together. It eliminates not only preemption, which arises from competition, but also the direct effect of competition. (3) The third type uses an open-loop equilibrium to remove preemption. For example, Hümernund et al. (2014) uses this concept to study preemption in the U.S. penicillin industry. Their open-loop equilibrium forces firms to pre-commit at the beginning of time to a strategy where their actions depend only on their own states regardless of their rivals’ states. This counterfactual essentially converts a dynamic game into a static game because firms cannot react to their rivals’ incumbency status in the market. In this regard, this type of counterfactual loses the dynamic aspect of strategic interactions between firms. In short, the counterfactuals that have been proposed in the literature cannot remove the effect of preemption without influencing other aspects of the equilibrium outcome.

To respond to this deficiency, this paper proposes a measure of preemption that captures only the effect of preemption. This is the main contribution of this paper. In addition, this paper also demonstrates that firms can preempt entry without a large exit cost, contrary to conventional wisdom.

My approach of capturing preemptive motives through a decomposition of firms’ motives in equilibrium conditions is reminiscent of that used in Besanko et al. (2014). However, their paper examines predatory pricing, which is a continuous decision, whereas I examine preemptive entry, which is a discrete choice. In addition, compared to Besanko et al. (2014), the decomposition in this paper is at a finer granularity: it emphasizes the impact of entry and exit costs on firms’ strategies, while Besanko et al. (2014) aggregates these factors with other predatory motives of firms. Furthermore, although the model in Besanko et al. (2014) accommodates entry and exit dynamic of firms, the definition of predation in that paper focuses on pricing and ignores other channels of predation, such as preemptive entry. In this regard, this paper complements the definition of predation in Besanko et al. (2014) through a formal characterization of predatory entry in a dynamic oligopoly game.

The balance of the paper is organized as follows: Section 2 below describes the general model of the dynamic oligopoly game between two firms in a spaceless market. Section 3 defines preemptive motives in a simplified version of the general model and constructs a counterfactual and measure of preemption based on the definition of preemptive motives. Extensions to more complex models
with stochastic demand and spatial markets are made at the end of Section 3. Section 4 examines how sunk costs influence preemption through a numerical analysis. Section 5 investigates how preemption varies across market size through an empirical application. This section extends the counterfactual of preemption to a model with five players. Section 6 compares the counterfactual of preemption developed in this paper to other counterfactuals that have been proposed in the literature. Section 7 concludes by summarizing the key results.

2.2 Model

The model in this paper is a discrete-time Markov dynamic oligopoly game between two retail chains. Following the standard structure in the EP framework, I assume that firms compete in two dimensions: (1) a dynamic dimension that involves the entry and exit of firms; (2) a static dimension, where firms compete in prices or quantities taking market structure, demand, and costs as given.

2.2.1 Static Competition

There are two firms in a market: A and B. They sell differentiated products that are substitutes. In any given period, a firm’s flow profits are summarized by a flow profit function $\pi(\cdot)$, which is an indirect profit function that comes from a static equilibrium of price or quantity competition. $\pi(\cdot)$ depends on the market structure, market demand and cost conditions.\footnote{for example, the size of population, consumers' socioeconomic characteristics, prices of capital, material and land, etc.} Let $s$ denote market structure and $z$ denote the exogenous market variables that represent market demand and cost information, then $\pi(\cdot)$ is a function $\pi_i : Z \times S \rightarrow \mathbb{R}$, where $Z$ is the set of $z$ and $S$ the set of $s$.

2.2.2 Action Set and Markov States

At the beginning of each period, firms can choose to be active or inactive. Let $i \in \{A, B\}$ denote the identity of a firm. Let $a_{it} \in \{0, 1\}$ denote firm $i$’s action at time $t$, with 1 indicating active and 0 inactive. The action space is $A \equiv \{0, 1\}$. The Markov state variables include exogenous market variable $z_t$, which evolves according to a Markov process, and market structure $s_t$. Let $s_{it} \in \{0, 1\}$ denote the active status of firm $i$ at the beginning of each period and $s_t \equiv (s_{At}, s_{Bt})$, then $S$ includes four elements: $S = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Both variables $z_t$ and $s_t$ are common knowledge to firms. In addition, firm $i$ also receives a private idiosyncratic shock $\varepsilon_{it}$, which relates to the fixed cost, entry cost and scrap value. The Markov state at time $t$ is thus a tuple $(s_{it}, z_t, \varepsilon_{it})$.

2.2.3 Timing

The game is an infinite horizon game.\footnote{Some may challenge the infinite-horizon setup by arguing that managers of firms might not look too far ahead and that a finite-horizon game should be more reasonable. While this argument may apply to humans, it does not apply to firms.} The sequence of events within each period is as follows:
• At the beginning of each period, both firms simultaneously make entry and exit decisions, and these decisions take effect right away. At this stage, firms pay entry costs \((EC)\) when enter and receive scrap values \((SV)\) when exit.

• Active firms in the market take the ex-post market structure as given and engage in static competition. At this stage, active firms receive flow profits \(\pi(\cdot)\).

An important feature of the setup here is that firms move simultaneously within each period. Many studies on preemption have assumed a sequential-move setting where the “preemptor” firm is predetermined and set as the first mover.\(^\text{12}\) The model in this paper deviates from sequential moves for two reasons: (1) there is no reason why one firm should act before the other; (2) a simultaneous-move setup allows the identity of the preemptor to be endogenously determined based on the economic fundamentals and beliefs of firms, instead of being pre-chosen and granted the first-mover advantage.

### 2.2.4 Entry costs and scrap values

The entry cost \((EC_i)\) and scrap value \((SV_i)\) are one-time costs firm \(i\) incurs when entering or exiting a market. The entry cost includes the expenses on setting up the store, such as purchasing or leasing the property and outfitting the building. The scrap value comes from the liquidation of the store, such as sales of real estate and equipment. The scrap value could be positive or negative. If negative, it can be interpreted as the exit cost. Since both entry cost and scrap values can change over time given the market condition, they are functions of \(z_t\): \(EC_i(z_t) : Z \rightarrow R\) and \(SV_i(z_t) : Z \rightarrow R\), \(\forall i \in \{A, B\}\).

### 2.2.5 Flow payoff functions

A firm’s total payoff per-period includes three components: the flow profit from the static competition \(\pi_i(\cdot)\), entry costs or scrap values \(EC_i\) and \(SV_i\), and the private shock \(\varepsilon_{it}\). Assume that \(\varepsilon_{it}\) is additively separable in the payoff function and \(i.i.d.\) across choice alternatives and time. Then firm \(i\)’s choice-specific flow payoff function is

\[
\tilde{\Pi}_i(a_{it}, a_{-it}, s_t, z_t, \varepsilon_t) \equiv \pi_i(a_{it}, a_{-it}, z_t) + \varepsilon_{it}(a_{it}) + ev_i(a_{it}, s_t, z_t),
\]

where

\[
ev_i(a_{it}, s_t, z_t) \equiv -(a_{it} - s_{it}) [a_{it}EC_i(z_t) + (1 - a_{it}) SV_i(z_t)].
\]

not necessarily work for corporations. A finite-horizon game assumes that the players know exactly when the end period is. For a human being, it is often reasonable to assume that one does not live beyond 110 years. For corporations, however, it is difficult to predict when they would disappear completely from the market. In particular, even although the managers of a company may not stay with the company forever, the resale value of the corporate tends to incentivize the managers to increase the value of the company as much as possible. Therefore, when modeling the long-term interactions between firms, an infinite horizon is more reasonable.

\(^{12}\)For example, in Igami (2015), the incumbent is chosen as the preemptor and is given the first mover advantage, so that in equilibrium, the incumbent will indeed preempt entry. Similarly, Zheng (2016) also presets one retail chain, the Blue firm, as the preemptor and models the Blue firm as the first mover in each period of the game.
To be precise in terminology, in the remaining text, I use “flow profit” when referring to $\pi_i(\cdot)$, and “flow payoffs” when referring to $\tilde{\Pi}_i(\cdot)$. I make this distinction because $\pi_i(\cdot)$ is a firm’s profit through sales to consumers\footnote{More specifically, it can be seen as a firm’s variable profits minus the fixed costs.} and in particular, it is the result of competition between firms. The competitive effect of one firm on another is transmitted through $\pi_i(\cdot)$ but not $\ev_i$ in $\tilde{\Pi}_i(\cdot)$. This distinction will become important in the development of the measure of preemption in later texts.

### 2.2.6 Assumptions

Two assumptions for the model have been discussed above. I formally state these assumptions below:

**ASSUMPTION 1:** The variable $z_t$ has a discrete and finite support. That is, $z_t \in Z \equiv \{z_1, z_2, \ldots, z_{|Z|}\}$, where $|Z|$ is a finite number. The transition of $z_t$ is completely exogenous, independent of firms’ actions. It follows a Markov process with a transition probability function $f_z(z_{t+1}|z_t) : Z \times Z \to [0, 1]$.

**ASSUMPTION 2:** The private shock $\varepsilon_t$ is independently and identically distributed over time and across firms. That is, the transition probability of the Markov states, $p(\cdot|\cdot)$, can be factorized into

$$p(s_{t+1}, z_{t+1}, \varepsilon_{t+1}|a_t, s_t, z_t, \varepsilon_t) = p_\varepsilon(\varepsilon_{t+1}) f_z(z_{t+1}|z_t) I \{ (s_{At+1}, s_{Bt+1}) = (a_{At}, a_{Bt}) \}$$

where $p_\varepsilon(\cdot)$ is the joint density function of $\varepsilon_{t+1}$ and $I \{ \cdot \}$ the indicator function. The independence of $\varepsilon$ across firms implies $p_\varepsilon(\varepsilon_{t+1}) = g_A(\varepsilon_{At+1}) g_B(\varepsilon_{Bt+1})$, where $g_i(\cdot)$ is a density function that is positive and continuous everywhere. The density function $g_i(\cdot)$ is known to all firms.

Assumption 1, the exogenous market variable assumption, implies that whether firm $i$ decides to open a store at a market does not alter how the consumers’ socioeconomic status will evolve in the market\footnote{This assumption is more or less realistic depending on the industry under examination. In the case of fast food restaurants, small coffee shops or convenient stores, this assumption is realistic since very rarely people move to a certain area just because there is a small coffee shop nearby. For ethnic grocery supermarket chains, however, this assumption may not be very realistic. Ethnic groups often prefer to live in areas where ethnic foods are available. This assumption is made for simplicity of the expression of the measure of preemption developed in this paper. It can be relaxed.}. Assumption 2 is a standard assumption in Markov dynamic models with entry and exit. It ensures that a pure strategy MPE exists (Doraszelski and Satterthwaite, 2010), and in empirical applications, it helps researchers avoid the endogeneity problem associated with unobservable common shocks.

### 2.2.7 Strategies and Conditional Choice Probabilities

Firms are assumed to play stationary Markov strategies; that is, they choose the same actions at the same state regardless of time. The time script of the model can thus be omitted. Let $\sigma = (\sigma_B(s, z, \varepsilon), \sigma_A(s, z, \varepsilon))$ be a pair of strategy functions, with $\sigma_i : S \times Z \times R^2 \to A$. For each strategy profile $\sigma$, we can define a set of conditional choice probabilities (CCPs) $P^\sigma =$
\{(P^σ_A(1|s,z), P^σ_B(1|s,z)) : s \in S, z \in Z\},
with
\[ P^σ_i(1|s,z) \equiv \Pr(\sigma_i(s,z,\varepsilon_i) = 1|s,z) = \int I\{\sigma_i(s,z,\varepsilon_i) = 1\} g_\varepsilon(\varepsilon_i) \, d\varepsilon_i \]

Since there are only two choices in firms’ action set, the conditional choice probabilities for \(a = 0\) is determined by \(P^σ_i(0|s,z) = 1 - P^σ_i(1|s,z)\). Given our assumption regarding \(g_\varepsilon(\cdot)\), there is a one-to-one mapping between \(\sigma_i(s,z,\varepsilon_i)\) and \(P^σ_i(1|s,z)\).

### 2.2.8 Firms’ Problems

Let \(\hat{\pi}^σ_i(a_i,s,z)\) denote firm \(i\)’s expected flow payoff without \(\varepsilon_i(a_i)\) if it chooses action \(a_i\), and the other firm behaves according to strategy \(\sigma_{-i}\). If firm \(i\)’s expectation of the other firm’s actions is consistent with \(\sigma_{-i}\), then

\[
\hat{\pi}^σ_i(a_i,s,z) = \sum_{a_{-i} \in A} P^σ_i(a_{-i}|s,z) \pi_i(a_i,a_{-i},z) + ev_i(a_i,s,z).
\]

Let \(\hat{V}^σ_i(s,z,\varepsilon_i)\) denote firm \(i\)’s value if it acts optimally now and in the future given the other firm’s strategy \(\sigma_{-i}.\) Firm \(i\) faces the following problem:

\[
\hat{V}^σ_i(s,z,\varepsilon_i) = \max_{a_i \in A} \left\{ \hat{\pi}^σ_i(a_i,s,z) + \varepsilon_i(a_i) + \beta \sum_{(s',z') \in S \times Z} \left[ \int \hat{V}^σ_i(s',z',\varepsilon_i') g_\varepsilon(\varepsilon_i') \, d\varepsilon_i' \right] f^σ_i(s',z'|s,z,a_i) \right\},
\tag{2.2.1}
\]

where \(\beta \in (0, 1)\) is the discount factor and \(f^σ_i(s',z'|s,z,a_i)\) is the transition probability from \((s,z)\) to \((s',z')\) if firm \(i\) chooses \(a_i\) and the other firm plays strategy \(\sigma_{-i}\). Given that \(z\) is exogenous, we can write:

\[
f^σ_i(s',z'|s,z,a_i) = f_z(z'|z) \sum_{a_{-i} \in A} P^σ_i(a_{-i}|s,z) I\{ (s',z_{-i}) = (a_i,a_{-i}) \}.
\]

Denote \(\int \hat{V}^σ_i(s,z,\varepsilon_i) g_\varepsilon(\varepsilon_i) \, d\varepsilon_i\) by \(V^σ_i(s,z)\). Based on the Bellman equation (2.2.1), we can write

\[
V^σ_i(s,z) = \int \max_{a_i \in A} \{ \hat{\pi}^σ_i(a_i,s,z) + \varepsilon_i(a_i) \} g_\varepsilon(\varepsilon_i) \, d\varepsilon_i,
\tag{2.2.2}
\]

where \(\hat{\pi}^σ_i(a_i,s,z) \equiv \hat{\pi}^σ_i(a_i,s,z) + \beta \sum_{(s',z') \in S \times Z} V^σ_i(s',z') f^σ_i(s',z'|s,z,a_i)\). The function \(V^σ_i(s,z)\) is often referred to as the continuation value.

### 2.2.9 Markov Perfect Equilibria

The equilibrium concept is Markov perfect equilibrium, which is formally defined below:

**Definition 1.** A stationary Markov perfect equilibrium (MPE) is a pair of strategy functions \(\sigma^* \)
such that for any firm \( i \in \{A, B\} \) and any state \((s, z, \varepsilon_i) \in Z \times S \times R^2\), \( \sigma_i^* \) solves the Bellman equation (2.2.1).

### 2.2.10 MPE in Probability Space

For the computation of a MPE in this model, it is convenient to define an equilibrium in probability space. Aguirregabiria and Mira (2007) show that every MPE is a fixed point in mapping \( \Psi(P) \equiv \{\Psi_i(1|z, s; P) : (z, s) \in Z \times S\} \), and every fixed point in \( \Psi(\cdot) \) is a MPE. The mapping \( \Psi(\cdot) \) has the following expression:

\[
\Psi_i(1|s, z; P) = \int I \left( 1 = \arg \max_{a \in A} \left\{ \tilde{\pi}_P^i(a, s, z) + \varepsilon_i(a) \right\} + \beta \sum_{(s', z') \in S \times Z} \Gamma_i(s', z'; P) f_P^i(s', z'|s, z, a) \right) g_i(\varepsilon_i) d\varepsilon_i, \tag{2.2.3}
\]

where \( \tilde{\pi}_P^i(a, s, z) = \pi_P^i(a, s, z) \) and \( f_P^i(s', z'|s, z, a) \) are the density and the cumulative distribution functions of \( \varepsilon_i \) and \( \varepsilon_i^* \) when \( P \) are the set of CCPs. \( \Gamma_i \) is the expected value at state \((s', z')\) (integrated over \( \varepsilon_i^* \)) if both firms behave according to their CCPs in \( P \) in the future.

The expression for \( \Gamma_i(s, z; P) \) is

\[
\Gamma_i(s, z; P) = \sum_{a_i \in A} P_i(a_i|s, z) \left[ \tilde{\pi}_P^i(a_i, s, z) + e_P^i(a_i, s, z) \right] + \beta \sum_{(s', z') \in S \times Z} \Gamma_i(s', z'; P) f_P^i(s', z'|s, z) \tag{2.2.4}
\]

The value vector \( \Gamma_i(P) = \{\Gamma_i(s, z; P) : (s, z) \in S \times Z\} \) takes the following form:

\[
\Gamma_i(P) = (I - \beta F_P)^{-1} \left\{ \sum_{a_i \in A} P_i(a_i) \left[ \tilde{\pi}_P^i(a_i) + e_P^i(a_i) \right] \right\}, \tag{2.2.5}
\]

where \( I \) is the identity matrix; \( F_P \) is the transition matrix of states \((s, z)\) induced by \( F_P P_i(a_i) \) is a vector of dimension \(|Z|\), with \( P_i(a_i) \equiv \{P_i(a_i|s, z) : (s, z) \in S \times Z\} \); \( \tilde{\pi}_P^i(a_i) \) is a 4\(|Z|\) dimensional vector that stacks the corresponding state-specific element \( \tilde{\pi}_P^i(a_i, s, z) \); \( e_P^i(a_i) \) is a vector of the expected value of \( \varepsilon_i(a_i) \) conditional on \( a_i \) being chosen at a state \((s, z)\). That is, \( e_P^i(a_i) \equiv \{e_P^i(a_i, s, z) : (s, z) \in S \times Z\} \) and \( e_P^i(a_i, s, z) \equiv E(\varepsilon_i(a_i)|s, z) \sigma_1(s, z, \varepsilon_i) = a_i \).

Hotz and Miller (1993) show that this conditional expectation \( e_P^i(a_i, s, z) \) is a function of \( P_i(a_i|s, z) \) and the density function \( g_i(\cdot) \) only\(^{15}\).\(^{16}\)

\(^{15}\)Each element in \( F_P \) is \( f_P^i(s'|s, z) = \int f(z'|z) \sum_{(a_A, a_B) \in A^2} P_A(a_A|s, z) P_B(a_B|s, z) I \{(s', s') = (a_A, a_B)\} \).

\(^{16}\)When \( z \) follows an extreme value type I distribution, the function form of \( e_P^i(a_i, s, z) \) is \( \sigma_e (\gamma - \ln P_i(a_i, s, z)) \), where \( \sigma_e \) is the variance of \( z \), and \( \gamma \) is the Euler constant. When \( z \) follows a joint normal distribution \( \mathcal{N}(0, \Omega) \), the functional form is \( \frac{\varphi^{-1}(P_i(a_i|s, z))}{\sqrt{\varphi(\varepsilon_i^*) - \varphi(\varepsilon_i^0)}} \), where \( \varphi(\cdot) \) and \( \Phi(\cdot) \) are the density and the cumulative
Given that there is a one-to-one mapping between $\sigma$ and $P$, for simplicity of notation, I use $P_i(s, z)$ and $V_i(s, z; P)$ without the $\sigma$ superscript to represent firm $i$’s CCPs and continuation values respectively in the remaining text.

\section*{2.3 Definition of Preemptive Motives in a Simplified Model}

I define preemptive motives through a decomposition of the equilibrium conditions for each firm. To illustrate this decomposition and the definition of preemptive motives as intuitively as possible, I use a simplified version of the model as describe in Section \pageref{sec:preemptive-motives-def}. In this simplified model, the vector of exogenous state variable $z$ is assumed to be constant over time. The Markov states are fully described by the four distinct market structure states $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$. For simplicity of notation, I omit exogenous market variable $z$ in the representation of the state.

The decomposition of a firm’s equilibrium conditions separates a firm’s marginal benefit of entry into the marginal benefit of entry in the current period, and that in all future periods. As discussed previously, preemptive motives are embedded in the marginal benefit of entry from the future. I show below that this marginal benefit from the future can be broken down further into 3 components:

1. Return on investment through savings in entry costs or growth in scrap values (SAV component henceforth). For example, McDonald’s return on real estate investment can be seen as part of this component.

2. Return on entry in terms of a firm’s net present value (NPV) of flow profits, while holding its rival’s probability of being active constant across its states (SING component henceforth). That is, this component of the return of entry for a firm is evaluated under the restriction that its rival’s CCPs do not depend on whether it is in the market or not. This component captures the impact of a firm’s entry behaviors on its own profits, assuming that its incumbency status has no effect on its rival’s behavior.

3. Return on entry in terms of a firm’s NPV of flow profits through sales to consumers taking into account the dependency of the rival’s strategy on this firm’s incumbency status (DET component henceforth).

Of all three components above, the third one is a firm’s preemptive motive. The larger the preemptive motive, the more aggressively a firm will enter the market in the current period, and that is the act of preemption. These components’ specific algebraic form and the detailed decomposition of a firm’s entry (or stay) motives in the equilibrium conditions are shown in the following subsection.

\subsection*{2.3.1 Decomposition and Definition}

Here I focus on firm A’s problem, and firm B’s is analogous. For notational simplicity, I remove subscript $A$ from payoff, value functions, and choice probabilities, and use only the subscript $B$ to distribution function of the standard normal.

\footnote{In particular, the rival’s CCPs at those states when this firm is in the market are set to be the same as those when this firm is not.}
indicate the rival’s choice probabilities. Based on the expression in (2.2.4), firm A’s problem at state $s$ can be written as

$$V (s) \equiv \max_{P(s) \in (0,1)} \left[ P(s) P_B (s) \pi (1,1) + P(s) (1 - P_B (s)) \pi (1,0) + (1 - P(s)) P_B (s) \pi (0,1) + (1 - P(s)) (1 - P_B (s)) \pi (0,0) + \beta \{P(s) P_B (s) (V (1,1) + P(s) (1 - P_B (s)) V (1,0) + (1 - P(s)) P_B (s) V (0,1) + (1 - P(s)) (1 - P_B (s)) V (0,0)\} + I \{s = 0\} (-P(s) \text{EC}) + I \{s = 1\} (1 - P(s)) \text{SV} + P(s) e^P (1) + (1 - P(s)) e^P (0) \right]$$

(2.3.1)

The first order condition for the above optimization problem is

$$A^{-1} (P(s)) = I \{s_A = 0\} (-\text{EC}) + I \{s_A = 1\} (-\text{SV}) + P_B (s) (\pi (1,1) - \pi (0,1)) + (1 - P_B (s)) (\pi (1,0) - \pi (0,0)) + \beta (P_B (s) (V (1,1) - V (0,1)) + (1 - P_B (s)) (V (1,0) - V (0,0)))$$

(2.3.2)

where $A (\cdot)$ is an invertible function that maps the choice specific value differences onto the probability space $[0,1]$. This condition in (2.3.2) is both a sufficient and necessary condition for the equilibrium CCP $P(s)$. The left-hand side (LHS) of (2.3.2) is a function of only $P(s)$. It is the derivative of the conditional expectation of the private shocks with respect to $P(s)$. It represents a marginal cost of entry in terms of the expected values of the private shocks. The right-hand side (RHS) of (2.3.2) is the marginal benefit of being active in terms of entry costs and scrap values, flow profits and continuation values.

The RHS of (2.3.2) includes the marginal benefit of entry (or stay) in the current period and that from all future periods. The marginal benefit of entry in the current period is represented by the first two lines: the first line $I \{s_A = 0\} (-\text{EC}) + I \{s_A = 1\} (-\text{SV})$ is the entry cost that firm A pays by entering or the scrap value it foregoes by choosing to stay; the second line $P_B (s) (\pi (1,1) - \pi (0,1)) + (1 - P_B (s)) (\pi (1,0) - \pi (0,0))$ is the marginal benefit of entry (or stay) from firm A’s expected flow profits in the current period. The third line $\beta (P_B (s) (V (1,1) - V (0,1)) + (1 - P_B (s)) (V (1,0) - V (0,0)))$ represents the expected marginal benefit of entry (or stay) from all future periods.

In the third line, the difference between the continuation values $V (1, s_B) - V (0, s_B), s_B \in \{0,1\}$ represents the gain from entry (or stay) through many sources. For example, if firm A anticipates the entry cost to grow in the future, it may decide to enter the market early to save on entry costs. This savings is embedded in $V (1, s_B) - V (0, s_B), s_B \in \{0,1\}$ but is not related to preemption. To separate firms’ preemptive motives from other entry motives of firms, I break down the continuation values even further. Based on equation (2.2.5), the vector of continuation values $V$ can be written as

$$V = \left( I - \beta F_p \right)^{-1} \left\{ P^p \pi + ev + e \right\},$$

(2.3.3)

\(^{18}\)See Proposition 2 in Aguirregabiria and Magesan (2016)

\(^{19}\)Note that in this simplified model, the entry costs and scrap values stay constant because the exogenous market conditions are assumed to be time-invariant. If we include time-varying exogenous market conditions and allow the entry costs and scrap values to change from state to state, then the savings in entry costs are included in the difference between the continuation values because the incumbency status is worth more when the entry cost or scrap value is large.
where $\pi$, $e\nu$ and $e$ are column vectors with:

$$
\begin{align*}
\pi &= (\pi (0,0), \pi (0,1), \pi (1,0), \pi (1,1))^T \\
e\nu &= (-P (0,0) EC, -P (0,1) EC, (1 - P (0,0)) SV, (1 - P (1,1)) SV)^T \\
e &= \Delta (P).
\end{align*}
$$

The three terms within the curly brackets of (2.3.3) represent various sources of flow payoffs for firm A. Specifically, $F^P \pi$ represents firm A’s expected flow profits from static competition; $e\nu$ is firm A’s expected flow payoff from entry and exit costs; and $e$ is firm A’s expected flow payoff from the private information shocks. Of all three terms, $F^P \pi$ is the only term that is relevant to firm A’s preemptive motives. This is because firm B’s strategy $P_B (\cdot)$ does not enter into the other two terms $e\nu$ and $e$ (as shown by (2.3.4)). Firm B’s behaviors influence firm A’s flow payoffs only through $F^P \pi$, and if firm A preempts the entry of firm B, its gain from preemption would be reflected through the $F^P \pi$ term. Let

$$
\begin{align*}
\Upsilon_\pi &= (I - \beta F^P)^{-1} F^P \pi, \\
\Upsilon_{e\nu} &= (I - \beta F^P)^{-1} (e\nu + e).
\end{align*}
$$

Then

$$V = \Upsilon_\pi + \Upsilon_{e\nu},$$

and $\Upsilon_\pi$ and $\Upsilon_{e\nu}$ are the net present values (NPVs) of firm A’s flow payoffs from various sources discussed above. This is because $(I - \beta F^P)^{-1}$ converts all flow payoffs into NPVs.

It is evident that $\Upsilon_\pi$ is the only term in $V$ that is related to firm A’s preemptive motives. The difference between the continuation values can then be written as the differences between these terms:

$$V (1, s_B) - V (0, s_B) = \Upsilon_\pi (1, s_B) - \Upsilon_\pi (0, s_B) + \Upsilon_{e\nu} (1, s_B) - \Upsilon_{e\nu} (0, s_B).$$

The last difference $\Upsilon_{e\nu} (1, s_B) - \Upsilon_{e\nu} (0, s_B)$ can be seen as the return on investment through savings in entry costs or growth in scrap values. It represents the SAV component of firm A’s marginal benefit of entry (or stay) from the future as outlined at the very beginning of this section.

The SING and DET components of the marginal benefit from the future are intertwined in the difference term $\Upsilon_\pi (1, s_B) - \Upsilon_\pi (0, s_B), s_B \in \{0, 1\}$. That is, $\Upsilon_\pi (1, s_B) - \Upsilon_\pi (0, s_B), s_B \in \{0, 1\}$ includes both firm A’s preemptive motives and the effect of firm A’s behaviors on its own NPV of flow profits, holding fixed firm B’s behaviors across firm A’s states. To see this more clearly, we can

\[\text{Note that } e = \Delta (P) \text{ is a function of firm A’s CCPs; the specific functional form of } \Delta (P) \text{ depends on the distribution of } \varepsilon. \text{ See footnote [16] for the forms of } \Delta (P) \text{ for the distributions of extreme value type I and joint normal.}\]

\[\text{Although the other term } \Upsilon_{e\nu} \text{ has } P_B (s) \text{ in its expression through } (I - \beta F^P)^{-1}, \text{ it merely represents that firm B’s CCPs can affect the transition of the Markov states, and due to firm A’s own behaviors at each state, firm A may pay (or receive) different expected values of entry costs (or scrap values). This is a very indirect effect of firm B’s CCPs on firm A’s payoffs. It does not bear a close link to firm A’s preemptive motives and therefore should not be identified as part of firm A’s preemptive motives.}\]

\[\text{This is because the interpretation of the private shocks } \varepsilon \text{ relates mainly to entry costs and scrap values. Although the private shock } \varepsilon \text{ is also related to the fixed costs, given that the expected value of } \varepsilon \text{ conditional on entry and exit is entirely endogenous, attributing the source of } \varepsilon \text{ mainly to entry and exit costs is innocuous for the purpose of defining preemptive motives.}\]
write out $\Upsilon_\pi (1, s_B)$ and $\Upsilon_\pi (0, s_B)$ in their recursive forms:

$$
\begin{align*}
\Upsilon_\pi (1, s_B) &= P (1, s_B) P_B (1, s_B) (\pi (1, 1) + \beta \Upsilon_\pi (1, 1)) \\
&+ P (1, s_B) (1 - P_B (1, s_B)) (\pi (1, 0) + \beta \Upsilon_\pi (1, 0)) \\
&+ (1 - P (1, s_B)) P_B (1, s_B) (\pi (0, 1) + \beta \Upsilon_\pi (0, 1)) \\
&+ (1 - P (1, s_B)) (1 - P_B (1, s_B)) (\pi (0, 0) + \beta \Upsilon_\pi (0, 0))
\end{align*}
(2.3.8)

$$

$$
\begin{align*}
\Upsilon_\pi (0, s_B) &= P (0, s_B) P_B (0, s_B) (\pi (1, 1) + \beta \Upsilon_\pi (1, 1)) \\
&+ P (0, s_B) (1 - P_B (0, s_B)) (\pi (1, 0) + \beta \Upsilon_\pi (1, 0)) \\
&+ (1 - P (0, s_B)) P_B (0, s_B) (\pi (0, 1) + \beta \Upsilon_\pi (0, 1)) \\
&+ (1 - P (0, s_B)) (1 - P_B (0, s_B)) (\pi (0, 0) + \beta \Upsilon_\pi (0, 0))
\end{align*}
(2.3.9)

$$

The equations [2.3.8] and [2.3.9] show that $\Upsilon_\pi (1, s_B)$ and $\Upsilon_\pi (0, s_B)$ differ not only in terms of firm B’s CCPs $P_B (s)$ but also firm A’s CCPs $P (s)$. In general, the difference between $P (1, s_B)$ and $P (0, s_B)$ represents the effect of firm A’s behaviour on the NPV of its own profits, and the difference between $P_B (1, s_B)$ and $P_B (0, s_B)$ captures the gain accrued to firm A through the entry deterrence effect.

In particular, if $P_B (1, s_B) < P_B (0, s_B)$ and $P (1, s_B) = P (0, s_B)$, $\Upsilon_\pi (1, s_B)$ would have less weight on firm A’s duopoly profits than $\Upsilon_\pi (0, s_B)$. This implies that $\Upsilon_\pi (1, s_B)$ is likely to be greater than $\Upsilon_\pi (0, s_B)$. That is, firm A’s NPV of flow profits at its incumbency state $(1, s_B)$ is more valuable than that at its non-incumbency state $(0, s_B).

If we replace $P_B (1, s_B)$ with $P_B (0, s_B)$ in $\Upsilon_\pi (1, s_B)$, then we can control for this entry-deterrence effect, and the term resulting from this replacement would capture the NPV of firm A’s flow profits at state $(1, s_B)$ due to firm A’s own entry and exit behaviors without affecting firm B’s. Let $\hat{\Upsilon}_\pi (1, s_B)$ denote this resulting term after the replacement. Then $\hat{\Upsilon}_\pi (1, s_B)$ can be written as

$$
\hat{\Upsilon}_\pi (1, s_B) = P (1, s_B) P_B (0, s_B) \left( \pi (1, 1) + \beta \hat{\Upsilon}_\pi (1, 1) \right) \\
+ P (1, s_B) (1 - P_B (0, s_B)) \left( \pi (1, 0) + \beta \hat{\Upsilon}_\pi (1, 0) \right) \\
+ (1 - P (1, s_B)) P_B (0, s_B) \left( \pi (0, 1) + \beta \hat{\Upsilon}_\pi (0, 1) \right) \\
+ (1 - P (1, s_B)) (1 - P_B (0, s_B)) \left( \pi (0, 0) + \beta \hat{\Upsilon}_\pi (0, 0) \right),
\tag{2.3.10}
$$

where $\hat{\Upsilon}_\pi (\cdot)$ are obtained by

$$
\hat{\Upsilon}_\pi = \left( I - \beta \hat{F}^P \right)^{-1} \hat{F}^P \pi,
\tag{2.3.11}
$$

and $\hat{F}^P$ replaces in $F^P$ all $P_B (1, s_B)$ with $P_B (0, s_B)$, $\forall s_B \in \{0, 1\}$.

The reason that $\hat{\Upsilon}_\pi (1, s_B)$ has other $\hat{\Upsilon}_\pi (\cdot)$ terms on the RHS of [2.3.10] is because $\Upsilon_\pi (1, s_B)$ has a recursive nature. Replacing $P_B (1, s_B)$ with $P_B (0, s_B)$ for all periods will change not only $\Upsilon_\pi (1, s_B)$, but also all other $\Upsilon_\pi (\cdot)$ terms. In short, $\hat{\Upsilon}_\pi (1, s_B)$ represents firm A’s NPV of flow profits at state $(1, s_B)$ when the entry deterrence effect of its incumbency status on firm B is unsuccessful. Since $\Upsilon_\pi (1, s_B)$ is that for the case when entry deterrence is successful, $\Upsilon_\pi (1, s_B) - \hat{\Upsilon}_\pi (1, s_B), \forall s_B \in \{0, 1\}$ then captures the gain in firm A’s NPV of flow profits through preemption. This is firm A’s
preemptive motives.

To summarize, firm A’s marginal benefit of entry (or stay) from the future can be represented by the difference between its continuation values $V(1, s_B) - V(0, s_B), s_B \in \{0, 1\}$, and this difference can be separated into three components:

$$V(1, s_B) - V(0, s_B) = \Upsilon_{\text{eve}}(1, s_B) - \Upsilon_{\text{eve}}(0, s_B)$$

$$+ \tilde{\Upsilon}_\pi(1, s_B) - \Upsilon_\pi(0, s_B)$$

$$+ \Upsilon_\pi(1, s_B) - \hat{\Upsilon}_\pi(1, s_B)$$

(2.3.12)

The first component on the first line is the return on investment through savings in entry costs or growth in scrap values (SAV). The second component on the second line is the return on entry in terms of a firm’s net present value (NPV) of flow profits due to firm A’s own entry and exit behaviors without affecting its rival (SING). The third component on the last line is the return on entry in terms of firm A’s NPV of flow profits through preemption (DET). This is firm A’s preemptive motive. Firm B’s preemptive motives can be analogously identified. Definition 2 below formally states the definition of both firms’ preemptive motives:

**Definition 2.** The preemptive motives of firm A are $\Upsilon_{\pi A}(1, s_B) - \hat{\Upsilon}_{\pi A}(1, s_B), s_B \in \{0, 1\}$, and those of firm B are $\Upsilon_{\pi B}(s_A, 1) - \hat{\Upsilon}_{\pi B}(s_A, 1), s_A \in \{0, 1\}$

It is worth noting that these preemptive motive terms can also be seen as the “sacrifice” in firms’ current-period flow profits by entering (or staying) aggressively today. For example, in firm A’s equilibrium conditions, as shown by equation (2.3.2), for a given level of $P(s)$, if firm A’s preemptive motives are large and positive, then its marginal benefit of entry (or stay) from all future periods would also be large and positive. This implies that firm A would be willing to make a “sacrifice” in its current-period flow profits when choosing a high level of $P(s)$. In other words, by entering (or staying) aggressively today, firm A is happy to make a short-term “sacrifice” because it could recoup all its loss from the preemptive gain in the future. In this respect, the preemptive motive terms represent both the long-term gain from preemption and the short-term “sacrifice” of preemption.

### 2.3.2 Measuring the Contribution of Preemption to Firms’ Behaviors

The definition in the previous section provides a measure of how much entry deterrence contributes to the NPV of a firm. At the same time, we are also interested in how much entry deterrence contributes to firms’ behaviors, and this section provides such a measure. To do so, I first construct a counterfactual for how firms will behave without preemptive motives. Then, by comparing the equilibrium with the counterfactual, I obtain a preemption index to measure the contribution of entry deterrence to firms’ behaviors.

---

23The terms $\Upsilon_{\pi B}(s_A, 1) - \hat{\Upsilon}_{\pi B}(s_A, 1), s_A \in \{0, 1\}$ are similarly constructed as those for firm A.
2.3.2.1 Counterfactual of Preemption

I design the counterfactual for preemption by forcing firms to ignore their preemptive motives and re-optimize their values. Forcing firms to ignore their preemptive motives is equivalent to setting the preemptive motive terms to 0 in firms’ equilibrium conditions. Here I use again firm A’s equilibrium conditions as an example to illustrate the construction of firms’ counterfactual CCPs.

Setting \( \Upsilon_\pi (1, s_B) - \tilde{\Upsilon}_\pi (1, s_B) = 0, s_B \in \{0, 1\} \) in firm A’s equilibrium conditions (2.3.2) implies that \( \Upsilon_\pi (1, s_B) = \tilde{\Upsilon}_\pi (1, s_B), s_B \in \{0, 1\}, \) and in their specific algebraic forms:

\[
\Upsilon_\pi (1, 0) = \frac{1}{\det (I - \beta F^P)} \det \begin{pmatrix} F_{c1}^P & F_{c2}^P & F^P \pi & F_{c4}^P \end{pmatrix} \\
= \tilde{\Upsilon}_\pi (1, 0) = \frac{1}{\det (I - \beta \hat{F}^P)} \det \begin{pmatrix} \hat{F}_{c1}^P & \hat{F}_{c2}^P & \hat{F}^P \pi & \hat{F}_{c4}^P \end{pmatrix} \\
\Upsilon_\pi (1, 1) = \frac{1}{\det (I - \beta F^P)} \det \begin{pmatrix} F_{c1}^P & F_{c2}^P & F_{c3}^P & F^P \pi \end{pmatrix} \\
= \tilde{\Upsilon}_\pi (1, 1) = \frac{1}{\det (I - \beta \hat{F}^P)} \det \begin{pmatrix} \hat{F}_{c1}^P & \hat{F}_{c2}^P & \hat{F}_{c3}^P & \hat{F}^P \pi \end{pmatrix},
\]

where \( F_{c1}^P \) and \( \hat{F}_{c1}^P, i = 1, 2, 3, 4, \) are the 1st to the 4th columns of matrices \( F^P \) and \( \hat{F}^P \) respectively.

For the condition \( \Upsilon_\pi (1, s_B) = \tilde{\Upsilon}_\pi (1, s_B), s_B \in \{0, 1\} \) to hold for any parameterization of the model, it has to be that \( F^P = \hat{F}^P. \) That is, in the counterfactual scenario, \( P_B (1, s_B) \) is replaced with \( P_B (0, s_B) \) in the \( \Upsilon_\pi (1, s_B) \) term. Again, due to the recursive nature of \( \Upsilon_\pi (s) \), if \( F^P = \hat{F}^P, \) then

\[
\Upsilon_\pi (s) = \tilde{\Upsilon}_\pi (s), s \in S
\]

This is to say that in the counterfactual, to eliminate firm A’s preemptive motives, we need to replace all \( \Upsilon_\pi (\cdot) \) with \( \tilde{\Upsilon}_\pi (\cdot) \) in firm A’s equilibrium conditions (2.3.2). This yields:

\[
A^{-1} (P (s)) = (s = 0) (-EC) + (s = 1) (-SV) \\
+ P_B (s) (\pi (1, 1) - \pi (0, 1)) + (1 - P_B (s)) (\pi (1, 0) - \pi (0, 0)) \\
+ \beta \left( P_B (s) \left( \tilde{\Upsilon}_\pi (1, 1) - \tilde{\Upsilon}_\pi (0, 1) \right) + (1 - P_B (s)) \left( \tilde{\Upsilon}_\pi (1, 0) - \tilde{\Upsilon}_\pi (0, 0) \right) \right) \\
+ \beta \left( P_B (s) (\pi_X (1, 1) - \pi_X (0, 1)) + (1 - P_B (s)) (\pi_X (1, 0) - \pi_X (0, 0)) \right) (2.3.13)
\]

\( \forall s \in S. \)

The set of conditions in (2.3.13) for firm A and a similar set of conditions for firm B jointly determine the counterfactual CCPs of both firms. These conditions differ from firms’ equilibrium conditions only in the \( \tilde{\Upsilon}_\pi (\cdot) \) terms, and they can be interpreted as follows:

When making entry and exit decisions, firm A considers the marginal benefit of entry (or stay) in both the current period and all future periods. Although firm A knows that firm B may enter (stay) less likely at state \((1, s_B)\) compared to state \((0, s_B)\), it does not take this information into consideration when calculating its NPVs of flow profits in all future periods, i.e. \( \Upsilon_\pi (s) \). In particular, firm A assumes that firm B’s likelihood of being in the market at state \((1, s_B)\) is the same as that

\[24\]Recall that \( \Upsilon_\pi \equiv (I - \beta F^P)^{-1} F^P \pi \) and \( \tilde{\Upsilon}_\pi \equiv (I - \beta \hat{F}^P)^{-1} \hat{F}^P \pi. \) Based on Cramer’s rule, specific elements of the \( \Upsilon_\pi \) and \( \tilde{\Upsilon}_\pi \) vectors can be written out in the form of the ratio between the determinants of the corresponding matrices.

\[25\]A similar set of conditions can be obtained for firm B.
for state \((0, s_B)\). This assumption eliminates the preemptive motives.

This counterfactual makes a distinction between a firm’s perception of its rival’s behaviors in future periods and the rival’s actual behaviors. These two are not the same in this counterfactual. The rival, say firm \(-i \in \{A, B\}\), can still hesitate at entering (or staying in) the market when firm \(i\) is an incumbent due to the direct competitive effect: for example, if firm \(i\) has a low scrap value and can commit to staying in the market after entry, then firm \(-i\) would be reluctant at entering the market because its profits in the current period would mostly likely come from duopoly profits. However, firm \(i\) does not perceive firm \(-i\) to behave this way when calculating its own NPV of flow profits, \(\Upsilon_{s_i}(s)\). In this aspect, firms are assumed to have bounded rationality in this counterfactual.

It is important to note that while eliminating preemptive motives, this counterfactual retains direct competitive effects. These effects are captured mostly through firms’ current-period competition as represented by the second lines in \((2.3.13)\). These direct competitive effects allow firms to react to their rivals’ entry costs and scrap values. For example, according to \((2.3.13)\), if firm A’s scrap value is smaller than its entry cost, then its CCPs at states with \(s_A = 1\) are likely bigger than those at states with \(s_A = 0\), i.e. \(P(1, s_B) > P(0, s_B), \forall s_B \in \{0, 1\}\). Knowing this, firm B is likely to enter (or stay in) the market less frequently at the states with \(s_A = 1\) than those states with \(s_A = 0\), i.e. \(P_B(1, s_B) < P_B(0, s_B)\). The same analysis also applies to firm B. If firm B’s scrap value is much lower than its entry cost, it is likely that \(P_B(s_A, 1) > P_B(s_A, 0)\) and \(P(s_A, 1) > P(s_A, 0), \forall s_A \in \{0, 1\}\). These behaviors patterns are naturally expected in a competition: when firms’ scrap values are less than their entry costs, a firm would enter (or stay at) a market less frequently if that market is already occupied by its rival.

This counterfactual also preserves the dynamic features of the game through the third and fourth lines of \((2.3.13)\), where firms’ considerations of future payoffs are also taken into account. In short, this counterfactual retains both the direct effects of competition and the dynamic features of the game.

---

\(^{26}\)This pattern of behavior can be easily inferred through the second line in firm B’s conditions, \(P(s) (\pi_B(1, 1) - \pi_B(1, 0)) + (1 - P(s)) (\pi_B(0, 1) - \pi_B(0, 0))\), where firm A’s CCP \(P(1, s_B)\) puts more weight on \(\pi_B(1, 1) - \pi_B(1, 0)\), the negative part of firm B’s marginal benefit of entry in the current period. Of course, firm B’s CCPs are also affected by the marginal benefit of entry in the NPV of future payoffs, but they usually work in the same direction.

\(^{27}\)One may argue that there are alternative ways to construct \(\Upsilon_{s}(\cdot)\) if the underlying principle of the counterfactual is the bounded rationality of firms. For example, \(P_B(1, s_B)\) and \(P_B(0, s_B)\) can be set to any number between 0 and 1 as long as they are set to be equal to each other. I here argue that that proposal is not a good way to design the counterfactual because fixing \(P_B(1, s_B)\) and \(P_B(0, s_B)\) to a pre-specified (or exogenous) number is equivalent to changing the economic fundamentals for firms in future periods; that is, we are changing firms’ profits in the future exogenously. However, preemption is an endogenous phenomenon: it depends on firms’ expectation of their profits in future periods, which change endogenously according to both their own behaviors and their expectations of their rivals’ behaviors. In this regard, pre-fixing \(P_B(1, s_B)\) and \(P_B(0, s_B)\) to an exogenous number is not consistent with the endogenous nature of preemption. Nonetheless, these alternative counterfactuals have been implemented for the numerical examples in Section 2.4 and they create results that are not as desirable as the one advocated in this paper. These materials are available from the author upon request.
2.3.2.2 Preemption Index

To compare firms’ equilibrium CCPs with those in the counterfactual, I propose two types of preemption indices: (1) probability-based preemption index and (2) duration-based preemption index. Within the duration-based preemption index, I differentiate the index for individual firms from that for all firms. The probability-based preemption index is a direct comparison of firms’ CCPs from the counterfactual to those in the equilibrium. The unit of measure is in probabilities. The duration-based index translates firms’ CCPs into the timing of entry and exit through a simple survival formula. These indices are different ways of presenting how firms’ behaviors will change without preemptive motives. In some cases, one index illustrates an aspect of the results better than another.

Probability-Based Preemption Index

This preemption index directly compares firms’ counterfactual CCPs with the equilibrium ones. This index has the following formula:

\[ PI_i = \sum_s \left( P^{ctf}_i(s) - P^{MPE}_i(s) \right) f(s), \forall i \in \{A,B\}, s \in S, \]

where \( P^{ctf}_i(s) \) is firm \( i \)'s probability of being active at state \( s \) in the counterfactual, \( P^{MPE}_i(s) \) is the one in the MPE, and \( f(s) \) is the probability mass function of the steady-state distribution calculated based on firms’ MPE CCPs.

In this formula, the differences between firms’ counterfactual and MPE CCPs are weighted by the steady-state distribution \( f(s) \). This is because in most applications, the number of Markov states is large, and it is difficult to present the CCP difference for each individual state. A weighted average is thus taken for the convenience of presentation.

The probability-based preemption index is useful for demonstrating how firms’ strategies change without preemptive motives; however, the probabilities themselves could be difficult to interpret in some context. A more intuitive way to present firms’ behaviors is in terms of the timing of entry or exit, and this is done in the duration-based preemption index.

Duration-Based Preemption Index

To convert firms’ CCPs into the timing of entry and exit, I use a simple duration analysis. The intuition behind the duration analysis is as follows: if a firm is not an incumbent in a market at a given state, its entry probability at that state informs us of how long on average it takes this firm to enter the market. Similarly, if a firm is an incumbent, its exit probability implies how long it takes this firm to exit the market.

The duration-based preemption index for individual firms has the following form:

\[ WD_i = \sum_s f(s) \left\{ I\{s_i = 0\} \left[ D^{ctf}_i(s) - D^{MPE}_i(s) \right] + I\{s_i = 1\} \left[ D^{MPE}_i(s) - D^{ctf}_i(s) \right] \right\}, \forall i \in \{A,B\}, \]

(2.3.14)
where

\[ D_{ctf}^i (s) = \begin{cases} \frac{1}{P_{ctf}^i (s)} & \text{if } s_i = 0 \\ 1 - \frac{1}{P_{ctf}^i (s)} & \text{if } s_i = 1 \end{cases} \]

\[ D_{MPE}^i (s) = \begin{cases} \frac{1}{P_{MPE}^i (s)} & \text{if } s_i = 0 \\ 1 - \frac{1}{P_{MPE}^i (s)} & \text{if } s_i = 1 \end{cases} \]

\( D_{ctf}^i \) and \( D_{MPE}^i \) are the average length of time for firm \( i \) not to change its status. Their formulas come from the survival rate of \( s_i \), a standard formula for expected duration of survival. For state \( s_i = 0 \), the survival rate is \( P_{ctf}^i (0|s) = 1 - P_{ctf}^i (s) \). For state \( s_i = 1 \), the survival rate is \( P_{ctf}^i (s) \).

By this construction, durations at state \( s_i = 0 \) represent the average time that firm \( i \) will wait before entering a market, and those at state \( s_i = 1 \) represent the average wait time before exiting a market. The weighted difference between counterfactual and MPE durations, \( WD_i \), represent the additional length of time that firm \( i \) would be absent from the market in the counterfactual compared to the equilibrium.

It should be noted that \( WD_i \) is constructed for an individual firm in isolation of their rivals. Although one firm may not be in a market, it is not necessarily true that the market is not served by another firm. To see how much additional time a market will be served by no firm at all in the counterfactual, I construct the duration-based preemption index for all firms. It has the following form:

\[ WAD = \sum_s f (s) \left\{ I \{s = (0, 0)\} \left[ AD_{ctf}^i (s) - AD_{MPE}^i (s) \right] + I \{s \neq (0, 0)\} \left[ AD_{MPE}^i (s) - AD_{ctf}^i (s) \right] \right\} \]

where

\[ AD_{ctf}^i (s) = \begin{cases} \frac{1}{1 - \prod_{i=1}^{\infty} \left(1 - P_{ctf}^i (s) \right)} & \text{if } s = (0, 0) \\ \prod_{i=1}^{\infty} \left(1 - P_{ctf}^i (s) \right) & \text{if } s \neq (0, 0) \end{cases} \]

\[ AD_{MPE}^i (s) = \begin{cases} \frac{1}{1 - \prod_{i=1}^{\infty} \left(1 - P_{MPE}^i (s) \right)} & \text{if } s = (0, 0) \\ \prod_{i=1}^{\infty} \left(1 - P_{MPE}^i (s) \right) & \text{if } s \neq (0, 0) \end{cases} \]

This index for all firms highlights the impact of preemption on the state of the overall market. It represents how much longer a market would remain unserved by any firm if firms behave non-preemptively. Examples of applying these preemption indices in an empirical analysis are shown in Section 2.5.

2.3.3 Extension to Models with Stochastic Demand, Multiple Firms and Multiple Locations

The definition and counterfactual of preemption developed in this section can be extended to models with stochastic demand, multiple product locations, and multiple players. Given that the algebraic expressions of these extensions are somewhat involved, I leave the details to the Appendix. Section 3 of the Appendix makes the extension to the general model with stochastic demand. Section 4 of the Appendix develops the extension for a model with two locations and stochastic demand. The
extension to a dynamic game with multiple players is shown in Subsection 2.5.4.1. In the following two sections, I implement the counterfactual and measure of preemption in a numerical exercise and an empirical application.

2.4 A Numerical Application

A frequently asked question about spatial preemption is what is the source of strategic commitment for preemption. This question is closely related to the seminal work by Judd (1985), who argues that sunk costs do not matter for preemption and a large exit cost is imperative for an incumbent to preempt entry. While Judd’s conclusion holds under the stylized model in that paper, it does not hold for the more realistic model in this paper. The numerical analysis below shows that sunk costs matter for preemption, and that firms do not need a large exit cost to preempt entry. In particular, the source of strategic commitment for firms is not limited to exit costs.

In the model of this paper, the source of strategic commitment includes two factors: (1) the sunk part of the entry cost, i.e. $EC - SV$ (hereafter the sunk cost), and (2) the “excessive” preemptive gain that is not warranted by economic fundamentals. The second factor arises under a setting of multiple equilibria. When multiple MPEs exist, some MPE exhibit a much higher level of aggression in firms’ strategies than others even though firms have the same sunk costs across MPEs. These multiple MPEs can be ranked by their level of aggression (or preemption) based on their distance to the counterfactual. The difference between the preemptive motive in the most preemptive MPE and that in the least preemptive one captures the “excessive” preemptive gain.

The key reason for the overturn of Judd’s (1985) conclusion hinges on two features of the model of this paper. One is the infinite-horizon setup, which allows firms to continuously commit to staying in the market after entry because there is always a future gain in doing so, while in Judd’s model, which is a finite-horizon game, the incumbent cannot commit in the last period. The other key feature of the model in this paper is the incomplete information setup. With incomplete information, firms’ strategies in terms of CCPs are probabilistic, and they respond much more sensitively to changes in scrap values (or exit costs) than the binary strategies in Judd’s model in the complete information setting. The greater sensitivity of firms’ strategies in response to changes in exit costs allows exit costs to have a bigger impact on firms’ preemptive behaviors in the incomplete information model. Therefore, in the model of this paper, a firm can preempt entry without a large exit cost.

Subsection 2.4.1 below describes the setup of the model parameters for the numerical examples in this section. Subsection 2.4.2 examines how firms’ equilibrium strategies change with respect to sunk costs ($EC - SV$), and by comparing the equilibrium strategies with those in the counterfactual, I illustrate how sunk costs influence preemption.

\[^{28}\text{In this paper, I relax the restriction of only one exit and no reentry opportunity in Judd’s model. I assume that firms have unlimited opportunities to enter and exit and re-enter and re-exit.}\]
2.4.1 Parameters of the Simplified Model

The model for the numerical analysis is the simplified model as described in Section 2.3. Firms’ flow profits $\pi_i(\cdot), i \in \{A, B\}$ are set as follows

<table>
<thead>
<tr>
<th></th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1 0.5, -0.1 4.5, 0.0</td>
</tr>
<tr>
<td></td>
<td>0 2.0, 2.0 5.0, 0.0</td>
</tr>
</tbody>
</table>

Table 2.1: Flow Profits $\pi_i(\cdot)$

These payoffs can be seen as coming from a Hotelling line market with two locations: L1 and L2. Consumers are evenly distributed along the line. Firm A has a store at L1, and both firms decide whether to enter L2 or not. The entry location is restricted to L2 only, and the total flow profit functions for each firm regarding their actions at L2 are specified in Table 2.1. The purpose of including the L1 location in the market is to make the cannibalization effect between firm A’s products at both locations easier to conceptualize. As shown in the table, because firm A already sells a product at L1, both its monopoly and duopoly profits from operating at both locations ($\pi_A(1, 0) = 4.5$ and $\pi_A(1, 1) = 0.5$) are smaller than those with being active at only L1 not L2 ($\pi_A(0, 0) = 5$ and $\pi_A(0, 1) = 2.0$) due to cannibalization.

This specification of firms’ flow profits are similar to the setting as in Judd (1985), where the incumbent firm A does not want to expand to location L2 without the threat of entry from firm B. With this specification, in a simultaneous-move game, as in the second last stage in Judd (1985), the incumbent firm A will choose not to be active at location L2, unless it has a large exit cost. As will be shown later, this result does not hold for the dynamic game in this paper.

Other parameters of the model are set as such: $EC_A = EC_B = 1$, $\beta = 0.95$, and $\varepsilon_i(a)$ follows an extreme value type I distribution with the location parameter 0 and scale parameter $\sigma_\varepsilon = 0.8$. Here I do not specify $SV_A$ and $SV_B$; instead, I let them gradually decline from the levels of entry costs, and in doing so, I examine how sunk costs influence firms’ preemptive behaviors.

2.4.2 How Sunk Costs Influence Preemption

Two numerical examples are shown in this section: Case 1 and 2. In Case 1, I decrease only firm A’s scrap value $SV_A$ from the level of $EC_A$ to $-11$, while holding firm B’s scrap value constant at $EC_B = 1$. Through this case, I examine specifically how a firm’s absolute value of sunk cost influences preemption. In Case 2, I reduce both firms’ scrap values at the same rate from $EC_A = EC_B = 1$ to $-11$, while maintaining a constant ratio of the two firms scrap values. Through this case, I investigate how the ratio of the two firms’ scrap values affects preemption. I find that both

---

Fixing the ratio of $SV_A$ to $SV_B$ at 1 is not necessary. The ratio can be much smaller or larger than 1. As long as the ratio is less than a threshold, i.e. firm B’s scrap value $SV_B$ is below a certain level or its sunk cost is larger than a threshold, the general results as shown by Case 2 hold. The numerical exercises for other ratio levels have been conducted and are available from the author upon request.
factors are important.

For both numerical exercises, to systematically explore the equilibrium and counterfactual correspondences along the parameter of scrap values, I use a homotopy method, which is explained in detail in Section 1 of the Appendix. The homotopy method requires a starting point for both the MPE correspondence and the counterfactual one. To obtain these starting points, I take advantage of the fact that the MPE for this game at the parameters \( EC_i = SV_i, \forall i \in \{A, B\} \) is also a counterfactual. The intuition behind the MPE overlapping with the counterfactual when \( EC_i = SV_i, \forall i \in \{A, B\} \) is that when firms’ scrap values and entry costs are equal, firms’ sunk costs are 0 and they cannot commit to staying in the market after entry. In particular, firms are not hesitant at leaving the market because upon exit, they can retrieve all their initial investment from the scrap values. Knowing that a firm cannot commit, its rival would not be afraid of entering the market at that firm’s incumbency state. In this sense, a firm’s incumbency state no longer has an entry deterrence power, and firms cannot preempt their rivals by aggressive entry.

I call the MPE that is also a counterfactual the benchmark MPE. The benchmark MPE is then non-preemptive. In particular, it is very easy to solve for computationally. Section 2 of the Appendix provides a simple algorithm for obtaining firms’ CCPs in a benchmark MPE. Because of the non-preemptive nature of the benchmark MPE and its computational simplicity, I use it as the starting point in the homotopy method to systematically trace out the equilibrium and counterfactual correspondences along the dimension of the scrap values.

**Case 1** \( SV_A \in [1, -11], SV_B = 1 \)

As mentioned above, here I decrease firm A’s scrap value gradually from \( EC_A \) to -11, while holding firm B’s constant at \( EC_B = 1 \). Since firm A’s scrap value is always less than its entry cost, firm A can commit to staying at L2 after entry to a certain extent, whereas firm B cannot. Under this setup, firm B would always respond to firm A’s incumbency status by reducing its probability of entry (or stay) at firm A’s incumbency state. Anticipating this, firm A will enter aggressively in the current period in order to preempt the entry of firm B.

Figure 1 plots the correspondence of firms’ equilibrium and counterfactual CCPs at each state against firm A’s scrap value. Firms’ equilibrium CCPs represent how they will behave with preemptive motives, and the counterfactual CCPs illustrate how their behaviors change once the preemptive motives are eliminated. The top four panels illustrate firm A’s CCPs at each state, and the bottom four show those for firm B. The blue lines on the graph represent the equilibrium correspondence, and the red lines the counterfactual one. As shown on the graphs, at the origin of the horizontal axis, both firms’ scrap values equal their entry costs, and the MPE and the counterfactual lines overlap at the benchmark MPE. In this non-preemptive benchmark MPE, firm A chooses to be active at L2 with a probability of around 14.2% and firm B with a probability of about 88.7% at

---

30 Section 2 of the Appendix provides a general proof of this result.

31 Because it is also a counterfactual where firms’ preemptive motives are 0.
every state. These probabilities are consistent with firms’ flow profit structure shown in Table 2.1 without preemptive motives, firm A would not enter L2 with a high probability, but firm B would.

As firm A’s scrap value declines, the MPE and the counterfactual correspondences start to diverge. The lower firm A’s scrap value, the further apart these two correspondences. In particular, firm A always enters (or stays) with a greater probability in the MPE than in the counterfactual, whereas firm B does the reverse. The gap between firms’ CCPs in the MPE and those in the counterfactual indicates that firm A acts preemptively, and the extent of preemption increases as firm A’s scrap value decreases.

Figure 1: MPE and Counterfactual CCPs in Case 1

To illustrate the preemptive motives behind firms’ equilibrium strategies, I plot in Figure 2 the preemptive motive terms as shown in Definition 2 for both firms. For convenience, I name the term $\Upsilon_{\pi_A} (1, 0) - \hat{\Upsilon}_{\pi_A} (1, 0)$ “firm A’s preemptive motive for $s_B = 0$,” and all the other terms are similarly labelled. The top two panels in Figure 2 represent the preemptive motive terms for firm A, and the bottom two for firm B. As can be seen, at the benchmark MPE, all firms’ preemptive motives are 0, which is consistent with the fact that the benchmark MPE is non-preemptive. For firm A, its preemptive motive grows as its scrap value decreases. In particular, as firm A’s scrap value reaches

---

\[32\] In the benchmark MPE, firms’ CCPs are the same across all market structure states. See Section 2 of the Appendix for an explanation and the general proof.
around \(-5.2\), there is a discontinuous jump in the preemptive motive terms from about 2 to 33. This jump can be seen as a change of regime in firm A’s preemptive motives from relatively weak to very strong. For firm B, the preemptive motive terms are always 0, which is consistent with the fact that firm B does not commit to staying in the market after entry.

An interesting phenomenon as shown in Figures 1 and 2 is that although firm A’s preemptive motives increase monotonically with the decline in firm A’s scrap value, firm A’s entry probabilities at states \((0, 0)\) and \((0, 1)\) follow a non-monotonic relationship with the scrap value. They first decrease as the scrap value declines, and once the scrap value passes the threshold at \(-5.2\), they start increasing. The initial decline in these entry probabilities indicates that although a lower scrap value increases firm A’s ability to commit and allows it to gain more from preemption, the growth in preemptive motives is not enough to induce a rise in firm A’s entry probabilities; there is another effect at work that makes firm A less likely to enter. This other effect is a direct structural effect. It reduces firm A’s expected value from entry as the scrap value declines because in case business fails in the future, firm A would have to incur a greater loss to exit. This direct structural effect thereby discourages firm A from entry. The first effect of the lower scrap value through the channel of rising preemptive motives can be called the preemptive strategic effect of the scrap value. In general, both the structural effect and the strategic effect are at work as the scrap value declines, and they work in different directions. If the structural effect dominates, firm A’s entry probabilities would decrease with a lower scrap value. If the strategic effect dominates, firm A’s entry probabilities would increase with a smaller scrap value. This dominance of the strategic effect explains why firm A’s entry probabilities increase with the decline in the scrap value after the scrap value reaches
Source of Strategic Commitment and Least Preemptive MPE

As shown in the analysis above, a lower scrap value provides a source of strategic commitment for firm A. It should be noted, however, for the same scrap value, firm A can have different levels of strategic commitment, as shown by the multiple MPEs that exist in the region between $SV_A = -1.8$ and $SV_A = -5.2$. Figure 3 enlarges the top third panel of Figure 1 and the top-left panel in Figure 2. As can be seen, when $SV_A = -3$, there are three MPEs (A, B, and C), and firm A’s strategies are the most aggressive in MPE A and the least aggressive in MPE C: its probability of staying at L2 is much higher in MPE A than in MPE B or in MPE C. This difference in firm A’s probability of staying implies different levels of strategic commitment across these MPEs, and the source of the difference in strategic commitment across MPEs comes from the “excessive” preemptive gain as firm A’s strategy moves from the least aggressive MPE C to the most aggressive MPE A. As shown in the right panel of Figure 3, the gap between the preemptive motive in MPE A and that in MPE C captures the “excessive” preemptive gain that is not warranted by the scrap value and other economic fundamentals.

Figure 3: Multiple MPEs

Since the counterfactual in this case is unique (Point D), we can rank the MPEs A, B and C by their levels of preemption based on their distance to the counterfactual (D). By this criterion, MPE C is the least preemptive of all MPEs. Notice that the least preemptive MPE, C, is connected to the benchmark MPE through a continuous mapping. This quality implies that if firms start from the benchmark MPE and then gradually adjust their strategies in response to the decline in scrap

\[ SV_A = -5.2 \]

---

33 The structural effect and the strategic effect of sunk costs and how they induce a non-monotonic relationship between a firm’s entry probabilities and sunk cost have also been documented in Cabral and Ross (2008), though in a context of a finite-horizon sequential-move game with Cournot competition.
values, they would end up at MPE C. This type of gradual adjustment makes MPE C maintain the relatively accommodative nature of the benchmark MPE, which is consistent with MPE C being the least preemptive. MPEs that are connected through a continuous equilibrium mapping are often called “the same type” of MPE in the literature. As shown here, the benchmark MPE and MPE C are indeed of the same type.

**Case 2** \( SV_A \in [1, -11], SV_B \in [1, -11] \)

In Case 2, I decrease both firm A and firm B’s scrap values at the same rate from \( EC_A = EC_B = 1 \) to \(-11\), while holding the ratio of firm A’s scrap value to firm B’s constant at \( 1/3 \). Under this setup, both firms can commit to staying at L2 after entry to a certain extent, and as a result, both firms’ incumbency status has an entry deterrence effect on their rivals. However, the direct structural effect of a lower scrap value can make firms respond to their own states more than their rivals’ states; this effect undermines the entry deterrence effect of the rival’s incumbency status. How much firm A would be able to preempt the entry of firm B or vice versa is not as straightforward as in Case 1.

Firms’ strategies are shown in Figure 4, which plots the correspondence of firms’ equilibrium and counterfactual CCPs at each state against both firms’ scrap values. It can be seen that firms behave mostly according to their own states in the MPE scenario, and although they do respond to their rivals’ incumbency status, the entry deterrence effect is not very large. As a result, firms’ strategies are very similar across the MPE and the counterfactual scenarios, which implies that firms’ preemptive motives are not strong in this case. This is further confirmed by Figure 5, which plots the correspondence of firms’ preemptive motives against scrap values. As can be seen, both firm’s preemptive motives are very small.

---

34See Aguirregabiria (2012) and Doraszelski and Escobar (2010).
35Again, here fix the ratio of \( SV_A \) to \( SV_B \) at 1 is not necessary. Firms’ preemptive behaviors show a similar pattern as those in Case 2. The numerical exercises for other ratio values have been conducted and are available from the author upon request.
36Some may ask the question how entry costs influence firms’ behaviors. A quick answer to that is that the entry costs will change the location of the benchmark MPE in Figures 1 and 4. We do not need a dynamic game to examine the effect of entry costs on firms’ likelihood of entry. A static game as shown in Table 5 in the Appendix can illustrate the intuition just as well.
2.4.2.1 Structural Effect and Strategic Effect of the Scrap Value

As mentioned earlier, how firms behave in equilibrium in both cases above can be explained by the joint effect of the direct structural effect and the preemptive strategic effect of a lower scrap value. The direct structural effect makes firms behave more according to their own states instead of their rivals’ states, and thereby undermine the preemptive strategic effect. We can identify these effects in the various components of firms’ equilibrium conditions, and explain more rigorously how scrap values influence firms’ equilibrium strategies and preemptive behaviors. Again, let us use firm A as an example. Firm A’s equilibrium conditions by components are shown below:

\[ A^{-1}(P(s)) = I\{s_A = 0\}(-EC) + I\{s_A = 1\}(-SV) \]

\[ + P_B(s)\left(\pi(1,1) - \pi(0,1)\right) + (1 - P_B(s))\left(\pi(1,0) - \pi(0,0)\right) \]

\[ + \beta \left( P_B(s) \left( \Upsilon_x(1,1) - \tilde{\Upsilon}_x(0,1) \right) + (1 - P_B(s)) \left( \Upsilon_x(1,0) - \tilde{\Upsilon}_x(0,0) \right) \right) \]

\[ + \beta \left( P_B(s) \left( \Upsilon_{eve}(1,1) - \tilde{\Upsilon}_{eve}(0,1) \right) + (1 - P_B(s)) \left( \Upsilon_{eve}(1,0) - \tilde{\Upsilon}_{eve}(0,0) \right) \right) \]

\[ + \beta \left( P_B(s) \left( \Upsilon_{eve}(1,1) - \tilde{\Upsilon}_{eve}(0,1) \right) + (1 - P_B(s)) \left( \Upsilon_{eve}(1,0) - \tilde{\Upsilon}_{eve}(0,0) \right) \right) , \quad (2.4.1) \]

\[ \forall s \in S. \]

As can be seen, the value of SV enters directly into two components: \( I\{s_A = 1\}(-SV) \) and \( \Upsilon_{eve}(1,s_B) - \Upsilon_{eve}(0,s_B) \), \( \forall s_B \in \{0,1\} \). The effect of these two components on firm A’s CCPs \( P(s) \) is the direct structural effect. When \( s_A = 0 \), only \( \Upsilon_{eve}(1,s_B) - \Upsilon_{eve}(0,s_B) \) captures the

---

37 Again, for simplicity of notation, I remove subscript A from firm A’s payoff, value functions, and choice probabilities.
The preemptive strategic effect of lower scrap values is captured by firm A’s preemptive motives $\Upsilon (1, s_B) - \tilde{\Upsilon} (1, s_B), s_B \in \{0, 1\}$. As mentioned before, these terms are influenced by firm B’s CCPs $P_B (1, s_B)$ and $P_B (0, s_B)$, and represent an indirect effect of SV. This indirect effect manifests itself through the direct competitive effect that firm A has on firm B, which can be shown by the current period competition in firm B’s conditions, $P (s) (\pi_B (1, 1) - \pi_B (1, 0)) + (1 - P (s)) (\pi_B (0, 1) - \pi_B (0, 0))$. In particular, as SV decreases, knowing that firm A is more likely to stay at L2 after entry, firm B will become increasingly reluctant to enter (or stay at) L2 when $s_A = 1$; when $s_A = 0$, however, firm B knows that firm A is less likely to enter L2 as SV declines, and thereby becomes more and more eager to enter (or stay at) L2. This type of behavior of firm B implies that the gap between $P_B (1, s_B)$ and $P_B (0, s_B)$ grows with a lower SV, leading to an increase in firm A’s preemptive motives. Therefore, we see that firm A’s preemptive motives increase as its scrap value declines in Figure 2.

Both the direct structural effect and the preemptive strategic effect influence firms’ strategies as their scrap values decrease. How the strategies play out in equilibrium depend on which effect overpowers the other. It is important to note that although the analysis here focuses on the scrap

\[38\] Firm B’s conditions are analogous to firm A’s, only with the superscript switch between A and B.
values, the scrap values should be interpreted as sunk costs, or the sunk part of the entry costs. The reason for this is clear: if we change the entry costs at the same time, scrap values alone would not have the same effect on firms’ preemptive behaviors as shown in the case studies. For example, at the extreme, if we decrease entry costs at the same rate as the scrap values and maintain the equality of the two parameters, i.e. $EC_i = SV_i$, then any equilibrium correspondence we trace out would consist only of benchmark MPEs, which are non-preemptive. That is, there would be no preemption in equilibrium even though the scrap values have been decreasing. Therefore, the scrap values here should be interpreted as sunk costs.

Summary of Results

The analysis above shows that preemption depends on both the absolute levels of sunk costs and the ratio between competitors’ sunk costs, and that the source of a firm’s strategic commitment includes two factors: sunk costs and “excessive” preemptive gain in the case of multiple equilibria. Another important insight shown by Case 1 is that for a firm to preempt entry, its exit cost does not have to be very large. For example, as shown in Figure 3, MPE A is a very preemptive equilibrium. The exit cost for firm A at this point, however, is only $3$. The main reasons for this result are summarized below:

1. The model in this paper is a dynamic game with an infinite horizon, while Judd’s (1985) is a finite-horizon game. In a finite horizon game, the incumbent firm cannot commit to staying in the market in the last period because there is no gain in future profits through preemption. Therefore, the only subgame perfect equilibrium in Judd’s model is no preemption unless the incumbent has a large exit cost. In the model in this paper, firms interact with each other repeatedly for an infinite number of periods, and there is always a future gain in profits through preemption.

2. The incomplete information structure of the model in this paper also helps preemption to occur with a relatively low level of sunk cost. Judd’s (1985) model is a complete information game. With complete information, firms’ strategies (in terms of probabilities assigned to each action) are often either 1 or 0, and preemption is defined mostly based on the equilibrium outcome, which is also a binary variable. With incomplete information, preemption is a matter of degree, not a binary outcome. In particular, because firms’ strategies are probabilistic, they change much more sensitively to the increase in sunk costs compared to those in a complete information setting, where the strategies could stay constant at 1 or 0 for a very large range of the sunk costs. Due to this sensitivity, firms’ preemptive motives change also gradually with the increase in sunk costs. Therefore, we could see preemptive behaviors even for sunk costs.

39It should be noted here that Judd (1985) argues that sunk cost is irrelevant to the game of preemption. His definition of sunk cost is different from the definition of sunk cost in this paper. The sunk cost that Judd (1985) refers to is equivalent to the entry cost in my paper. By Judd’s (1985) argument, the level of entry cost should be also irrelevant for preemption. However, this is not the case in the context of my model. My model allows the activity of “reentry,” while Judd’s model does not. In my model, in the post-entry periods, firms’ repeated interactions involve both entry and exit activities, and therefore entry cost still matters for preemption.
that are not large. For example, in Case 1, all MPEs show preemptive behaviors from firm A, although in some MPEs, firm A’s preemptive behaviors are more pronounced than others.

It is also worth noting that when comparing firms’ behaviors in Case 1 with those in Case 2, there is a clear winner and a loser of preemption in Case 1, but not so in Case 2. In Case 1, without preemptive motives, firm A enters the market less aggressively and firm B more frequently. In Case 2, however, the pattern is not very clear. As shown by the third column of panels in Figure 4, for Case 2 at state (1, 0), both firms choose to be active less frequently in the counterfactual. This does not mean that both firms lose in Case 2, but rather implies that preemptive motives create a race between firms. Because both firms have large sunk costs and both their incumbency states have an entry deterrence effect, both firms have an incentive (though small) to preempt their rivals in the equilibrium. Once their preemptive motives are removed, however, both firms lose the incentive to go into the market earlier than their rivals, resulting in a decline in the entry probabilities of both firms. This phenomenon of a preemptive race will be discussed in greater detail in the following empirical application.

2.5 An Empirical Application: Preemptive Race in Canadian Burger Industry

This section applies the counterfactual and measure of preemption to analyze the preemptive race in the Canadian burger chain industry. The research question under examination is how the effects of preemption on firms’ behaviors vary across different sizes of markets. There are five Canadian burger chains: McDonald’s, A&W, Harvey’s, Burger King, and Wendy’s. The empirical results show that with preemptive motives, all firms in all market types try to preempt their rivals by entering a market early or exiting late. However, once preemptive motives are eliminated, all firms would delay entry and accelerate exit. On average, a market can go on unserved by any firm for up to 74 more years in the counterfactual. The effects of preemption vary non-monotonically across different market sizes, with medium-size markets being affected much more than very small or large markets. Given that there are five burger chains in this empirical analysis, this section also extends the counterfactual and measure of preemption to more than two players.

2.5.1 Data and Summary Statistics

The data of the Canadian burger industry comes from the dataset published by Igami and Yang (2016). The dataset contains the locations and timing of opening and closing of outlets for five large chains: McDonald’s, A&W, Harvey’s, Burger King, and Wendy’s. The dataset has a panel structure; it covers 400 markets in seven major Canadian metropolitan areas (Toronto, Montreal, Vancouver, Calgary, Edmonton, Winnipeg, and Ottawa) from 1970 to 2005. The data also include the population, average income and property value information for each market in each year.
In this empirical analysis, I focus on firms’ decisions of being active or inactive in a market. Table 2.2 summarizes firms’ active and inactive status across time-location markets as well as their entry and exit activities. As can be seen, McDonald’s was a dominant firm and active in over one third of the time-location markets. The number of markets where it is active is twice as many as the second largest firm A&W. McDonald’s had also been the first entrant in many markets. On average, its entry timing was about 2 to 5 years earlier than the other firms. McDonald’s also never exited a market, in contrast to the other firms, who all did at some point. These summary statistics indicate that McDonald’s is very different from the other firms.

Table 2.2: Firms’ Status and Entry and Exit Activities Across Time-Location Markets

<table>
<thead>
<tr>
<th></th>
<th>Number of Market-Year Observations</th>
<th>McDonald’s</th>
<th>A&amp;W</th>
<th>Burger King</th>
<th>Harvey</th>
<th>Wendy’s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td></td>
<td>4,787</td>
<td>2,344</td>
<td>1,156</td>
<td>1,848</td>
<td>1,003</td>
</tr>
<tr>
<td>Inactive</td>
<td></td>
<td>9,213</td>
<td>11,656</td>
<td>12,844</td>
<td>12,152</td>
<td>12,997</td>
</tr>
<tr>
<td><strong>Activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry</td>
<td></td>
<td>265</td>
<td>152</td>
<td>97</td>
<td>98</td>
<td>89</td>
</tr>
<tr>
<td>Exit</td>
<td></td>
<td>0</td>
<td>80</td>
<td>34</td>
<td>52</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>14,000</td>
<td>14,000</td>
<td>14,000</td>
<td>14,000</td>
<td>14,000</td>
</tr>
</tbody>
</table>

2.5.1.1 Preliminary Regression

To explore the competitive effects between firms, I ran a preliminary logit regression of firms’ actions against their own states and their rivals’ states as well as market condition variables and firms’ fixed effects. To capture heterogeneity across markets, I use a time-invariant market fixed effect. The regression results indicate that there are substantial competitive effects between firms; however, the effects have the right sign only when market heterogeneity is accounted for. The logit regression is specified as follows:

\[
a_{imt} = \begin{cases} 
1 (\text{active}) & \text{if } y^{*}_{imt} > 0 \\
0 (\text{inactive}) & \text{else}
\end{cases} \\
y^{*}_{imt} = \theta_0 + \theta_1 s_{imt} + \theta_2 \left( \sum_{j \neq i} s_{jmt} \right) + \theta_3 Pop_{mt} + \theta_4 Income_{mt} + \gamma_i + \mu_m + \nu_{imt}
\]

where \( y_{imt} \) is firm \( i \)'s action at time \( t \) in market \( m \); \( s_{imt} \) is firm \( i \)'s own active state and \( s_{jmt} \) the rival firm \( j \)'s active state. \( \left( \sum_{j \neq i} s_{jmt} \right) \) is therefore the total number of active rival firms. Here I assume that a firm’s decision to be active or not depends on the total number of competitors but not on the identity of the competitors. The relationship between action \( a_{imt} \) and the state is that \( a_{imt} = s_{imt+1} \); that is, the action taken today becomes the firm’s state tomorrow. \( Pop_{mt} \) is different from Igami and Yang (2016), which examines firms’ behaviors through the number of outlets they open in a market.
and \( Income_{mt} \) are the log population and income levels in market \( m \) at time \( t \) respectively.

\( \gamma_i \) is the firm fixed effect, \( \mu_m \) market fixed effect, and \( \nu_{int} \) is an i.i.d cost shock that follows a logit distribution. The parameter \( \theta \)'s are the coefficients.

The logit regression results are shown in Table 2.3. As can be seen, without market fixed effects, the competitive effect between a firm and its rivals is positive, but once market fixed effects are controlled for, the competitive effect is negative, a sign that is consistent with the theory of competition. The change in the sign of competitive effects can be explained by the fact that larger markets tend to attract more firms. At a cross-section level, without controlling for market heterogeneity, firms would appear to cluster. For a study of preemption, it is very important that the sign of the competitive effect of a firm's presence on its rivals' profits is negative because firms will not participate in preemption otherwise.

### Table 2.3: Preliminary Logit Regression Estimates

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own status ( \theta_1 )</td>
<td>8.438***</td>
<td>8.127***</td>
</tr>
<tr>
<td></td>
<td>(0.0876)</td>
<td>(0.0991)</td>
</tr>
<tr>
<td>Number of active rival firms ( \theta_2 )</td>
<td>0.247***</td>
<td>-0.905***</td>
</tr>
<tr>
<td></td>
<td>(0.0422)</td>
<td>(0.0652)</td>
</tr>
<tr>
<td>Population ( \theta_3 )</td>
<td>0.139*</td>
<td>3.672***</td>
</tr>
<tr>
<td></td>
<td>(0.0586)</td>
<td>(0.389)</td>
</tr>
<tr>
<td>Income ( \theta_4 )</td>
<td>0.441***</td>
<td>1.361**</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.426)</td>
</tr>
<tr>
<td>Burger King fixed effect ( \gamma_{bk} )</td>
<td>-0.374**</td>
<td>-0.445***</td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>Harvey fixed effect ( \gamma_{hvy} )</td>
<td>-0.282*</td>
<td>-0.384**</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.116)</td>
</tr>
<tr>
<td>Wendy's fixed effect ( \gamma_{wdy} )</td>
<td>-0.321**</td>
<td>-0.454***</td>
</tr>
<tr>
<td></td>
<td>(0.121)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>McDonald's fixed effect ( \gamma_{mcd} )</td>
<td>1.248***</td>
<td>1.264***</td>
</tr>
<tr>
<td></td>
<td>(0.0986)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>Constant ( \theta_0 )</td>
<td>-10.901***</td>
<td>-58.178***</td>
</tr>
<tr>
<td></td>
<td>(1.344)</td>
<td>(3.289)</td>
</tr>
<tr>
<td>Market Fixed Effects</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>70,000</td>
<td>69,475</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

\(* p < 0.05, ** p < 0.01, *** p < 0.001\)

Table 2.3 also indicates that McDonald's firm fixed effect at 1.248 is much larger than the other firms. Burger King's, Harvey's and Wendy's fixed effects are of a similar magnitude in the -0.28 to -0.38 range, slightly lower than that of A&W (normalized to 0). Given that McDonald's behaves very differently from the others and the other firms act relatively similarly, I treat McDonald's as

\[\text{Although the dataset includes property value information for each market, the estimated coefficient of property value is insignificant and thereby is excluded from the regression.}\]
its own unique firm and the other firms symmetrically in the structural model as described in the following section.

2.5.2 Model Specification

The dynamic model to capture firms’ strategic entry and exit behaviors is very similar to the one as described in Section 2.2, with only two slight changes: (1) there are five firms instead of two; (2) there is an added one-period time-to-build assumption to capture the lag time between firms’ decisions and the realization of those decisions. The one-period time-to-build assumption implies that if a firm decides to change its status at the beginning of a period, it incurs the cost of entry or exit in the current period, but its action does not take effect until the beginning of the next period. Extending the counterfactual and measure of preemption in Section 2.3 to include these two features is trivial, and it will be discussed in detail in Section 2.5.4.1.

Firms’ payoff functions can be specified as

\[
\Pi_{int}(a, s_{mt}, z_{mt}, \varepsilon_{int}|\theta) = \pi_{int}(s_{mt}, z_{mt}) - a(a - s_{imt})\theta_{ec} + \varepsilon_{int}(a),
\]

where \(i \in \{mcd, other\}\) represents the identity of a firm. \(a \in \{0, 1\}\) is the action of being active or inactive. \(s_{mt}\) is the market structure state, and \(s_{imt}\) is the active status of firm \(i\). \(z_{mt} \equiv \{z_{pop}, z_{inc}\}\) is a vector of exogenous market variables, which include population and income, and is assumed to transition according to a Markov process. \(\theta\) is the vector of the structural primitives of the model. Specifically, \(\theta_{ec}\) is the entry cost of opening a shop and is assumed to be the same for all firms.\(^{43}\) Given that fixed costs, entry costs and scrap values cannot be separately identified from each other (Aguirregabiria and Suzuki, 2014), firms’ scrap values here are normalized to 0. \(\varepsilon_{int}(a)\) is the idiosyncratic shock to firms’ fixed costs, entry costs or scrap values. It follows an extreme value type I distribution with location parameter 0 and scale parameter 1. Firms’ profits from static competition can be specified as below:

\[
\pi_{int}(s_{mt}, z_{mt}|\theta) = s_{imt} \left[ \theta_{mcd}D_{i \text{ mod}} + \theta_{m} + \theta_{c} \left( \sum_{j \neq i} s_{jmt} \right) + \theta_{pop}z_{pop} + \theta_{inc}z_{inc} \right], \tag{2.5.1}
\]

where \(D_{i \text{ mod}} \equiv I \{i = mcd\}\) is a McDonald’s firm dummy, and \(\theta_{mcd}\) is the McDonald’s firm fixed effect to capture its distinct behaviors from the other firms. \(\theta_{m}\) is the market fixed effect, and it

\(^{42}\)It usually takes time for firms to set up an outlet once having decided to open a shop or to liquidate the assets once having decided to close.

\(^{43}\)Alternatively, one could specify a different entry cost for McDonald’s. This specification would require including an interaction term between \(s_{imt}\) and a McDonald’s firm dummy in the preliminary logit regression as shown above. This is because the entry cost has an effect on a firm’s actions only through its state; a firm is reluctant to change its state unless its sunk cost is substantial. Therefore, including an interaction term between \(s_{imt}\) and McDonald’s dummy is necessary if we are to attribute McDonald’s unique behaviors to its entry costs. This alternative logit regression has been run, but unfortunately the data does not contain enough variation to identify the coefficient for the interaction term. As a result, I assume that all firms have the same entry costs and attribute McDonald’s unique behaviors only to its profitability in the market.
can be seen as the time-invariant component of the other firms’ profit functions in market \( m \). \( \theta_c \) is the competitive effect of the number of active rival firms on firm \( i \)'s profit. \( \theta_{\text{pop}} \) and \( \theta_{\text{inc}} \) capture how firms’ variable profits are related to the population and income levels in a market.

To satisfy Assumption 1 in Section 2.2.6, I discretize the population and income variables. I assume that they follow a bivariate AR(1) process, where the autoregressive parameters are homogeneous across markets but the means vary. Specifically, I demean the population and income for each market first, and then pool the demeaned variables from all markets and divide each variable into 3 quantiles. This discretization gives 9 states in total for the exogenous variable \( z_{mt} \). The discretized values for each variable are \( \{0, 1, 2\} \). The market-specific mean population and income are subsumed into the market fixed effect \( \theta_m \). To obtain the transition matrix of state \( z_{mt} \), I use the method in Tauchen (1986).

Given this model setup, firms solve the following Bellman equation by choosing an optimal entry and exit strategy \( \sigma \)

\[
V_{im}(s_m, z_m, \varepsilon_{im}, \sigma|\theta) = \max_{\sigma} \{ \Pi_i(s_m, z_m, \varepsilon_{im}, \sigma|\theta) + \beta E [V_i(s'_m, z'_m, \varepsilon'_{im}, \sigma|\theta) | s_m, z_m, \sigma] \}
\]

2.5.3 Identification and Estimation

The identification of the structural primitives comes from various features of the data. \( \theta_{\text{ec}} \) is identified from the difference in probability of being active between active firms and potential entrants. \( \theta_c \) is identified from the dependency of firms’ CCPs on the number of competitors in a market for a given market. \( \theta_m \) is identified from the variation of firms’ choice probabilities across geographic markets for a given market structure, population and income levels, and firms’ identity; the long panel feature of the data with 35 years of observations for each market is particularly helpful for the identification of \( \theta_m \). \( \theta_{\text{med}} \) is identified from McDonald’s entry and exit behaviors that appear much more aggressive than the other firms. \( \theta_{\text{pop}} \) and \( \theta_{\text{inc}} \) are identified from the variations in firms’ entry and exit probabilities across different population and income levels.

To estimate the structural primitives of the model, I use the two-step estimator as introduced by Hotz and Miller (1993) with the identifying assumption that firms’ behaviors come from a single equilibrium within each market. In the first step, I conduct a semi-nonparametric approximation of the CCPs for all the firms in each geographic market. A logit regression with market fixed effects, which is similar to the one in Section 2.5.1.1 but with discretized population and income variables and McDonald’s dummy, provides such an approximation. I then use the CCPs from the first

---

44 I experimented with other numbers of discretization and obtained the Bayesian Information Criterion for each case, and found that there was no loss in information by choosing the number 3.

45 The time subscript \( t \) is omitted here because we assume that all state variables follow a stationary Markov process. The state symbols with apostrophes represent the states in the next period.

46 Ideally, the reduced form estimates of CCPs in the first step would have to allow all the slope parameters (and not only the intercept) to be market specific, i.e., to run a separate logit regression model for each market. I have tried this specification; unfortunately, the data does not contain enough variation to identify market-specific parameters in order to give reasonable predictions for CCPs. As a result, I employ the simple but biased approach of using market fixed effects to account for market heterogeneity in the first stage, while fully aware of its caveats. More sophisticated techniques such as those introduced by Kasahara and Shimotsu (2009) and Arcidiacono and Miller (2011) can be
step to construct the continuation values associated with each independent variable and then plug them into firms’ choice specific value functions in the second-stage logit regression.

The structural parameter estimates from the second stage are shown in Table 2.4. As can be seen, the competitive effect parameter $\theta_c$ is estimated to be -0.148. The coefficients associated with population and income are 0.0349 and 0.177 respectively. The entry cost is -8.294 and McDonald’s firm fixed effect 0.129; all estimates are statistically significant at the 0.001 level. The market fixed effect $\theta_m$ is also estimated in the second stage, and their distribution is plotted in Figure 6. As shown, the distribution of the market fixed effect is slightly skewed, with most of the markets falling within the -0.25 to 0.5 range; the average market fixed effect is -0.0576.

To give the structural estimates in Table 2.4 a meaningful interpretation, I normalize the estimates by the average market fixed effect and average population and income levels. It is easy to see from equation (2.5.1) that mean $(\theta_m) + \theta_{pop} + \theta_{inc}$ represents the monopolistic profit for a non-McDonald’s firm in an average-sized market. Normalizing the structural estimates by this term will give the estimates meaningful economic interpretations: the normalized values will represent the marginal effect of each independent variable on a firm’s profit relative to a non-McDonald’s firm’s monopolistic profit in an average-sized market. Table 2.5 shows these normalized estimates. The ratio $\frac{\theta_c}{\text{mean}(\theta_m) + \theta_{pop} + \theta_{inc}}$ is -95.9%, meaning that the entry of one additional rival firm in an average-sized market will decrease a non-McDonald’s firm’s monopolistic profit by 95.9%; $\frac{\theta_{pop}}{\text{mean}(\theta_m) + \theta_{pop} + \theta_{inc}}$ implies that if population goes up to the largest level, a non-McDonald’s firm’s monopolistic profit in an average market will increase by 22.6%; similarly, $\frac{\theta_{inc}}{\text{mean}(\theta_m) + \theta_{pop} + \theta_{inc}}$ indicates that the monopolistic profit will grow by 114.7% if income rises to the highest level. The ratio $\frac{\theta_{mc}}{\text{mean}(\theta_m) + \theta_{pop} + \theta_{inc}}$ suggests that the entry cost is equivalent to 53.75 years of a non-McDonald’s firm’s monopolistic profit in an average market, and $\frac{\theta_{mcd}}{\text{mean}(\theta_m) + \theta_{pop} + \theta_{inc}}$ implies that McDonald’s monopolistic profit is 83.6% more than those of the other firms in an average market. All estimates seem to be within a reasonable range based on these interpretations, except that the entry cost on the high end: it is somewhat implausible that an outlet of a non-McDonald’s firm in an average market would take 53.75 years to recover its initial setup cost. This large entry cost estimate may be due to persistent market shocks that are not accounted for in this model, and their effects are captured by the entry cost parameter.

The market fixed effect $\theta_m$ can be seen as representing the time-invariant market size. As mentioned earlier, it represents the time-invariant component of the other firms’ profit functions. This time-invariant component can come from both the demand side and the cost side. If we assume that firms’ fixed costs are constant across markets, then $\theta_m$ captures the effect of the unobserved time-invariant market demand on a firm’s profit. The same interpretation also applies for McDonald’s. For this reason, in the following analysis, I use $\theta_m$ to represent market size when examining the relationship between preemption and market size.
Table 2.4: Structural Parameter Estimates

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of active rival firms $\theta_c$</td>
<td>-0.148***</td>
</tr>
<tr>
<td></td>
<td>(0.00149)</td>
</tr>
<tr>
<td>Population $\theta_{pop}$</td>
<td>0.0349***</td>
</tr>
<tr>
<td></td>
<td>(0.00155)</td>
</tr>
<tr>
<td>Income $\theta_{inc}$</td>
<td>0.177***</td>
</tr>
<tr>
<td></td>
<td>(0.00177)</td>
</tr>
<tr>
<td>Entry cost $\theta_{ec}$</td>
<td>-8.294***</td>
</tr>
<tr>
<td></td>
<td>(0.00977)</td>
</tr>
<tr>
<td>McDonald’s firm fixed effect $\theta_{mcd}$</td>
<td>0.129***</td>
</tr>
<tr>
<td></td>
<td>(0.00525)</td>
</tr>
<tr>
<td>Market Fixed Effect</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>69,475</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 6: Distribution of Market Fixed Effects $\theta_m$
Table 2.5: Normalized Structural Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of active rival firms ( L_{m} ) ( \theta_{m} )</td>
<td>-95.9%</td>
</tr>
<tr>
<td>Population ( L_{m} \theta_{pop} ) ( \theta_{inc} )</td>
<td>22.6%</td>
</tr>
<tr>
<td>Income ( L_{m} \theta_{pop} ) ( \theta_{inc} )</td>
<td>114.7%</td>
</tr>
<tr>
<td>Entry cost ( L_{m} \theta_{inc} )</td>
<td>-33.75</td>
</tr>
<tr>
<td>McDonald’s firm FE ( L_{m} \theta_{mcd} )</td>
<td>83.6%</td>
</tr>
</tbody>
</table>

2.5.4 Effect of Preemption by Market Size

To examine how preemption varies across market size, I first construct the counterfactual scenarios of how firms would behave without preemptive motives and then compare firms’ counterfactual CCPs with those from the equilibrium as observed in the data. To simplify the counterfactual analysis, I assume the markets are time invariant at their highest population and income levels. Although the counterfactual can be constructed for time-varying markets as explained in Section [2.3.3], I make this simplification because it is innocuous for this particular application. The transition matrix of the population and income variables suggests that if a market starts at the state with the highest population and income levels, it is likely to stay there. In particular, in the steady-state distribution of markets, about 96% of the markets have the highest levels of population and income. Therefore, at the highest population and income state, a market can be seen as time invariant, and the market fixed effect \( \theta_{m} \) will capture the variation in market size related to demand.

With this simplification, I solve first for firms’ equilibrium CCPs in the dynamic game with time-invariant markets, and then construct the counterfactual to obtain firms’ counterfactual CCPs. The comparison between these two sets of CCPs shows that firms’ counterfactual CCPs of choosing to be active are lower than their equilibrium ones at every single state for all geographic markets. This means that without preemptive motives, firms delay entering a market but speed up exit. Section 2.5.4.1 below specifies the counterfactual for this model with five players and the one-period time-to-build assumption. Section 2.5.4.2 illustrates how the effects of preemption vary across market size.

2.5.4.1 Extension of Measure to Multiple Firms

The counterfactual in this model can be constructed by solving the following conditions for firm \( i \):

\[
\ln \left( \frac{P_{i}(s_{m})}{1 - P_{i}(s_{m})} \right) = (1 - s_{im}) \theta_{ic} + \beta \left[ \sum_{s'_{-im}} \left( \Upsilon_{ic} \left(s'_{im} = 1, s'_{-im}, \theta_{m} \right) - \Upsilon_{ic} \left(s'_{im} = 0, s'_{-im}, \theta_{m} \right) \right) f_{im} \left(s'_{-im} \mid s_{m} \right) \right] \\
+ \beta \left[ \sum_{s'_{-im}} \left( \Upsilon_{ieve} \left(s'_{im} = 1, s'_{-im}, \theta_{m} \right) - \Upsilon_{ieve} \left(s'_{im} = 0, s'_{-im}, \theta_{m} \right) \right) f_{im} \left(s'_{-im} \mid s_{m} \right) \right], \quad (2.5.2)
\]

\( \forall s_{m} \in S_{m} \)

Note that the log form on the left hand side of the equation is due to the assumption of extreme value type I errors.
where \( s_m = (s_{im}, s_{-im}) \) is the market structure state, \( s_{im} \) is the state of firm \( i \) and \( s_{-im} \) is a vector of the active status of all firms other than \( i \), \( S_m \) is the set of all market structure states, \( s'_{-im} \) denotes the rivals’ states tomorrow, and \( f_{im} (s'_{-im}|s_m) \) is the transition probability of the rivals’ states to \( s'_{-im} \) from today’s state \( s_m \). Given that firms simultaneously decide to enter or exit, the rivals’ transition probability \( f_{im} (s'_{-im}|s_m) \) is independent of firm \( i \)’s actions. \( \hat{\Upsilon}_{1\pi} (\cdot) \) and \( \Upsilon_{1\pi} (\cdot) \) are similarly defined as in Section 2.3.2. In particular, \( \hat{\Upsilon}_{1\pi} (\cdot) \) is constructed as follows:\(^{48}\)

\[
\hat{\Upsilon}_{1\pi} = (I - \beta \hat{F}_m)^{-1} \pi_{im}, \tag{2.5.3}
\]

where \( \pi_{im} \) is a vector that stacks firm \( i \)’s flow profit in market \( m \) at various states and \( \hat{F}_m \) is a “modified” transition matrix, in which all of the rival firms’ CCPs \( P_{-i} (s_{im} = 1, s_{-im}) \) is replaced with \( P_{-i} (s_{im} = 0, s_{-im}) \).\(^{49}\) Firms’ CCPs under the counterfactual of no preemptive motives can thus be obtained by solving the system of equations in (2.5.2) jointly for all firms.

### 2.5.4.2 Effect of Preemption by Market Size

To compare firms’ counterfactual and equilibrium CCPs, I present the two preemption indices as described in Section 2.3.2. Figure 7 plots the probability-based preemption index for each firm against the market size (\( \theta_m \)). The blue line in the figure indicates the preemption index for McDonald’s and the red for the other firms. As can be seen, in all markets, the preemption index is negative for all firms, which implies that all firms enter or stay in the market with a lower probability without preemptive motives. In this case, preemption does not create a clear winner or loser, but results in a race: in order to reduce the probability of the rivals’ being active in a market, a firm will enter (or stay in) a market aggressively.

A distinct pattern in Figure 7 is that the effect of preemption is not monotonic with respect to market size. In particular, in very large and small markets, firms are affected less by preemptive motives than in medium sized markets. Furthermore, the effects of preemption peak at different market sizes for different firms. For McDonald’s, the effect is the highest at around -0.25. For the other firms, the peak is near 0.16.

An interesting feature of Figure 7 is that the overall magnitude of the effect of preemption on firms’ choice probabilities seems to be higher for other firms than for McDonald’s. This phenomenon

---

48 The algebraic expressions for \( \Upsilon_{1\pi} (\cdot) \) are almost identical to those in equation (2.3.5), only that the transition matrix \( F^P \) is a 32 by 32 matrix in this case.

49 At first glance, it may appear that the counterfactual condition (2.5.2) is missing the current-period competition compared to the one in equation (2.3.13), which is constructed without the one-period time-to-build assumption. Recall that the second line in equation (2.3.13) represents the current-period competition and the direct competitive effect. The condition (2.5.2) does not include this line; however, it should be noted that the current-period competition is still accounted for in the \( (2.3.13) \). In particular, it is embedded in the first line of \( \beta \sum_{s'_{-im}} \left( \hat{\Upsilon}_{1\pi} (s'_{im} = 1, s_{-im}) - \Upsilon_{1\pi} (s_{im} = 0, s_{-im}) \right) \beta f_{im} (s_{-im}|s_m) \), in which the \( \Upsilon_{1\pi} \) term, \( \hat{\Upsilon}_{1\pi} = \pi_{im} + \beta \hat{F}_m \Upsilon_{1\pi} \), includes a deterministic profit \( \pi_{im} \) that is not affected by the replacement of CCPs in \( \hat{F}_m \). If we take out this deterministic \( \pi_{im} \) from the \( \Upsilon_{1\pi} \) term, we can create a line of expression that is very similar to the second line in (2.3.13). That is, the current period competition and direct competitive effect are accounted for in the above counterfactual condition (2.3.2).
seems to contradict the pattern shown in the data where McDonald’s exhibits clearly more aggressive entry behaviors than other firms (see Table 2.2). This anomaly is due to the nonlinear relationship between the logit choice probabilities and the difference in firms’ choice-specific values. For CCPs greater than 0.5, the larger the CCPs are, the greater preemptive motives need to be in order to generate the same marginal increase in the CCPs. Because McDonald’s has much higher CCPs for staying in the market after entry in the equilibrium, even though its preemptive motives can be large, its CCPs do not decrease as much at the margin in the counterfactual compared to the other firms. For this reason, McDonald’s appears to be influenced by preemptive motives less than the other firms under this index. This oddity is corrected once we use the duration-based preemption index for firms, which accounts for the nonlinear relationship between CCPs and preemptive motives.

Figure 8 plots the duration-based preemption index against the market size for each firm. As shown, in terms of durations, preemption affects McDonald’s much more than the other firms. The effect for McDonald’s (the blue line) is higher than that for the other firms (the red line) for almost all market sizes greater than -0.25. In addition, the figure also suggests a non-monotonic relationship between the effect of preemption and market size, although the pattern is not as pronounced as that shown by the probability-based index. In this regard, both indices are good at presenting some aspect of the results.

As mentioned earlier, the duration-based preemption index can be interpreted as the average years of additional absence from a market for a firm. Based on this interpretation, McDonald’s can stay away from a market for up to 85 years, and the other firms for up to 56 years. At first glance, these numbers seem large. However, this order of magnitude is consistent with the data.

As a robustness check, I plot firms’ CCPs in terms of durations at various estimation stages: (1) the empirical durations from the first-stage semi-nonparametric estimation, (2) the structural durations from the second-stage structural estimation, and (3) the time-invariant large market durations based on firms’ equilibrium CCPs for time-invariant markets. The durations here can be interpreted as the average wait time before each firm enters or exits a market. To be more precise,
I plot the entry durations and exit durations for each firm separately. Again these durations are weighted by the steady-state distributions. The relationship between these durations and market size is shown in Figure 9. The blue lines represent the empirical durations, the red the structural ones and the green for the time-invariant large market. The top two panels are the durations for McDonald’s and the bottom two are for the other firms.

As shown in Figure 9, all firms’ durations are in the order of hundreds. In particular, the empirical duration of McDonald’s could go up to 1,000 years. This reflects the pattern observed in the data that McDonald’s never exits. Although McDonald’s exit duration for time-invariant large market (the green line) is substantially lower than the empirical one for larger markets, the magnitude of this duration is still in the range of 400-600 years. Therefore, for the effect of preemption in duration, a maximum of 85 years for McDonald’s and 56 years for other firms are consistent with the data.

The duration index above is for individual firms in isolation of the behaviors of their rivals. To obtain how long on average a market will be served by no firm at all in the counterfactual compared to the equilibrium, I plot the duration-based preemption index for all firms in Figure 10. As can be seen, without preemptive motives, some markets can go unserved by any firm for an additional 74 years. Again, the relationship between the effect of preemption and the market size is non-monotonic.

To conclude, the results in this section show that the effects of preemption discourage all firms from entering or staying in a market. The magnitude of the effects is non-monotonic across the market size. In particular, medium sized markets are affected much more than small or large markets. This non-monotonic pattern is very reasonable because in both small markets or large markets.

---

As shown in Figure 9, the green and red lines are fairly similar. This is because every market has a high probability of evolving into a state with the highest population and income levels. As a result, the dynamic game with time-varying markets and the game with time-invariant large market size yield somewhat similar CCPs. This confirms that assuming time-invariant markets in the counterfactual analysis is innocuous.
markets, firms’ behaviors are less influenced by their rivals’ incumbency status and preemption has a smaller effect in those cases. In small markets, firms do not want to enter a market anyway no matter whether their opponents are active in the market or not, in which case, a firm has a smaller incentive to preempt its rivals. In large markets, all firms want to enter the market regardless of their rivals’ status because there is enough demand for all firms. In that case, preemption is less effective. Therefore, the effect of preemption is most pronounced in medium sized markets.

2.6 Comparison to Other Definitions in the Literature

In this section, I compare the definition of preemptive motives in my paper to other definitions proposed in the literature. Since the definitions of preemption imply the counterfactuals, I focus on discussing the counterfactuals here instead of the definitions themselves. As mentioned before,
a number of papers have studied spatial preemption, including Igami (2015), Igami and Yang (2016), Zheng (2016), Yang (2015), and Hünermund et al. (2014). All the counterfactuals in these papers change other aspects of the equilibrium outcome while eliminating or reducing the effect of preemption. Compared to these counterfactuals, the counterfactual in this paper has the advantage of preserving both the direct competitive effect and the dynamic features of the game while removing the effect of preemption. Now let us compare these counterfactuals in detail.

Igami (2015) uses a finite-horizon model with sequential moves in each period to study preemption in the hard disk drive industry. The “preemptor” firms are the incumbents and are modeled as the first-movers. The “preempted” firms are potential entrants, who move after observing the incumbents’ actions within each period. The counterfactual in Igami (2015) makes the potential entrants ignore the entry of the incumbents at the new product location. Applying that principle to the simplified model in this paper is equivalent to setting $P_A(s), \forall s \in S$ to 0 if we treat firm A as the preemptor. Then firm B’s problem becomes the problem of a single agent who faces no competition, as implied by its equilibrium conditions:

$$
A^{-1}(P_B(s)) = 1(s_B = 0)(-EC_B) + 1(s_B = 1)(-SV_B) + (\pi_B(0, 1) - \pi_B(0, 0)) + \beta(\nabla_B(0, 1) - \nabla_B(0, 0)) \quad (2.6.1)
$$

$s \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}$,

where $\nabla_B(0, 1)$ and $\nabla_B(0, 0)$ are calculated based on (2.2.5), but with $P_A(s) = 0, \forall s \in S$ for all current and future periods.

Note that in Igami’s (2015) model, the potential entrants in are short-lived: they either become incumbents after entry or disappear permanently if choose not to enter. His model is thus very different from the one in this paper, where both firms are long-lived. However, the principle for removing preemption used in Igami (2015) can still be compared to the one in this paper.
As can be seen, the counterfactual in Igami (2015) removes competition faced by the potential entrant completely, and thereby eliminates not only the preemptive effects, but also the direct competitive effects. This is not the case for the counterfactual in this paper. As discussed earlier, the direct competitive effects are retained especially by the current period competition as represented by the second line of (2.3.13).

The same principle from Igami (2015) is also used in Igami and Yang (2016) to remove preemption. In particular, the authors eliminate the competitive effects of McDonald’s on other chains. The model in Igami and Yang (2016) is very similar to the model in this paper, only with more than two firms and the option of more stores. The implication of the counterfactual in that paper is analogous to what has been discussed above. In particular, any competitive effect that McDonald’s has on other chains is eliminated, including both the direct competitive effects and the preemptive effects.

Two additional issues are worth noting about the counterfactual in Igami and Yang (2016). First, there is no reason that McDonald’s is the only preemptor in every type of market. As shown in the numerical analysis in Section [2.4], the identity of the preemptor depends on the economic fundamentals of the market. Although Yang (2015) provides empirical evidence that suggests that McDonald’s has been implementing preemptive strategies, the test of preemption proposed in that paper is based on a three-stage game, which is fundamentally different from the dynamic game framework used in Igami and Yang (2016). The second issue is that although the counterfactual in Igami and Yang (2016) removes the competitive effect of McDonald’s on other chains, there is no guarantee that other chains are not playing preemptive strategies against each other. In other words, the preemptive effects are not examined in full in Igami and Yang (2016).

Similar to Igami (2015), Zheng (2016) also uses a dynamic game with sequential moves to model preemption. The firms in Zheng (2016) are two large retail chains, Blue and Red, competing over locations with spatially differentiated products. Within each period, the Blue firm moves first and the Red second. Again this sequential-move setup gives the pre-selected preemptor, the Blue firm, the first mover advantage, so that in equilibrium the Blue firm would indeed become the preemptor. The counterfactual in this paper uses a one-period deviation standard: it takes the location that the Blue firm chooses out of the Red firm’s choice set for one period. The idea behind this counterfactual is that if the Blue firm knows that the Red is not entering the location in that period, would the Blue firm have delayed its decision to enter, given that it faces a cannibalization effect by expanding spatially? Applying the principle of Zheng’s counterfactual to the simplified model in this paper is equivalent to setting \( P_B(s) = 0 \) in the current period in firm A’s equilibrium conditions:

\[
A^{-1}(P_A(s)) = I\{s_A = 0\}(-EC_A) + I\{s_A = 1\}(-SV_A) + (\pi_A(1,0) - \pi_A(0,0)) + \beta(V_A(1,0) - V_A(0,0))
\]

\[
s \in \{(0,0),(0,1),(1,0),(1,1)\},
\]

where \( V_A(0,1) \) and \( V_A(0,0) \) are calculated based on equation (2.2.5), with firms CCPs \( P_i(s) \), \( \forall s \in S, i \in \{A,B\} \) unchanged for future periods.

As can be seen, this counterfactual removes the direct competitive effect faced by firm A in
the current period, but does not necessarily capture the preemptive motives in future periods. Nevertheless, this counterfactual could dampen aggressive entry behaviors of firm A; for example, if the effect of cannibalization \((\pi_A (1,0) - \pi_A (0,0))\) is very strong, then firm A would choose to enter (or stay) less frequently than it would in the equilibrium scenario. Zheng (2016) notes that this counterfactual represents the lower bound of preemptive incentives. Compared to Zheng’s (2016) counterfactual, the one in this paper captures firms’ preemptive motives to the full extent.

Hünermund et al. (2014) models the process of an industry shakeout in a dynamic oligopoly game with many firms that make simultaneous entry and exit decisions. The authors use an open-loop equilibrium as the counterfactual. In the open-loop equilibrium, firms are forced to behave only according to their own states regardless of their rivals’ status in the market. They choose their strategies in terms of CCPs at the beginning of time to maximize their NPVs of payoffs at the initial state and commit to these strategies in all future periods. Applying their open-loop equilibrium concept to the simplified model in this paper implies \(P_A (s_A, 1) = P_A (s_A, 0), \forall s_A \in \{0, 1\}\) and \(P_B (1, s_B) = P_B (0, s_B), \forall s_B \in \{0, 1\}\). This pattern of behavior of firms is very different from what is seen in the counterfactual in this paper. As shown by both numerical examples, Case 1 and Case 2, in Section 2.4, firms do react to their rivals’ incumbency status through direct competitive effects. In this regard, the open-loop equilibrium in Hünermund et al. (2014) undermines the direct effect of competition.

Overall, compared to the counterfactuals proposed in the literature, the counterfactual in this paper has the advantage of eliminating only the effect of preemption while preserving both the direct competitive effect and the dynamic features of the game.

2.7 Conclusion

This paper proposes an approach to measure preemptive entry in a dynamic oligopoly game. I first define preemptive motives by decomposing firms’ equilibrium conditions and isolating the preemptive gain from all other entry motives of firms. By forcing firms to ignore their preemptive motives and re-optimize their values, I then obtain a counterfactual that eliminates the effect of preemption without changing the structural parameters of the game. Compared to other approaches in the literature, this method has the advantage of preserving the direct competitive effect and the dynamic features of the game while removing the indirect strategic effect of competition.

I implement this method in a numerical experiment and in an empirical application. In the numerical exercise, I study the influence of entry and exit costs on preemption and find that firms can successfully preempt entry without large exit costs, a result contrary to the conventional wisdom. In the empirical application, I examine how the effects of preemption vary across market size in the Canadian burger industry and find that the effects are much stronger in medium-sized markets than small or large markets.

\(^{52}\)This is so particularly when \((\pi_A (1,0) - \pi_A (0,0)) < (\pi_A (1, 1) - \pi_A (0, 1))\).

\(^{53}\)For a detailed definition of open-loop equilibrium, see Fudenberg and Tirole (1991)
I further extend the counterfactual and measure of preemption to dynamic oligopoly games with a greater number of competitors and in spatial markets that include multiple product locations. Through these extensions, I show that the counterfactual and measure of preemption developed in this paper is fairly flexible and robust to different model specifications.
Bibliography


Hollenbeck, B. 2015. The Spread of horizontal chains: Efficiency or information? Working paper, Anderson School of Management, UCLA.


APPENDIX

.1 Homotopy Method

The homotopy method has been used to explore equilibrium manifolds in much of the computational theory literature; for example, Besanko et al. (2010), Besanko et al. (2014) and Borkovsky et al. (2012). Borkovsky et al. (2010) provides a detailed explanation of the theory of the homotopy method and a step by step manual on how to implement the homotopy method by using the HOMPACK90 software. Without reinventing the wheel, I hereby sketch the main ideas of the homotopy method.

Essentially the method solves a complex system of equations by using a known solution to a simpler problem. For example, if we want to find a MPE for any given set of parameters in the model, we can first set firms’ scrap values to their entry costs, and then obtain a benchmark MPE by solving the one-shot game as shown in Table 6. Once we have the benchmark MPE (denoted by $P^f$) in hand, we can define a system of equations as follows:

$$
\begin{align*}
H(P, \tau) &= P - \Psi(P; [(1 - \tau) ec + sv]) = 0, \\
P(\tau = 0) &= P^f, 0 \leq \tau \leq 1,
\end{align*}
$$

where $[(1 - \tau) ec + sv]$ represents the parameter values for scrap value in $\Psi(\cdot)$. It is a linear combination of the entry cost vector $ec$ and scrap value vector $sv$ in the model of interest. When $\tau = 0$, the parameter values for scrap value in $\Psi(\cdot)$ are set to $ec$, and we know that the benchmark MPE is a solution to the system of equations in (1.1). When $\tau = 1$, the parameter values of scrap value reflect the real scrap values $sv$, and the equations in (1.1) become the equilibrium conditions for our real model of interest. Our aim is to find a solution to this system of equations for $\tau = 1$.

To do so, we can start from the know solution $P^f$ for $\tau = 0$, and trace out a solution path for $H(P, \tau) = 0$ by increasing $\tau$ gradually from 0 to 1. The ending point of the path is then an equilibrium solution to the real model. Because the system of equations in (1.1) is very complex, we cannot obtain an analytical solution for $P$ as a function of $\tau$. To overcome this difficulty, we can trace out the solution path for $H(P, \tau) = 0$ numerically by first rewriting $(P, \tau)$ as a function of a

---

54 Recall that $\Psi(\cdot)$ is the mapping as represented by (2.2.3).

55 Specifically, $ec = \begin{pmatrix} EC_A \\ EC_B \end{pmatrix}$, $sv = \begin{pmatrix} SV_A \\ SV_B \end{pmatrix}$. 
path parameter \( p \), i.e. \( (P(p), \tau(p)) \). The path parameter \( p \) can be seen as the distance travelled on the path from the starting point \( P(p = 0) = P^f \). The solution path is then an equilibrium correspondence

\[
H^{-1} = \{(P(p), \tau(p)) : H(P(p), \tau(p)) = 0\},
\]

which can be obtained by numerically solving the following system of differential equations:

\[
\frac{\partial H(P(p), \tau(p))}{\partial p} p'(p) + \frac{\partial H(P(p), \tau(p))}{\partial \tau(p)} \tau'(p) = 0
\]

Varying \( p \) from 0 till \( \tau(p) = 1 \) gives us a MPE in the real model\(^{56}\).

### 2. Some Counterfactuals Are MPEs when \( EC = SV \)

In general, the counterfactuals of the game are not MPEs. However, when firms’ entry costs are equal to their scrap values, i.e. \( EC_i = SV_i, \forall i \in \{A, B\} \), some counterfactuals coincide with MPEs. The reason for this is that when \( EC_i = SV_i \), firm \( i \) is not hesitant at exiting the market after entry because upon exit, firm \( i \) can retrieve all its initial investment in \( EC_i \) from the scrap values \( SV_i \). Knowing that, its competitor firm \( -i \) is not afraid of entering the market at firm \( i \)’s incumbency state. This implies that firm \( -i \)’s CCPs would be the same regardless of firm \( i \)’s incumbency status. Since the only difference between a counterfactual and an equilibrium is whether or not we replace \( P_B(1, s_B) \) with \( P_B(0, s_B) \) and \( P_A(s_A, 1) \) with \( P_A(s_A, 0) \), if some equilibria already have the characteristics of \( P_B(1, s_B) = P_B(0, s_B) \), \( \forall s_B \in \{0, 1\} \) and \( P_A(s_A, 1) = P_A(s_A, 0), \forall s_A \in \{0, 1\} \), then they must be also counterfactuals.

I formally state the existence of such MPEs in Proposition 1 below:

**Proposition 1.** When \( EC_i = SV_i, \forall i \in \{A, B\} \), there exist MPEs that satisfy

\[
P_A(s) = P_A(s'), P_B(s) = P_B(s'), \forall s, s' \in S
\]

**Proof.** See Subsection 2.1 below.

The intuition of the proof is that when \( EC_i = SV_i, \forall i \in \{A, B\} \), the term \( 1(s_i = 0)(-EC_i) + 1(s_i = 1)(-SV_i) \) in the equilibrium conditions (2.3.2) for firm A and a similar one for firm B becomes \(-EC_i \), and the equilibrium condition at every state has exactly the same algebraic form. In other words, at every state, firms face exactly the same tradeoffs. As a result, there exists a MPE where firms act completely the same at every state, i.e. \( P_A(s) = P_A(s'), P_B(s) = P_B(s'), \forall s, s' \in S \). I call these MPEs the benchmark MPEs.

For reasons discussed earlier, the benchmark MPEs are also counterfactuals, and therefore non-preemptive by definition. Furthermore, these benchmark MPEs have the desirable property that firms behave as if they are playing a static simultaneous-move game. In particular, firms’ CCPs in

---

\(^{56}\)For a form of solution to (1.4) and an algorithm of numerically tracing the path, please see Zangwill and Garcia (1981) (Chapters 2, 15 and 16).
the benchmark MPEs can be obtained from solving a one-shot game. This property is summarized in Proposition 2 below:

**Proposition 2.** Let \( \{ P_i (s) = \bar{P}_i, \forall i \in \{ A, B \} , s \in S \} \) be firms’ CCPs of a benchmark MPE. Then \( \{ \bar{P}_i, \forall i \in \{ A, B \} \} \) is the equilibrium CCPs of the following one-shot game:

where \( \varepsilon_i (a) \) is i.i.d and follows an extreme value type I distribution.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0</td>
<td>( \pi_A (1, 1) - (1 - \beta) EC_A + \varepsilon_A (1), \pi_B (1, 1) - (1 - \beta) EC_B + \varepsilon_B (1) )</td>
<td>( \pi_A (1, 1) - (1 - \beta) EC_A + \varepsilon_A (1), \pi_B (1, 1) - (1 - \beta) EC_B + \varepsilon_B (1) )</td>
</tr>
<tr>
<td></td>
<td>( \pi_A (0, 1) + \varepsilon_A (0), \pi_B (0, 1) + (1 - \beta) EC_B + \varepsilon_B (1) )</td>
<td>( \pi_A (0, 0) + \varepsilon_A (0), \pi_B (0, 1) + (1 - \beta) EC_B + \varepsilon_B (1) )</td>
</tr>
</tbody>
</table>

Furthermore, let \( \{ \bar{P}_A, \bar{P}_B \} \) be the equilibrium CCPs of the above one-shot game, then \( \{ \bar{P}_i (s) = \bar{P}_i, \forall i \in \{ A, B \} , s \in S \} \) are the CCPs of a benchmark MPE.

**Proof.** See Subsection 2.2 below.

The intuition for the proof is that because firms behave the same regardless of the state, the Markov state as a payoff relevant history no longer matters. The game is then static in nature. In particular, because firms freely enter and exit the market, their real cost of entry is \((1 - \beta) EC_i\); they pay \( EC_i \) upon entry and receive a scrap value \( SV_i = EC_i \) when exit in the next period. The static game in Table 6 takes this real cost of entry into consideration, and subtracts it from firms’ flow profits for the actions that involve entry.

**2.1 Proof of Proposition 1**

**Proof.** When \( EC_i = SV_i, \forall i \in \{ A, B \} \), the term \((s_1 = 0) (-EC_i) + 1 (s_1 = 1) (-SV_i)\) becomes \(-EC_i\) in the equilibrium conditions for firm A and firm B, and these two sets of equations can be written as follows:

\[
A^{-1} (P_A (s)) = -EC_A + P_B (s) (\pi_A (1, 1) - \pi_A (0, 1)) + (1 - P_B (s)) (\pi_A (1, 0) - \pi_A (0, 0)) + \beta (P_B (s) (V_A (1, 1) - V_A (0, 1)) + (1 - P_B (s)) (V_A (1, 0) - V_A (0, 0))) 
\]

\[
A^{-1} (P_B (s)) = -EC_B + P_A (s) (\pi_B (1, 1) - \pi_B (0, 1)) + (1 - P_A (s)) (\pi_B (1, 0) - \pi_B (0, 0)) + \beta (P_A (s) (V_B (1, 1) - V_B (0, 1)) + (1 - P_A (s)) (V_B (1, 0) - V_B (0, 0))) 
\]

It is easy to see from the above equations that at any state \( s \), the tradeoff that firm \( i \) faces is exactly the same. As long as their opponent firm \( i \)'s CCPs are the same across all states, i.e. \( P_{-i} (s) = P_{-i} (s'), \forall s, s' \in S \), it is necessary that firm \( i \) will behave the same at every state, \( P_i (s) = P_i (s'), \forall s, s' \in S \).

Now we need to show that CCPs that satisfy \( P_i (s) = P_i (s'), \forall i \in \{ A, B \} , s, s' \in S \) exist as a solution to the system of equations in (2.1) and (2.2). To establish existence, I use a constructive
proof. Suppose that the solution $P_i(s) = P_i(s')$, $\forall i \in \{A, B\}$, $s, s' \in S$ does exist, then the state $s$ no longer matters. The system of equations in (2.1) and (2.2) can be collapse into two equations:

$$
A^{-1}(\tilde{P}_A) = -EC_A + \tilde{P}_B(\pi_A(1,1) - \pi_A(0,1)) + (1 - \tilde{P}_B)(\pi_A(1,0) - \pi_A(0,0)) + \beta\left(\tilde{P}_B(V_A(1,1) - V_A(0,1)) + (1 - \tilde{P}_B)(V_A(1,0) - V_A(0,0))\right) \tag{2.3}
$$

$$
A^{-1}(\tilde{P}_B) = -EC_B + \tilde{P}_A(\pi_B(1,1) - \pi_B(0,1)) + (1 - \tilde{P}_A)(\pi_B(0,1) - \pi_B(0,0)) + \beta\left(\tilde{P}_A(V_B(1,1) - V_B(0,1)) + (1 - \tilde{P}_A)(V_B(0,1) - V_B(0,0))\right) \tag{2.4}
$$

If a solution $(\tilde{P}_A, \tilde{P}_B)$ exists for this system of equations in (2.3) and (2.4), then the CCPs $P_i(s) = \tilde{P}_i, \forall i \in \{A, B\}, s \in S$ must be a solution to (2.1) and (2.2). So the key to the proof is to show that a solution $(\tilde{P}_A, \tilde{P}_B)$ exists for (2.3) and (2.4). Note that (2.3) and (2.4) can be transformed into a fixed-point mapping:

$$
\tilde{P}_A = A\left(u_A(\tilde{P}_B)\right) \tag{2.5}
$$

$$
\tilde{P}_B = A\left(u_B(\tilde{P}_A)\right), \tag{2.6}
$$

where

$$
u_A(\tilde{P}_B) = -EC_A + \tilde{P}_B(\pi_A(1,1) - \pi_A(0,1)) + (1 - \tilde{P}_B)(\pi_A(1,0) - \pi_A(0,0)) + \beta\left(\tilde{P}_B(V_A(1,1) - V_A(0,1)) + (1 - \tilde{P}_B)(V_A(1,0) - V_A(0,0))\right)
$$

$$
u_B(\tilde{P}_A) = -EC_B + \tilde{P}_A(\pi_B(1,1) - \pi_B(0,1)) + (1 - \tilde{P}_A)(\pi_B(0,1) - \pi_B(0,0)) + \beta\left(\tilde{P}_A(V_B(1,1) - V_B(0,1)) + (1 - \tilde{P}_A)(V_B(0,1) - V_B(0,0))\right)
$$

Since the mapping $A(u_i(\cdot))$ maps the compact and convex space of firms’ choice probabilities $[0,1]^2$ into itself. By Brouwer’s fixed point theorem, there exists a fixed point in this mapping. In particular, (2.5) and (2.6) ensure that this fixed point cannot be at the border of $[0,1]^2$. This ensures that the fixed point is indeed the solution to the system of equations (2.3) and (2.4). This completes the proof.

### 2.2 Proof of Proposition 2

**Proof.** In the previous section, it’s been shown that $\{\tilde{P}_i, \forall i \in \{A, B\}\}$ or $(\tilde{P}_A, \tilde{P}_B)$ satisfy (2.5) and (2.6). If we can show that in benchmark MPEs,

$$
V_A(1, s_B) - V_A(0, s_B) = EC_A, \forall s_B \in \{0, 1\}
$$

$$
V_B(s_A, 1) - V_B(s_A, 0) = EC_B, \forall s_A \in \{0, 1\}, \tag{2.7}
$$


then (2.5) and (2.6) are also the equilibrium conditions for the one-shot game in Table 6. Then it would follow that \( \bar{P}_A, \bar{P}_B \) is the equilibrium solution to the one-shot game.

It is easy to show that conditions in (2.7) hold. Notice that based on (2.3.1), in equilibrium,

\[
V_A(s) = P_A(s) P_B(s) \pi_A(1,1) + P_A(s)(1-P_B(s)) \pi_A(1,0) \\
+ (1-P_A(s)) P_B(s) \pi_A(0,1) + (1-P_A(s))(1-P_B(s)) \pi_A(0,0) \\
+ \beta[P_A(s) P_B(s) V_A(1,1) + P_A(s)(1-P_B(s)) V_A(1,0) \\
+ (1-P_A(s)) P_B(s) V_A(0,1) + (1-P_A(s))(1-P_B(s)) V_A(0,0)]] + 1(s_A=0)(-P_A(s) EC_A) + 1(s_A=1)(1-P_A(s)) SV_A \\
+ P_A(s) e_A^P(1) + (1-P_A(s)) e_A^P(0)
\]

(2.8)

Since \( P_i(s) = \bar{P}_i, \forall i \in \{A,B\}, s \in S \) and \( EC_A = SV_A \), the following is true:

\[
V_A(1,s_B) - V_A(0,s_B) = (1-\bar{P}_A) SV_A + \bar{P}_A EC_A \\
= EC_A, \forall s_B \in \{0,1\}
\]

Similarly for firm B, \( V_B(s_A,1) - V_B(s_A,0) = EC_B, \forall s_A \in \{0,1\} \). This completes the first half of the proof.

Now consider an equilibrium CCPs of the one-shot game \( (\bar{P}^*_A, \bar{P}^*_B) \), it must be true that it satisfies the following conditions:

\[
\bar{P}^*_A = A(u_A(\bar{P}^*_B)) \\
\bar{P}^*_B = A(u_B(\bar{P}^*_A))
\]

(2.9) (2.10)

where

\[
u_A(\bar{P}^*_B) = -(1-\beta) EC_A \\
+ \bar{P}^*_B(\pi_A(1,1) - \pi_A(0,1)) + (1-\bar{P}^*_B)(\pi_A(1,0) - \pi_A(0,0)) \\
u_B(\bar{P}^*_A) = -(1-\beta) EC_B \\
+ \bar{P}^*_A(\pi_B(1,1) - \pi_B(1,0)) + (1-\bar{P}^*_A)(\pi_B(0,1) - \pi_B(0,0))
\]

For the CCPs with the construct \( \{\bar{P}_i(s) \equiv P^*_i, \forall i \in \{A,B\}, s \in S \} \), we can calculate the terms \( V_A(1,s_B; \hat{P}) - V_A(0,s_B; \hat{P}) \) and \( V_B(s_A,1; \hat{P}) - V_B(s_A,0; \hat{P}) \) based on (2.8)

\[57\] and these terms satisfy condition (2.7). We can then substitute \( \beta EC_A \) with \( \beta(V_A(1,s_B; \hat{P}) - V_A(0,s_B; \hat{P})) \) and \( \beta EC_B \) with \( \beta(V_B(s_A,1; \hat{P}) - V_B(s_A,0; \hat{P})) \) in (2.9) and (2.10), and we get

\[
\bar{P}_A(s) = A(u_A(\bar{P}_B(s))) \\
\bar{P}_B(s) = A(u_B(\bar{P}_A(s)))
\]

(2.11) (2.12)

\[57\] More strictly speaking, these terms should be \( \Gamma_A(1,s_B; \hat{P}) - \Gamma_A(0,s_B; \hat{P}) \) and \( \Gamma_B(s_A,1; \hat{P}) - \Gamma_B(s_A,0; \hat{P}) \). I use the notation here a little loosely. Note that equation (2.8) works for both equilibrium scenarios and out of equilibrium scenarios; that is, \( \Gamma_i(s; \hat{P}) \) has exactly the same expression as \( V_i(s; \hat{P}) \) (see (2.6)).
where

\[
\begin{align*}
    u_A \left( \hat{P}_B (s) \right) &= - EC_A \\
    &\quad + \hat{P}_B (s) \left( \pi_A (1, 1) - \pi_A (0, 1) \right) + \left( 1 - \hat{P}_B (s) \right) \left( \pi_A (1, 0) - \pi_A (0, 0) \right) \\
    &\quad + \beta \left( \hat{P}_B (s) \left( V_A \left( 1, 1; \bar{P} \right) - V_A \left( 0, 1; \bar{P} \right) \right) + \left( 1 - \hat{P}_B (s) \right) \left( V_A \left( 1, 0; \bar{P} \right) - V_A \left( 0, 0; \bar{P} \right) \right) \right)
\end{align*}
\]

\[
\begin{align*}
    u_B \left( \hat{P}_A (s) \right) &= - EC_B \\
    &\quad + \hat{P}_A (s) \left( \pi_B (1, 1) - \pi_B (0, 1) \right) + \left( 1 - \hat{P}_A (s) \right) \left( \pi_B (0, 1) - \pi_B (0, 0) \right) \\
    &\quad + \beta \left( \hat{P}_A (s) \left( V_B \left( 1, 1; \bar{P} \right) - V_B \left( 0, 1; \bar{P} \right) \right) + \left( 1 - \hat{P}_A (s) \right) \left( V_B \left( 0, 1; \bar{P} \right) - V_B \left( 0, 0; \bar{P} \right) \right) \right)
\end{align*}
\]

These conditions are equivalent to the equilibrium conditions for the dynamic game as shown in (2.1) and (2.2). Therefore \( \{ \hat{P}_i (s) \equiv \hat{P}^*_i, \forall i \in \{ A, B \}, s \in S \} \) constitutes the CCPs of a benchmark MPE. This completes the proof.

\[
\square
\]

.3 Extension of Definition to General Model

The general model in Section 2.2 differs from the simplified model in Section 2.3 only in that it incorporates stochastic demand and costs through multiple values in the exogenous market condition variable \( z \). Including multiple values in \( z \) does not significantly alter the definition of preemptive motives. Here I develop the definition and counterfactual of preemption in the context of the general model.

Following a similar procedure as in Section 2.3.1 I focus on firm A’s problem first and identify preemptive motives in firm A’s equilibrium conditions. Firm B’s problem is analogous. For the ease of writing out close-form expressions, I again assume that the private shock \( \varepsilon \) follows the extreme value type I distribution. It should be noted, however, that the definition of preemptive motives does not depend on the specification of the distribution of \( \varepsilon \).

At each state, firm A’s integrated value function is

\[
V_A (s, z) \equiv \max_{P_A (s, z) \in (0, 1)} \{ P_A (s, z) P_B (s, z) \pi_A (1, 1, z) + P_A (s, z) (1 - P_B (s, z)) \pi_A (1, 0, z) \\
+ (1 - P_A (s, z)) P_B (s, z) \pi_A (0, 1, z) + (1 - P_A (s, z)) (1 - P_B (s)) \pi_A (0, 0, z) \\
+ \beta \sum_{z'} \left[ P_A (s, z) P_B (s, z) V_A (1, 1, z') + P_A (s, z) (1 - P_B (s, z)) V_A (1, 0, z') \right] f_A (z'|z) \\
+ (1 - P_A (s, z)) P_B (s, z) V_A (0, 1, z') + (1 - P_A (s, z)) (1 - P_B (s, z)) V_A (0, 0, z') \} f_A (z'|z) \\
+ I \{ s_A = 0 \} (- P_A (s, z) EC_A (z)) + I \{ s_A = 1 \} (1 - P_A (s, z)) SV_A (z) \\
+ P_A (s, z) \sigma_\varepsilon (\gamma - \ln P_A (s, z)) + (1 - P_A (s, z)) \sigma_\varepsilon (\gamma - \ln (1 - P_A (s, z))) \}
\]

where \( z \) represents today’s market size and \( z' \) is tomorrow’s.

The first order condition is
\( A^{-1} (P_A (s, z)) = I \{ s_A = 0 \} (\alpha C_A (z)) + I \{ s_A = 1 \} (\alpha V_A (z)) + P_B (s, z) (\pi_A (1, 1, z) - \pi_A (0, 1, z)) + (1 - P_B (s, z)) (\pi_A (1, 0, z) - \pi_A (0, 0, z)) + \beta \sum_{s'} (P_B (s, z) (V_A (1, 1, z') - V_A (0, 1, z')) + (1 - P_B (s, z')) (V_A (1, 0, z') - V_A (0, 0, z'))) f_z (z'|z) \) 

(3.1)

Again, firms’ preemptive motives are embedded in the difference between the continuation values \( V_A (1, s_B, z') - V_A (0, s_B, z'), \forall s_B \in \{0, 1\}, z' \in Z \). Similar to what we did in Section 2.3.1, we can break down these difference terms. First, we rewrite \( V_A \) in the matrix form with three components:

\[
V_A = \left( I - \beta F^P \right)^{-1} \left\{ \tilde{F}^P \pi_A + \epsilon v_A + e_A \right\},
\]

where \( \pi_A \) is a \( 4|Z| \) column vector that stacks the corresponding state-specific element \( \pi_A (s, z') \); \( \epsilon v_A \) is the vector that stacks the expected entry cost \( EC_A (z') \) paid or scrap value \( SV (z') \) received at each corresponding state; \( e_A = \sum_{a \in A} |P_A (a) * \sigma (\gamma - ln P_A (a))| \); \( F^P \) is the transition matrix of the states; \( \tilde{F}^P \) is a diagonal block matrix with the following expression

\[
\tilde{F}^P = \begin{bmatrix}
G(z_1) & 0 & \cdots & 0 \\
0 & G(z_2) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & G(z_{|Z|})
\end{bmatrix},
\]

where \( G(z_n), n = 1, \ldots , |Z| \) is a \( 4 \times 4 \) matrix with each element

\[
g(s'|s, z_n) = \sum_{(a_A, a_B) \in A^2} P_A (a_A|s, z_n) P_B (a_B|s, z_n) I \{ (s'_A, s'_B) = (a_A, a_B) \},
\]

and \( s \) represents today’s market structure and \( s' \) tomorrow’s market structure.

Again let \( \Upsilon_{\pi_A} \equiv \left( I - \beta F^P \right)^{-1} \tilde{F}^P \pi_A, \Upsilon_{\epsilon v_A} \equiv \left( I - \beta F^P \right)^{-1} \epsilon v_A, \) and \( \Upsilon_{e_A} \equiv \left( I - \beta F^P \right)^{-1} e_A \). Then \( \Upsilon_{\pi_A} \) is the only term that is related to firms’ flow profits from the spatial competition. Firm A’s preemptive motive is embedded in the difference terms \( \Upsilon_{\pi_A} (1, s_B, z') - \Upsilon_{\pi_A} (0, s_B, z'), \forall s_B \in \{0, 1\}, z' \in Z \). To separate out the preemptive motives, we can replace \( P_B (1, s_B, z') \) with \( P_B (0, s_B, z') \) in all \( \Upsilon_{\pi_A} (\cdot) \) terms to create \( \tilde{\Upsilon}_{\pi_A} (\cdot) \) terms:

\[
\tilde{\Upsilon}_{\pi_A} \equiv \left( I - \beta \tilde{F}^P \right)^{-1} \tilde{F}^P \pi_A,
\]

where \( \tilde{F}^P \) and \( \tilde{F}^P \) are the results of replacing in \( F^P \) and \( F^P \) all \( P_B (1, s_B, z') \) with \( P_B (0, s_B, z') \), \( \forall s_B \in \{0, 1\}, z' \in Z \).

Then \( \Upsilon_{\pi_A} (1, s_B, z') - \tilde{\Upsilon}_{\pi_A} (0, s_B, z'), \forall s_B \in \{0, 1\}, z' \in Z \) would capture the gain accrued to firm A through the entry deterrence effect on firm B. They are firm A’s preemptive motives. Firm B’s preemptive motives can be similarly constructed.

**Definition 3.** The preemptive motives of firm A are \( \Upsilon_{\pi_A} (1, s_B, z) - \tilde{\Upsilon}_{\pi_A} (1, s_B, z), \forall s_B \in \{0, 1\}, z \in Z \) and those of firm B are \( \Upsilon_{\pi_B} (s_A, 1, z) - \tilde{\Upsilon}_{\pi_B} (s_A, 1, z), \forall s_A \in \{0, 1\}, z \in Z \).
To eliminate firms’ preemptive motives in the counterfactual, we need to set \( \Upsilon_{\pi A} (1, s_B, z') - \tilde{\Upsilon}_{\pi A} (0, s_B, z') = 0 \), \( \forall s_B \in \{0, 1\}, z' \in \mathbb{Z} \), and let firms re-optimize their values. Again due to the recursive nature of the \( \Upsilon_{\pi i} (\cdot) \) terms, this operation is equivalent to replacing all \( \Upsilon_{\pi i} (\cdot) \) terms with the corresponding \( \tilde{\Upsilon}_{\pi i} (\cdot) \) terms.

4 Extension of Definition to a Model With Two Locations

Suppose a market is a Hotelling line with two locations at the extremes: L1 and L2. Consumers are evenly distributed along the line. Firms can sell products at either locations. Products sold at the two locations can be seen as differentiated substitutes. In this regard, firms’ competition at one location will influence their profits at the other location. Including multiple locations in one market is to control for this type of correlation between firms’ activities in different submarkets. This type of correlation often takes the form of cannibalization; that is, a firm’s product at one location often takes away demand from the another location.

Compared to a one-location model, firms have more in their action set and the set of the market structure states is much larger. Section 4.1 below describes the model setup for a market with two locations. Section 4.2 demonstrates the decomposition of firms’ equilibrium conditions to isolate out the preemptive motives. Section 4.3 sets up the counterfactual of preemption.

4.1 Model

4.1.1 Players, Actions and Markov States

There are two firms: A and B. At the beginning of each period, firms can choose to be active or inactive at either L1 or L2 or both. Let \( i \in \{A, B\} \) denote a firm’s identity and \( a_{ilt} \in \{0, 1\} \) denote firm \( i \)’s action at location \( l \in \{L1, L2\} \) at time \( t \), with 1 indicating active and 0 inactive. Firm \( i \)’s action \( a_{ilt} \) is a vector \( a_{ilt} \equiv (a_{i1t}, a_{i2t}) \), and firm \( i \)’s action space is \( \mathcal{A} \equiv \{(0, 0), (1, 0), (0, 1), (1, 1)\} \).

The Markov states are payoff relevant states, including states that are common knowledge to both firms: exogenous variable \( z_t \) and market structure \( s_t \), and a private information shock \( \varepsilon_{ilt} \), which relates to the fixed cost, entry cost and scrap value. The Markov state at time \( t \) is a tuple \( (s_t, z_t, \varepsilon_{ilt}) \). Let \( s_{ilt} \) denote the active status of firm \( i \) at location \( l \) at the beginning of each period, with \( s_{ilt} \in \{0, 1\} \). Again, 1 indicates active and 0 inactive. Firm \( i \)’s active status in the market is summarized by the vector \( s_{it} \equiv (s_{i1t}, s_{i2t}) \), and the market structure is \( s_t = (s_{A1t}, s_{B1t}) \). Let \( S \) denote the set of market structures, with \( S \equiv \{(s_{A1l}, s_{A2l}, s_{B1l}, s_{B2l}) \mid s_{ilt} \in \{0, 1\}, \forall l \in \{1, 2\}\} \).

4.1.2 Timing of Events and Entry and Exit Costs

Firms interact with each other repeatedly for an infinite horizon. At the beginning of each period, both firms simultaneously make entry and exit decisions, and these decisions take effect right away. Each time a firm enters a location, it pays a one-time cost (EC), and every time it exits, it receives a one-time scrap value (SV). These entry costs and scrap values can be location specific. Let
\(EC_{il}\) and \(SV_{il}\) denote firm \(i\)'s entry cost and scrap value at location \(l\) respectively. Then the vector \(EC_i \equiv (EC_{i1}, EC_{i2})\) describes firm \(i\)'s entry cost structure in the market, and similarly \(SV_i \equiv (SV_{i1}, SV_{i2})\) for the scrap values. Both entry cost and scrap value can change over time due to price variations across different markets; in particular, they are functions of the state variable \(z_t\), with \(EC_i (z_t) : Z \rightarrow R^2\) and \(SV_i (z_t) : Z \rightarrow R^2, \forall i \in \{A, B\}\).

### 4.1.3 Flow Payoff Functions

A firm's total payoff per-period includes three components: (1) the flow profit \(\pi_i (\cdot)\) from a static spatial competition, where firms compete in prices or quantity with spatially differentiated products, (2) the entry costs or scrap values \(EC_i\) and \(SV_i\), and (3) the private shock \(\varepsilon_{it}\). Assume that \(\varepsilon_{it}\) is additively separable in the payoff function and \(i.i.d.\) across choice alternatives, we can write firm \(i\)'s choice-specific flow payoff function as

\[
\tilde{\Pi}_i (a_{it}, a_{-it}, s_t, z_t, \varepsilon_t) \equiv \pi_i (a_{it}, a_{-it}, z_t) + \varepsilon_{it} (a_{it}) + ev_{it} (a_{it}, s_t, z_t),
\]

where

\[
ev_{it} (a_{it}, s_t, z_t) \equiv -\langle (a_{it} - s_{it}) * a_{it} \rangle, EC_i (z_t) - \langle (a_{it} - s_{it}) * (1 - a_{it}) \rangle, SV_i (z_t)\),
\]

where the symbol \(*\) represents the element-by-element product, \(\langle \cdot, \cdot \rangle\) means the inner product of two vectors.

### 4.1.4 Strategies and conditional choice probabilities

Firms are assumed to play stationary Markov strategies; that is, they choose the same actions at the same state regardless of time. The time script of the model can thus be omitted. Let \(\sigma = (\sigma_B (s, z, \varepsilon), \sigma_A (s, z, \varepsilon))\) be a pair of strategy functions, with \(\sigma_i : Z \times S \times R^d \rightarrow A\). For each strategy profile \(\sigma\), we can define a set of conditional choice probabilities (CCPs)

\[
P^\sigma = \{(P^\sigma_A (s, z), P^\sigma_B (s, z)) : s \in S, z \in Z\}, \text{ where}\n\]

\[
P^\sigma_A (s, z) \equiv \{P^\sigma_A ((1, 0) | s, z), P^\sigma_A ((0, 1) | s, z), P^\sigma_A ((1, 1) | s, z))\}
\]

\[
P^\sigma_B (s, z) \equiv \{P^\sigma_B ((1, 0) | s, z), P^\sigma_B ((0, 1) | s, z), P^\sigma_B ((1, 1) | s, z))\}
\]

and each individual CCP is defined as follows

\[
P^\sigma_i (a | s, z) \equiv \text{Pr} (\sigma_i (s, z, \varepsilon_i) = a | s, z) = \int I \{\sigma_i (s, z, \varepsilon_i) = a\} g_i (\varepsilon_i) d\varepsilon_i,
\]

\[
\forall a \in A \setminus \{0, 0\}, i \in \{A, B\}
\]

The set of CCPs \(P^\sigma\) described above excludes \(P^\sigma_i ((0, 0) | s, z)\) because it is uniquely determined by the other three CCPs at the same \((s, z)\) state, i.e. \(P^\sigma_i ((0, 0) | s, z) = 1 - \sum_{a \in A \setminus \{0, 0\}} P^\sigma_i (a | s, z)\). Given our assumption regarding \(g_i (\cdot)\), there is a one-to-one mapping between \(\sigma_i (s, z, \varepsilon_i)\) and \(P^\sigma_i (s, z)\).

### 4.1.5 Firms' problems

Let \(\tilde{\pi}_i (a_i, s, z)\) denote firm \(i\)'s expected flow payoff without \(\varepsilon_i (a_i)\) if it chooses action \(a_i\), and the other firm behaves according to strategy \(\sigma_{-i}\). If firm \(i\)'s expectation of the other firm's actions is
consistent with $\sigma_{-i}$, then

$$\hat{\pi}_i^\sigma (a_i, s, z) = \sum_{\sigma_{-i} \in A} P_{\sigma_{-i}}^\sigma (a_{-i} | s, z) \pi_i (a_i, a_{-i}, z) + cv_i (a_i, s, z).$$

Let $\hat{V}_i^\sigma (s, z, \epsilon_i)$ denote firm $i$’s value if it acts optimally now and in the future given the other firm’s strategy $\sigma_{-i}$. Firm $i$ faces the following problem:

$$\hat{V}_i^\sigma (s, z, \epsilon_i) = \max_{a_i \in A} \left\{ \hat{\pi}_i^\sigma (a_i, s, z) + \epsilon_i (a_i) + \beta \sum_{(s', z') \in S \times Z} \int \hat{V}_i^\sigma (s', z', \epsilon_i) g_i (\epsilon_i) d\epsilon_i f_i^\sigma (s', z' | s, z, a_i) \right\}, \tag{4.1}$$

where $\beta \in (0, 1)$ is the discount factor and $f_i^\sigma (s', z' | s, z, a_i)$ is the transition probability from $(s, z)$ to $(s', z')$ if firm $i$ chooses $a_i$ and the other firm plays strategy $\sigma_{-i}$. Given that $z$ is exogenous, we can write:

$$f_i^\sigma (s', z' | s, z, a_i) = f_z (z' | z) \sum_{a_{-i} \in A} P_{\sigma_{-i}}^\sigma (a_{-i} | s, z) I \{(s', z', a_i) = (a_i, a_{-i})\}.$$

Denote $\int \hat{V}_i^\sigma (s, z, \epsilon_i) g_i (\epsilon_i) d\epsilon_i$ by $V_i^\sigma (s, z)$. Based on the Bellman equation (4.1), we can write

$$V_i^\sigma (s, z) = \max_{a_i \in A} \{v_i^\sigma (a_i, s, z) + \epsilon_i (a_i)\} g_i (\epsilon_i) d\epsilon_i, \tag{4.2}$$

where $v_i^\sigma (a_i, s, z) \equiv \hat{\pi}_i^\sigma (a_i, s, z) + \beta \sum_{(s', z') \in S \times Z} V_i^\sigma (s', z') f_i^\sigma (s', z' | s, z, a_i)$. The function $V_i^\sigma (s, z)$ is called the integrated value function.

### 4.1.6 Markov perfect equilibria

The Bellman equation (4.1) says that given a strategy of the other firm $\sigma_{-i}$, firm $i$ maximizes its current and expected future utilities at every state. The chosen strategy $\sigma_i^*$ by firm $i$ is the best response to the other firm’s strategy $\sigma_{-i}$. To form an equilibrium, $\sigma_{-i}$ has to be the best response to $\sigma_i^*$ as well. Formally, a stationary Markov perfect equilibrium can be defined as

**Definition 4.** A stationary Markov perfect equilibrium (MPE) is a pair of strategy functions $\sigma^*$ such that for any firm $i \in \{A, B\}$ and any state $(s, z, \epsilon_i) \in Z \times S \times R^4$, $\sigma_i^*$ solves the Bellman equation (4.1).

Again, similar to the model with one location, firms’ strategies can be represented in probability space. The representation and notation for this model with two locations are almost identical to those in Section 2.2.10 and thereby will not be repeated here.

### 4.2 Decomposition and Definition

Following a similar procedure as in Section 2.3.1 I focus on firm A’s problem first and identify preemptive motives in firm A’s equilibrium conditions. Firm B’s problem is analogous. Here, for simplicity of notation, I normalize firms’ flow profits when not in the market to 0; that is, $\pi_A ((0, 0), s_B, z) = \pi_B (s_A, (0, 0), z) = 0$.

For the ease of writing out close-form expressions, I assume that the private shock $\epsilon$ follows the extreme value type I distribution with location parameter 0 and scale parameter $\sigma_\epsilon$. At each state,
firm A’s integrated value function is

$$V_A(s, z) = \max_{P_A(a_A \neq (0,0)|s,z)} \sum_{a_A \in A} P_A(a_A|s,z) \times$$

$$\left\{ \sum_{a_B \in A} P_B(a_B|s,z) \left[ \pi_A(a_A, a_B, z) + \beta \sum_{z'} V_A(a_A, a_B, z') f_z(z'|z) \right] \right\}$$

$$- (1 - s_A) a_{A1} EC_{A1} + (1 - a_{A1}) s_A SV_{A1} - (1 - s_{A2}) a_{A2} EC_{A2} + (1 - a_{A2}) s_{A2} SV_{A2}$$

$$+ \sigma_z (\gamma - \ln(P_A(a_A|s,z)))$$

where

$$P_A((0,0)|s,z) = 1 - \sum_{a_A \in A\setminus(0,0)} P_A(a_A|s,z),$$

$$P_B((0,0)|s,z) = 1 - \sum_{a_B \in A\setminus(0,0)} P_B(a_B|s,z).$$

The equilibrium conditions for each $P_A(a_A \neq (0,0)|s,z)$ are

$$\sigma_z \ln \frac{P_A(a_A|s,z)}{P_A((0,0)|s,z)} = - (1 - s_A) a_{A1} EC_{A1} + (1 - a_{A1}) s_A SV_{A1}$$

$$- (1 - s_{A2}) a_{A2} EC_{A2} - (1 - a_{A2}) s_{A2} SV_{A2}$$

$$+ \sum_{a_B \in A} P_B(a_B|s,z) \left[ \pi_A(a_A, a_B, z) - \pi_A((0,0), a_B, z) \right]$$

$$+ \beta \sum_{a_B \in A} P_B(a_B|s,z) \left[ \sum_{z'} [V_A(a_A, a_B, z') - V_A((0,0), a_B, z')] f_z(z'|z) \right],$$

where

$$P_A((0,0)|s,z) = 1 - \sum_{a_A \in A\setminus(0,0)} P_A(a_A|s,z),$$

$$P_B((0,0)|s,z) = 1 - \sum_{a_B \in A\setminus(0,0)} P_B(a_B|s,z).$$

(4.3)

Here I consider that the incumbency status of a firm at any location in the overall market has a deterrence effect on its rival. This is because whenever a firm is already in the market and is committed to staying in the market, its rival’s overall expected payoffs from entering the market declines. The definition of preemptive motives in the general model therefore focuses on the gain from detering its rival from entering the overall market. Based on the first order conditions, firm A’s preemptive motive is embedded in the difference term $V_A(a_A, a_B, z') - V_A((0,0), a_B, z')$, $\forall a_B \in A, a_A \in A\setminus(0,0), z' \in Z$, which represents the marginal benefit of entry into the market in the future. To rewrite these terms in terms of Markov states instead of actions, we have $V_A(s_A, s_B, z) - V_A((0,0), s_B, z), \forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z$.

Similar to what we did in Section 2.3.1, we can further break down these difference terms. First, we can rewrite $V_A$ in the matrix form with three components:

$$V_A = (I - \beta F)^{-1} \left\{ \hat{F} \pi_A + \epsilon \nu_A + e_A \right\},$$

where $\pi_A$ is a $16|Z|$ column vector that stacks the corresponding state-specific element $\pi_A(s, z)$; $\epsilon \nu_A$ is the vector that stacks the expected entry cost $EC_A(z)$ paid (or scrap value $SV(z)$ received)
at each corresponding state; \( e_A = \sum_{a \in A} |P_A(a) \ast \sigma_x (\gamma - lnP_A(a))|; \) \( F^P \) is the transition matrix of the states; \( \tilde{F}^P \) is a diagonal block matrix with the following expression

\[
\tilde{F}^P = \begin{bmatrix}
G(z_1) & 0 & \cdots & 0 \\
0 & G(z_2) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & G(z_{|Z|})
\end{bmatrix},
\]

where \( G(z_n), n = 1, \cdots |Z| \) is a 16 \( \times \) 16 matrix with each element

\[
g(s'|s,z_n) = \sum_{(a_A,a_B) \in A^2} P_A(a_A|s,z_n) P_B(a_B|s,z_n) I \{ (s',s_B') = (a_a,a_B) \}
\]

Again let \( \Upsilon_{\pi_A} \equiv (I - \beta F^P)^{-1} \tilde{F}^P \pi_A, \) \( \Upsilon_{e\pi_A} \equiv (I - \beta F^P)^{-1} e\pi_A, \) and \( \Upsilon_{\pi_A} \equiv (I - \beta F^P)^{-1} e_A. \) Then \( \Upsilon_{\pi_A} \) is the only term that is related to firms’ flow profits from the spatial competition. Firm A’s preemptive motive is embedded in the difference terms \( \Upsilon_{\pi_A}(s_A,s_B,z) - \Upsilon_{\pi_A}(0,0,s_B,z), \) \forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z.

Since in this definition, I focus on a firm’s gain from deterring its rival from participating in the overall market, for firm A, this gain has to come from the fact that \( P_B((0,0)|s_A,s_B,z) > P_B((0,0),(0,0),s_B,z), \forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z. \) That is, the probability of firm B choosing not to be active in the market is higher when firm A is an incumbent in the market, regardless of the location. To extract this entry deterring effect on firm A’s NPV of flow profits, I replace \( P_B((0,0)|s_A,s_B,z) \) with \( P_B((0,0),(0,0),s_B,z), \forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z \) when calculating \( \Upsilon_{\pi_A}. \) In other words, firm A's incumbency status in the market will not reduce firm B’s likelihood of entering (staying in) the market. Furthermore, because firm A has 4 choices of actions, the choice probabilities associated with other actions would have to add up to 1 when combined with \( P_B((0,0),(0,0),s_B,z). \) To do so, I prorate the other probabilities over the remaining probability of \( 1 - P_B((0,0),(0,0),s_B,z). \) Specifically, I create a new set of choice probabilities \( \tilde{P}_B(\cdot) \) for each state \( (s_A,s_B,z), \forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z: \)

\[
\tilde{P}_B(a_B|s_A,s_B,z) \equiv \begin{cases} 
P_B((0,0),(0,0),s_B,z) & \text{if } a_B = (0,0) \\
\frac{(1-P_B((0,0),(0,0),s_B,z))}{P_B(a_B|s_A,s_B,z)} & \text{if } a_B \neq (0,0)
\end{cases}
\]

This set of probabilities represents an alternative behavior of firm B where firm A’s incumbency status has no impact on firm B’s decision to be active or inactive in the overall market. With this set in hand, I further construct the term

\[
\tilde{\Upsilon}_{\pi_A} \equiv (I - \beta \tilde{F}^P)^{-1} \tilde{F}^P \pi_A,
\]

where \( \tilde{F}^P \) and \( \tilde{F}^P \) are the results of replacing in \( F^P \) and \( \tilde{F}^P \) all the \( P_B(a_B|s_A,s_B,z) \) with \( \tilde{P}_B(a_B|s_A,s_B,z), \forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z. \) Then \( \tilde{\Upsilon}_{\pi_A}(s_A,s_B,z) - \tilde{\Upsilon}_{\pi_A}(s_A,s_B,z), \forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z \) would capture the gain accrued to firm A through the entry deterrence effect on firm B. This is because \( \tilde{\Upsilon}_{\pi_A}(s_A,s_B,z) \) results from the set of probabilities \( \tilde{P}_B(\cdot), \) where firm A’s incumbency status has no preemptive effect on firm B. I thereby define firm A’s and firm B’s preemptive motives as
Definition 5. The preemptive motive of firm A is $\pi_A(s_A, s_B, z) - \hat{\pi}_A(s_A, s_B, z), \forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z$, and that of firm B is $\pi_B(s_A, s_B, z) - \hat{\pi}_B(s_A, s_B, z), \forall s_A \in S_A, s_B \in S_B \setminus (0,0), z \in Z^{58}$

Note that this definition is not limited to the case with the private shock $\varepsilon$ following an extreme value type I distribution. In fact, this definition applies to any distribution that satisfies Assumption 2 in Section 2.2.6. This is because when firms choose their actions, they look at the marginal benefits of taking one action versus another. Definition 5 captures the components of preemptive motives in those marginal benefits. For any distribution of $\varepsilon$, these marginal benefits come from the differences between the overall payoffs associated with each action. Therefore, this definition of preemptive motive should not be restricted only to cases with the extreme value type I distribution.

4.3 Counterfactual

The counterfactual for this model is similarly constructed as that in Section 2.3.2. When firms calculate the expected NPV of their flow profits in the future, they assume that their incumbency status in the market at any location would not discourage their rival’s entry (or stay) in the market. In particular, they assume that their rival’s likelihood of entry at the states with their incumbency status to be the same as those when they are not in the market. That is, when firm A calculates $\pi_A(s, z), \forall s \in S$, it replaces all $P_B(a_B|s_A, s_B, z)$ with $\hat{P}_B(a_B|s_A, s_B, z)$, $\forall s_B \in S_B, s_A \in S_A \setminus (0,0), z \in Z$; and similarly, when firm B calculates $\pi_B(s, z), \forall s \in S$, it replaces all $P_A(a_A|s_A, s_B, z)$ with $\hat{P}_A(a_A|s_A, s_B, z)$, $\forall s_A \in S_A, s_B \in S_B \setminus (0,0), z \in Z^{59}$

The conditions for the counterfactual are thereby

$$
\sigma_A \ln \frac{P_A(a_A|s_A, s_B, z)}{P_A(\{(0,0)|s_A, s_B\}}} = - (1 - s_{A1}) a_{A1} EC_{A1}(z) - a_{A1}s_{A1} SV_{A1}(z)
- (1 - s_{A2}) a_{A2} EC_{A2}(z) - a_{A2}s_{A2} SV_{A2}(z)
+ \sum_{a_B \in A} P_B(a_B|s, z) [\pi_A(a_A, a_B, z) - \pi_A((0,0), a_B, z)]
+ \beta \sum_{a_B \in A} P_B(a_B|s, z) \left[ \sum_{z'} \left( \hat{\pi}_{A1}(a_A, a_B, z') - \hat{\pi}_{A1}(\{(0,0), a_B, z\}) \right) f_z(z|z') \right]
+ \beta \sum_{a_B \in A} P_B(a_B|s, z) \left[ \sum_{z'} \left( \hat{\pi}_{A2}(a_A, a_B, z') - \hat{\pi}_{A2}(\{(0,0), a_B, z\}) \right) f_z(z|z') \right]
+ \beta \sum_{a_B \in A} P_B(a_B|s, z) \left[ \sum_{z'} \left( \hat{\pi}_{A1}(a_A, a_B, z') - \hat{\pi}_{A1}(\{(0,0), a_B, z\}) \right) f_z(z|z') \right] \tag{4.4}
$$

\[58\] Again, $\pi_B(s_A, s_B, z) - \hat{\pi}_B(s_A, s_B, z), \forall s_A \in S_A, s_B \in S_B \setminus (0,0), z \in Z$ terms are similarly constructed as their counterparts for firm A.

\[59\] Here $\hat{P}_A(a_A|s_A, s_B, z), \forall s_A \in S_A, s_B \in S_B \setminus (0,0), z \in Z$ are also similarly constructed as $\hat{P}_B(a_B|s_A, s_B, z).$
\[
\sigma \ln \frac{P_B(a_B|s,z)}{P_B((0,0)|s,z)} = - (1 - s_{B1}) a_{B1} EC_{B1}(z) - a_{B1} s_{B1} SV_{B1}(z)
\]
\[
- (1 - s_{B2}) a_{B2} EC_{B2}(z) - a_{B2} s_{B2} SV_{B2}(z)
\]
\[
+ \sum_{a \in A} P_A(a_A|s,z) [\pi_B(a_A, a_B, z) - \pi_B(a_A, (0,0), z)]
\]
\[
+ \beta \sum_{a \in A} P_A(a_A|s,z) \left( \sum_{z'} \left[ \hat{\Upsilon}_{\pi B}(a_A, a_B, z') - \hat{\Upsilon}_{\pi B}(a_A, (0,0), z') \right] f_z (z'|z) \right)
\]
\[
+ \beta \sum_{a \in A} P_A(a_A|s,z) \left( \sum_{z'} \left[ \Upsilon_{\pi B}(a_A, a_B, z') - \Upsilon_{\pi B}(a_A, (0,0), z') \right] f_z (z'|z) \right)
\]
\[
+ \beta \sum_{a \in A} P_A(a_A|s,z) \left( \sum_{z'} \left[ \Upsilon_{\pi B}(a_A, a_B, z') - \Upsilon_{\pi B}(a_A, (0,0), z') \right] f_z (z'|z) \right)
\] \text{(4.5)}

This system of equations determine firms’ CCPs in the counterfactual where firms’ preemptive motives are eliminated.
Chapter 3

Aggressive Growth in Retail: A Trade-off Between Deterrence and Survival?
3.1 Introduction

In the retail industry, aggressive and preemptive expansion by incumbent firms may have potential long-run benefits in the form of entry deterrence (Eaton and Lipsey, 1979; Schmalensee, 1978). Researchers have recently quantified empirically the underlying preemptive motives that drive rapid entry and expansion in the retail industry (e.g., Fang, 2016; Igami and Yang, 2016; Zheng, 2016). However, this potential benefit largely depends on the extent to which their strategies actually prevent rival firms from entering. Preemptive efforts may be highly ineffective (and even harmful) for the incumbent if the potential entrant is strong enough to withstand intensified incumbent presence. Failure to prevent rivals from entering ex post reduces the market’s future viability, which may ultimately jeopardize outlet survival in the long-run as a natural consequence would be market exit. Survival of outlets is important for the retail chains given the costs associated with both entry and exit. Therefore, a managerially relevant question for incumbent retailers is whether or not there exists a trade-off between aggressive entry and long-run survival.

A better understanding about the presence of a trade-off would help incumbent retailers characterize potential risks associated with aggressive entry. We aim to study such risks by developing a simple analytical framework to determine whether there is a positive or negative relationship between an incumbent retailer’s aggressive entry into markets and long-run survival after entry. The framework we use can be described as a dynamic game of entry between two firms across local and isolated markets. Under the context of this game, we then obtain aggressive and non-aggressive entry strategies using a decomposition technique developed by Fang (2016). By characterizing strategies as such, we can compare measures of survival between these types of strategies. The measure of survival we develop looks at the share of outlets an incumbent is able to keep active in light of an unexpected productivity shock that makes the rival suddenly stronger, where we define rival strength as their ability to shield themselves from business-stealing effects associated with the incumbent’s presence. Our numerical analysis demonstrates that depending on how strong the rival becomes, aggressive entry can lead to either higher or lower survival rates. If the rival does not become too strong, then aggressive entry can actually help the incumbent keep more of its outlets open. In contrast, if the rival becomes too strong, then aggressive entry may lead to a situation in which an incumbent has to close many of its stores.

To apply these insights to an actual industry, we turn to the fast casual restaurant sector and focus on the two main chains in Texas, Taco Cabana (the incumbent) and Chipotle (the entrant).

---

1Preemptive motives have also been studied empirically in other industry contexts (e.g., Chiu, 2012; Ellison and Ellison, 2011; Igami, 2016; Schmidt-Dengler, 2006).
2See, for example, “Retailers Need to Close Some Doors to Survive” (Wall Street Journal, June 17, 2016).
3Exit is particularly risky if retail expansion is financed primarily with recapitalization via debt, given that retailers may face pressure to achieve a sufficiently high (and stable) return that pays off the debt’s interest; and this practice of debt-based financing is not uncommon in retail (Love, 1995). For a recent example of debt-financed retail expansion, see Wilson (2016).
4Avoiding overly aggressive entry strategies have become a priority for some retailers, like Wal-Mart. See, for example, “Wal-Mart Takes Its Time on Expanding in Africa” (Wall Street Journal, February 4, 2017).
5This decomposition technique was somewhat inspired by Besanko, Doraszelski and Kryukov (2014).
This setting matches our analytical framework, as the two restaurant chains offer very close substitutes of menu items, compete in spatially proximate areas, and entry (and exit) are the most important strategic decisions they have to make. It is also important to note that all of the outlets for both Taco Cabana and Chipotle are company-owned, which then allows us to abstract away from the additional strategic layer involving organizational form as it has been shown that preemptive motives affect the franchising mix decision (e.g., Nishida and Yang, 2016). With monthly data that tracks entry and revenue from 1993 to 2015 across census tracks, we estimate the dynamic game of entry that we originally used to obtain analytical insights about the aggression-survival trade-off. Access to revenue information is particularly important for our application as we need to assess how strong the entrant is in order to determine the extent to which the aggression-survival trade-off is a risk. Our estimates reveal that Chipotle is a stronger competitor to Taco Cabana, as the presence of Chipotle leads to greater business-stealing that hurts Taco Cabana’s revenue than the extent to which Taco Cabana’s presence hurts Chipotle’s revenues. Taken together, these estimates would suggest that a trade-off may exist between aggressive entry and survival for Taco Cabana, and thus, the incumbent should be cognizant of such risks.

This paper contributes to both the analytical and empirical work on early entry strategies. Related analytical work includes that of Shen and Villas-Boas (2010), who rationalize why firms enter at faster rates than demand growth during early periods. We complement this analytical work by providing new insights about the long-run benefits or risks of early entry. Empirical work in about early entry has largely been focusing on decomposing early mover advantages or disadvantages (e.g., Adams, Gans, Hayes and Lampe, 2016; Blevins, Khwaja and Yang, 2016; Boulding and Christen, 2003; Boulding and Christen, 2008; Boulding and Christen, 2009; van Heerde, Srinivasan and Dekimpe, 2010; Nishida, 2016) as well as responses by incumbents to the threat of entry (e.g., Goolsbee and Syverson, 2008; Seamans, 2012; Shankar, 2006; West, 1981). Our work narrows in on the risks specific to the preemptive benefit associated with being a first-mover, as we depart from the implicit assumption that preemptive entry will effectively deter rivals from entering, and instead focus on the possibility that such strategies will not always achieve their intended objectives. More generally, our work is related to recent research about dynamic retail entry and expansion (e.g., Arcidiacono, Bayer, Blevins and Ellickson, 2016; Beresteau, Ellickson and Misra, 2010; Blevins, Khwaja and Yang, 2016; Fang, 2016; Hollenbeck, 2016; Igami and Yang, 2016; Suzuki, 2013; Zheng, 2016; Yang, 2015; Yang, 2016). This past work has largely been focused on quantifying various elements of profit using entry and expansion data. We add to this research by comparing the long-term viability of equilibrium entry strategies available to managers (i.e., aggressive versus non-aggressive strategies).

6We point out that using the inferred sensitivity to competition from entry data alone may not accurately reveal the entrant’s strength, as insensitivity to the incumbent’s presence could indeed be because of the entrant’s strength, or instead, the entrant’s obliviousness to competition (e.g., Goldfarb and Xiao, 2011; Goldfarb and Yang, 2009).

7Not only do we see this in the relationship between revenue and rival presence, but also in the reduced form policy functions that approximate entry decisions.

8Other examples of preemption games include Fudenberg and Tirole (1985), Reinganum (1981), and Riordan (1992).
Chapter 3

Our paper proceeds as follows. The modelling framework we rely on is introduced in Section 2. Section 3 describes the analytical exercise we undertake to study the relationship between aggressive entry and survival. Our empirical application is introduced in Section 4, where we provide details about the fast casual industry along with key data patterns. After discussing the empirical setting, we describe the estimation approach used in Section 5. Section 6 summarizes the main findings from estimation. We conclude in Section 7.

3.2 Model

We describe the dynamic game of retail entry in this section. The model is based on the dynamic oligopoly framework set forth by Ericson and Pakes (1995). First, we discuss the actions and payoffs for the firms in our model. Next, we lay out strategies that the firms may follow. Finally, we provide an equilibrium concept for the dynamic game.

3.2.1 Actions and Payoffs

Consider a retail chain industry with two forward-looking firms, \( i \) and \( j \). Firms are forward looking with a discount factor of \( \beta \in (0,1) \), and seek to maximize their long-run discounted payoffs in an infinite horizon game.

At the beginning of each time period \( t \), the firms can choose whether or not to be active in a local market, whereby their decision takes effect right away. We let \( a_{it} \in \{0,1\} \) denote firm \( i \)'s action at time \( t \). The action space is \( \mathcal{A} \equiv \{0,1\} \). The payoff relevant state variables include the market size \( z_t = (z_{it}, z_{jt}) \) and market structure \( s_t \). The active status of firm \( i \) is denoted with \( s_{it} \in \{0,1\} \). Then \( s_t = (s_{it}, s_{jt}) \), and \( S = \{(0,0), (0,1), (1,0), (1,1)\} \). Both firms can observe \( z_t \) and \( s_t \). Finally, as in typical games of incomplete information, we allow for a private idiosyncratic shock \( \varepsilon_{it} \). Therefore, the payoff relevant states for firm \( i \) can be summarized by the tuple \( (s_t, z_t, \varepsilon_{it}) \). There is a cost of entry, which we denote as \( EC_i \). Entry costs may include expenses such as acquiring property to house the outlet. Finally, there is a scrap value from leaving a market, which we denote as \( SV_i \). The scrap value may come from the amount recovered from liquidating an outlet (e.g., real estate value, equipment).

Given the action space, payoff relevant states, we can define the flow payoff functions. The revenue generated is denoted as \( R_i(\cdot) \). Assuming that \( \varepsilon_{it} \) is additively separable in the profit function and i.i.d., we write the payoffs as

\[
\Pi_i(a_{it}, a_{jt}, s_{it}, z_{it}, \varepsilon_{it}) \equiv R_i(a_{it}, a_{jt}, z_{it}) + C_i(a_{it}, s_{it}) + \varepsilon_{it}(a_{it}) \tag{3.2.1}
\]

where we define the sunk costs as \( C_i(\cdot) = (a_{it} - s_{it})[a_{it}EC_i + (1 - a_{it})SV_i] \).
3.2.2 Strategies

We assume that the firms play stationary Markov strategies. As such, we omit the time subscript for notational simplicity. Let $\sigma = (\sigma_i(s,z,\varepsilon), \sigma_j(s,z,\varepsilon))$ define the strategy functions for the two firms. For each strategy profile $\sigma$, the conditional choice probabilities (CCPs) can be defined as

$$P_\sigma^i(a_i = 1|s,z) \equiv \Pr(\sigma_i(s,z,\varepsilon_i = 1|s,z) = \int \{\sigma_i(s,z,\varepsilon_i = 1|s,z) = 1\} f_i(\varepsilon_i) d\varepsilon_i. \quad (3.2.2)$$

By construction, $P_\sigma^i(a_i = 0|s,z) = 1 - P_\sigma^i(a_i = 1|s,z)$.

3.2.3 Equilibrium

In our subsequent descriptions, we summarize the payoff relevant states as $X = (s,z,\varepsilon)$. Given these payoff relevant states, the firms choose their strategies so as to solve the following recursive optimization problem:

$$V_i(X,\sigma) = E[\Pi_i(X,\sigma(X)) + \beta E(V_i(X',\sigma)|X,\sigma_i(X),\sigma_j(X)]. \quad (3.2.3)$$

Firms follow a Markov Perfect Equilibrium such that the strategy profile satisfies the following condition for all $i$:

$$V_i(X,\sigma|\sigma_i^*,\sigma_j^*) \geq V_i(X,\sigma|\sigma_i,\sigma_j^*) \quad (3.2.4)$$

for all $\sigma_i$ and all states $X$, where $V_i(\cdot)$ is a Bellman equation defined using the following recursive expression:

The strategy profile in equilibrium is defined such that no firm $i$ has an incentive to deviate from the optimal strategy $\sigma_i^*$, in that no alternative strategy ($\sigma_i$) yields higher expected discounted profits than $\sigma_i^*$ while its rivals use strategies $\sigma_j^*$.

3.2.4 Definition and Measure of Deterrence

The definition and measure of deterrence used in this paper are based on Fang (2016), which derives a measure of preemptive motives for a dynamic discrete oligopoly game. In particular, Fang (2016) identifies the preemptive motive component in firms’ equilibrium conditions, and by forcing firms to ignore their preemptive motives when making entry decisions, Fang (2016) obtains a counterfactual that eliminates the effect of preemption without changing the underlying economic fundamentals. Compared to other recent attempts to quantify preemption in the literature (e.g., Igami and Yang, 2016; Zheng, 2016), Fang’s approach has the advantage of preserving the direct competitive effect and the dynamic features of the game. For this reason, we choose to employ her measure in this study.

In brief, Fang (2016) identifies the preemptive motives through a decomposition of a firm’s marginal benefit of entry in the equilibrium conditions. The marginal benefit of entry consists of a component in the current period and the one from all future periods. Preemptive motives
are embedded in that from all future periods because preemption usually represents a short-term sacrifice but long term gain. Fang (2016) then breaks down the marginal benefit of entry from the future further into 3 components and identifies the last one as being attributed to preemptive motives:

1. The capital investment component, which comes from the entry costs and scrap values that a firm pays upon entry or receives upon exit. A firm can optimize its timing of entry by anticipating changes in entry costs and scrap values in the future. If it expects that a sharp increase in entry costs in the future, it might choose to enter a market early in order to save costs. Similarly if it expects a high return from scrap values upon exit at a particular time in the future, it could also choose to enter today. This marginal benefit from entry is part of the future payoffs but is irrelevant to entry deterrence.

2. The single-agent-sales component, which represents a firm’s net present value of streams of profits through sales in all future periods when its incumbency status has no impact on its rival’s entry behavior. This component alone can motivate a firm to enter a market. As long as the stream of profits in the future is large enough to cover the initial entry cost, a firm would be willing to enter today. This component does not account for the return on a firm’s future profits through manipulating its action today in order to affect its rival’s entry behavior in the future and thereby is not relevant to entry deterrence.

3. The entry-deterrence component, which captures the gain in a firm’s net present value of future profits through the effect of its incumbency status on its rival’s entry behaviors. This component represents a firm’s preemptive motives. For example, if a firm’s being active in the market discourages its rival from entering the market in the next period, then by entering today, this firm can expect a higher probability of earning monopolistic profits in the future; as a result, this firm will enter the market aggressively today.

To eliminate preemptive motives, Fang (2016) proposes replacing the rival’s probability of entry at a firm’s active state with that at this firm’s inactive state in the future part of the marginal benefit of entry in the equilibrium conditions. Through this replacement, a firm’s incumbency status does not affect its rival’s behavior in the future, and preemptive motives are thereby shut down. Based on these new equilibrium conditions, firms will choose another set of optimal strategies, only that this new set of strategies is now clear of entry deterrence motives. These new strategies provide a counterfactual for how firms will behave without the intent to deter entry.

In the context of the model in this paper, the counterfactual scenario can be constructed by
rewriting firms’ equilibrium conditions as follows:

\[
F^{-1} \left( \frac{P_i(1|s,z)}{P_i(0|s,z)} \right) = 1_{(s_i=0)EC_i} + 1_{(s_i=1)(-SV_i)} + P_j(1|s,z)(R_i(1,1,z) - R_i(0,1,z)) + P_j(0|s,z)(R_i(1,0,z) - R_i(0,0,z))
\]

\[
+ \beta \sum_{z'} f(z'|z) \left[ P_j(1|s,z)(\bar{T}_{R_i}(1,1,z') - \bar{T}_{R_i}(0,1,z')) + P_j(0|s,z)(\bar{T}_{R_i}(1,0,z') - \bar{T}_{R_i}(0,0,z')) \right]
\]

\[
+ \beta \sum_{z'} f(z'|z)P_j(1|s,z)(\psi_{evi}(1,1,z') - \psi_{evi}(0,1,z')) + P_j(0|s,z)(\psi_{evi}(1,0,z') - \psi_{evi}(0,0,z'))
\]

\[
+ \beta \sum_{z'} f(z'|z)P_j(1|s,z)(\psi_{ei}(1,1,z') - \psi_{ei}(0,1,z')) + P_j(0|s,z)(\psi_{ei}(1,0,z') - \psi_{ei}(0,0,z'))]
\]

\forall s \in S, z \in Z. \quad (3.2.5)

where \( \psi_{evi} (\cdot) \) and \( \psi_{ei} (\cdot) \) come from \( \psi_{evi} \equiv (I - \beta F^P)^{-1} e_{v_i} \) and \( \psi_{ei} \equiv (I - \beta F^P)^{-1} e_{i} \) respectively, and \( e_{v_i} \) is the vector that stacks the expected entry cost \( EC_i \) or scrap value \( SV_i \) at each corresponding state, and \( e_i \) is a vector of the expected values of the error terms \( \varepsilon \) conditional on the actions taken at each state. \( F^P \) is the transition matrix of all the states, and \( f(z'|z) \) is the transition probability of the exogenous states from \( z \) to \( z' \). The \( \bar{T}_{R_i} (\cdot) \) terms are created by the following equation:

\[
\bar{\psi}_{R_i} \equiv (I - \beta \bar{F}_i^P)^{-1} \bar{F}_i^P R_i, \quad (3.2.6)
\]

where \( R_i \) is a vector that stacks the corresponding state-specific element \( R_i (s,z) \); \( \bar{F}_i^P \) is the result of replacing in \( F^P \) all \( P_j (a|1, s_j, z') \) with \( P_j (a|0, s_j, z') \), \( \forall s_j \in \{0,1\}, z' \in Z \), and \( \bar{F}_i^P \) is a diagonal block matrix with the following expression:

\[
\bar{F}_i^P = \begin{bmatrix}
G_i(z_1) & 0 & \cdots & 0 \\
0 & G_i(z_2) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & G_i(z_Z)
\end{bmatrix} \quad (3.2.7)
\]

where \( G_i(z_n), n = 1, \ldots |Z| \) is a \( 4 \times 4 \) matrix with each element

\[
g_i (s'|s, z_n) = \sum_{(a_s, a_j) \in A^2} P_i (a_s | s, z_n) P_j (a_j | 0, s_j, z_n) I \{ (s'_i, s'_j) = (a_i, a_j) \}. \quad (3.2.8)
\]

The right hand side of equation (3.2.5) contains the various components of firm \( i \)'s entry motives. Overall, the first two lines reflect firm \( i \)'s marginal benefit of entry in the current period, and the last 3 lines represent that in the future periods. In particular, the third line represents the single-agent-sales component of the firm’s marginal benefit of entry from the future. This interpretation comes from that the \( \bar{\psi}_{R_i} (\cdot) \) terms are created by replacing \( P_j (a|1, s_j, z') \) with \( P_j (a|0, s_j, z') \), \( \forall s_j \in \{0,1\}, z' \in Z \) in firm \( i \)'s expectation of how the state will transition in the future; that is, when firm \( i \) calculates its future expected profits, it expects (wrongly) that its state of being active has no impact on its rival’s entry probability. The \( \bar{\psi}_{R_i} (\cdot) \) then captures firm \( i \)'s net present value of sales only through its own entry and exit behaviors but not through the impact on its rival’s behaviors.
The last two lines of equation 3.2.5 represent the capital investment component of the firm’s marginal benefit of entry in the future periods. As explained earlier, the \( \Upsilon_{evi}(\cdot) \) and the \( \Upsilon_{ei}(\cdot) \) terms are created based on firm \( i \)'s entry costs and scrap values and the idiosyncratic shocks to these costs.

As can be seen, of the last 3 lines of equation 3.2.5, the marginal benefit of entry from the future does not contain the entry-deterrence component. Equation 3.2.5 thus leaves out firm \( i \)'s entry-deterrence motives. In addition, the new set of equilibrium conditions in Equation 3.2.5 preserves the direct competitive effect and the dynamic nature of firms’ interactions. The direct competitive effect of firm \( j \) on firm \( i \)'s CCPs is captured by the first two lines of equation 3.2.5 because they represent the current period competition where firms make simultaneous decisions based on their expectations (not observations) of their rival’s actions. The dynamic nature of firm \( i \)'s forward-looking behavior is captured by the capital investment and single-agent-sales components in the last 3 lines, which represent marginal payoffs from the future.

In summary, the only component that is left out in this new set of equilibrium conditions is the strategic entry-deterrence component. Thus, solving this new set of equilibrium conditions for each firm will give us the counterfactual CCPs which firms choose without the preemptive motives.

### 3.3 Assessing the Aggression-Survival Trade-Off

This section illustrates numerically the relationship between aggression and survival. First, Subsection 3.3.1 discusses the concept of survival used in this paper and describes the measure of survival. Then Subsection 3.3.2 shows through two numerical examples that the relationship between aggression and survival is not clear cut. Under some conditions, aggression harms survival, and there is a trade-off between the two. However, under some other conditions, aggression can in fact help a firm survive.

#### 3.3.1 Measuring Survival

The concept of survival in our paper relates to how an incumbent firm’s strategy will change if its rival suddenly becomes more competitive. Suppose an incumbent chain plays an aggressive strategy against its rival, and their strategies form an equilibrium. In the steady-state distribution, the incumbent would densely populate the markets with many outlets. However, if the rival receives a positive productivity shock which makes it permanently more competitive, the incumbent would need to change its strategy, and firms’ new strategies would lead to a new equilibrium. In the new steady-state distribution, the incumbent may not be able operate as many outlets as before. This is because a more competitive rival could render the incumbent’s original aggressive strategy ineffective, and the incumbent may not be able to stop the rival from entering the market by densely populating the market. A stronger competition from the rival means that the incumbent may have to close many outlets, incurring a large loss. Had it played a more accommodative strategy from the beginning, it would not have opened as many outlets, its loss may not be as severe. This is the
context in which we examine the relationship between aggression and survival.

Specifically, our measure of survival is constructed as follows: Let \( f^\sigma(s) \) denote the probability of state \( s \) in the steady-state distribution formed by a strategy profile \( \sigma \), and \( \mu^\sigma_{In} \) denote the proportion of the markets in which the incumbent firm \( In \) is active; that is, \( \mu^\sigma_{In} = \sum_s f^\sigma(s)I\{s_{In}=1\} \). The measure \( \mu^\sigma_{In} \) can be seen as the expected number of the incumbent’s outlets in a market. Let \( \mu^{MPE}_{In} \) denote the incumbent’s expected number of outlets in the steady-state distribution resulting from the MPE before the rival becomes stronger, and \( \mu^{MPEa}_{In} \) denote the one after, then our measure of survival \( Sv^{MPE}_{In} \) is

\[
Sv^{MPE}_{In} = \frac{\mu^{MPEa}_{In}}{\mu^{MPEb}_{In}}
\]

That is, the survival rate measures what percentage of the incumbent’s outlets survive after the change.

To examine the relationship between aggression and survival, we need to obtain the survival rate of the incumbent for a counterfactual scenario where the incumbent plays a non-preemptive strategy both before and after the rival becomes stronger. The construction of that counterfactual has been described in Section 3.2.4. Let \( \mu^{Cftb}_{In} \) denote the incumbent’s expected number of outlets in the steady-state distribution in the counterfactual before the change and \( \mu^{Cfta}_{In} \) after, then the survival rate under the non-preemptive scenario is then:

\[
Sv^{Cft}_{In} = \frac{\mu^{Cfta}_{In}}{\mu^{Cftb}_{In}}
\]

Comparing \( Sv^{MPE}_{In} \) with \( Sv^{Cft}_{In} \) will allow us to evaluate the relationship between aggression and survival. The following subsection demonstrates in two numerical exercises that aggression sometimes helps the incumbent firm survive a stronger rival, but sometimes it severely harms the incumbent. What dictates whether aggression leads to one scenario versus another is the degree to which a rival becomes stronger.

### 3.3.2 Is There a Trade-Off Between Aggression and Survival?

In general, this relationship depends on the range of parameters of firms’ profit functions. When the rival becomes stronger but not strong enough to make the incumbent firm’s aggressive strategies ineffective, the heightened competition could give a greater incentive to the incumbent for preemption. That is, the incumbent could act even more aggressively in order to deter the entry of the rival. This case is shown in Subsection 3.3.2.1. Once the rival grows more competitive to a point that renders the incumbent firm’s aggressive strategies less effective, meaning the incumbent’s competitive effect on the rival is weaker and the incumbent could incur a much greater sacrifice for preemptive behaviors, the incumbent will change its strategy to a much more accommodative one. In that case, the incumbent will be forced to close many of its outlets and suffer a great loss. This scenario is discussed in Subsection 3.3.2.2. In the former case, aggression helps the incumbent firm survive, but
in the latter case, aggression harms the incumbent. Had the incumbent behaved non-preemptively, its loss would be much less.

The following tables and paragraphs describe the structural primitives of the economic setting before the rival becomes stronger. We choose a setup that resembles a growing market. In particular, we let the exogenous state variable $z$ take on two values $z \in \{z_1, z_2\}$, and it follows a Markov process with the transition matrix

$$F_z = \begin{bmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{bmatrix}$$

At the smaller market size state $z_1$, firms’ profit functions are such that without the threat of entry from the rival $En$, the incumbent does not want to enter the market very frequently. The incumbent would enter the market frequently in the small market only to preempt the entry of its rival. As the market grows to $z_2$, both firms want to be active in the market if their rivals are not; however, they do not want to be active at the same time because the competitive effects between the two firms would make both of their profits very low. Specifically, firms’ profit functions are shown in Table 3.1.

As shown in Table 3.1 at the small market size $z_1$, the incumbent’s profit for being active is always lower than that of being inactive: $4.5 < 5.0$ and $0.5 < 2.0$. The rival, on the other hand, makes a greater profit being active in the market if the incumbent is not: $2.0 > 0.0$. At the larger market size $z_2$, both firms make more profit being active in the market if their rivals are not. Note that in the table, the incumbent’s profit functions are such that even when it is not in the market, its profits are still positive, unlike the rival whose profits for being inactive are always normalized to 0. Giving the incumbent a positive profit for not being active in the market is consistent with the notion that the incumbent should already have outlets in other markets, and its is looking to expand in this new market. Therefore, its per-period profits should be positive even if it does not expand to this new market.

---

9We choose a growing market setting to make the numerical exercises reflect more of the real world. Furthermore, a growing market is also where preemption has caught the most attention in the literature (Eaton and Lipsey, 1979; Shen and Villas-Boas, 2010).

<table>
<thead>
<tr>
<th>State $s$</th>
<th>Incumbent’s Revenue $R_{In}$</th>
<th>Entrant’s Revenue $R_{En}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(0, 0, z_1)$</td>
<td>5.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$(0, 1, z_1)$</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>$(1, 0, z_1)$</td>
<td>4.5</td>
<td>0.0</td>
</tr>
<tr>
<td>$(1, 1, z_1)$</td>
<td>0.5</td>
<td>-0.1</td>
</tr>
<tr>
<td>$(0, 0, z_2)$</td>
<td>5.5</td>
<td>0.0</td>
</tr>
<tr>
<td>$(0, 1, z_2)$</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>$(1, 0, z_2)$</td>
<td>7.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$(1, 1, z_2)$</td>
<td>1.0</td>
<td>-2.2</td>
</tr>
</tbody>
</table>
Firms’ entry costs and scrap values are set below:

\[ EC_{In}(z_1) = 1 \quad EC_{In}(z_2) = 1 \quad EC_{En}(z_1) = 1 \quad EC_{En}(z_2) = 1 \]
\[ SV_{In}(z_1) = -4 \quad SV_{In}(z_2) = -4 \quad SV_{En}(z_1) = 1 \quad SV_{En}(z_2) = 1 \]

Given these values, the incumbent firm’s sunk cost is 5 while the entrant’s is 0. With a relatively large positive sunk cost, the incumbent firm can commit strategically. In particular, it can commit to staying in the market after entry. The rival, however, cannot commit to staying in the market after entry with 0 sunk cost because upon exit, it can always recoup its initial investment \( EC_{En} \) from its scrap values. In other words, no structural primitives guarantee that the rival can stay in the market after entry. Under this setup, the incumbent firm is the only firm that can play an aggressive strategy in equilibrium. We restrict our numerical examples to this setting to focus on the relationship between aggression and survival for the incumbent firm only.

Firms’ discount factor \( \beta \) is set to 0.95 and the variance of the private information shock \( \sigma_{\epsilon} \) is 0.8. With these structural primitives, we solve for firms’ CCPs for the preemptive equilibrium and non-preemptive counterfactual scenarios respectively. Firms’ probabilities of being active at each state under each scenario are shown in Table 3.2.

As shown in Table 3.2, the incumbent firm enters the market at all states much more frequently in the preemptive MPE scenario than in the non-preemptive counterfactual scenario. The rival, on the other hand, enters the market much less frequently in the MPE. Firms’ strategies as shown by these CCPs indicate that the incumbent plays an aggressive entry strategy in the MPE in order to keep the rival out of the market most of the time. In particular, in the steady-state distribution associated with the preemptive MPE, the incumbent firm is active in 99.77% of the markets while the rival active in only 15.71% of the markets. Under the counterfactual scenario, however, the incumbent is active in only 0.2031% of the markets in the steady-state distribution, and the rival 91.81%.

Now suppose the rival firm receives a productivity shock and becomes permanently more com-

| State s   | \( P_{In}^{MPE}(a = 1|s, z) \) | \( P_{En}^{MPE}(a = 1|s, z) \) | \( P_{In}^{CFT}(a = 1|s, z) \) | \( P_{En}^{CFT}(a = 1|s, z) \) |
|-----------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| (0, 0, z_1) | 0.1890                        | 0.8745                        | 0.0003759                     | 0.9196                        |
| (0, 1, z_1) | 0.1890                        | 0.8745                        | 0.0003759                     | 0.9196                        |
| (1, 0, z_1) | 0.9952                        | 0.4564                        | 0.1712                        | 0.8795                        |
| (1, 1, z_1) | 0.9952                        | 0.4564                        | 0.1712                        | 0.8795                        |
| (0, 0, z_2) | 0.8516                        | 0.1158                        | 0.0007458                     | 0.9194                        |
| (0, 1, z_2) | 0.8516                        | 0.1158                        | 0.0007458                     | 0.9194                        |
| (1, 0, z_2) | 0.9997                        | 0.05673                       | 0.8194                        | 0.1342                        |
| (1, 1, z_2) | 0.9997                        | 0.05673                       | 0.8194                        | 0.1342                        |
petitive. In particular, its competitive effects on the incumbent firm change their profit functions for the large market size $z_2$. The following examples illustrate two situations: the first one is when the rival firm’s enhanced competitiveness does not hurt the incumbent enough such that the incumbent adopts a more accommodative equilibrium strategy. In this case, the rival’s stronger competitive effects make the incumbent act even more aggressively, especially at the small market size state $z_1$. Compared to the incumbent’s strategies in the non-preemptive counterfactual, its aggressive strategies help it survive. The second numerical example shows that if the rival grows even more competitive and the incumbent’s competitive effect on the rival is no longer substantial, then the incumbent’s preemptive strategies would be much less effective. Furthermore, the rival’s heightened competitive effects on the incumbent implies that the incumbent would incur a greater sacrifice if it engages in head-to-head competition with the rival by acting aggressively. Under this scenario, the incumbent would adopt a more accommodative strategy in equilibrium, and in the new steady-state distribution, the incumbent is active in much fewer markets, which implies that it would have to close many outlets in face of a more formidable competitor. Had it behaved non-preemptively at the beginning, its loss would be much less. The details of these two numerical examples are described in the following two subsections.

### 3.3.2.1 Scenario #1: Positive Relationship Between Aggression and Survival

In this numerical example, firms’ profit functions at the large market size $z_2$ change to those values shown in Table 3.3.

Compared to the values in Table 3.1, the rival now steals more customers from the incumbent firm. In particular, the rival’s profit at state $(0, 1, z_2)$ is bigger than before 3.0 v.s. 2.0, and the incumbent’s profit is smaller than before 1.5 v.s. 2.0. At state $(1, 1, z_2)$, the rival’s profit is -1.2 instead of -2.2 previously and the incumbent’s profit is lower at 0.5 instead of 1. Although the rival is stronger, the incumbent still has a substantial competitive effect on the rival: if the incumbent is in the market, it could turn the rival’s profit to -1.2. To avoid a head-to-head competition with the incumbent, the rival would try avoiding markets where the incumbent is active. This means that preemptive strategies could still be effective and have an important impact.

We obtain firms’ MPE CCPs under these new parameter values, and compare them with firms’ MPE CCPs before the change. Firms’ CCPs before and after under the MPE scenarios are shown in Table 3.4.

<table>
<thead>
<tr>
<th>State $s$</th>
<th>Incumbent’s Revenue $R_{In}$</th>
<th>Rival’s Revenue $R_{En}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(0, 0, z_2)$</td>
<td>5.5</td>
<td>0.0</td>
</tr>
<tr>
<td>$(0, 1, z_2)$</td>
<td>1.5</td>
<td>3.0</td>
</tr>
<tr>
<td>$(1, 0, z_2)$</td>
<td>7.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$(1, 1, z_2)$</td>
<td>0.5</td>
<td>-1.2</td>
</tr>
</tbody>
</table>
As shown in Table 3.4, the incumbent’s probabilities of being active at those states for the small market size $z_1$ are now much bigger than before. In particular, at states $(0, 0, z_1)$ and $(0, 1, z_1)$, the incumbent’s entry probabilities change to 0.2889 from 0.1890. This indicates that the stronger rival at the larger market size $z_2$ now makes the incumbent want to act even more aggressively ex-ante, i.e. before the market demand takes off.

The incumbent’s CCPs at $(0, 0, z_2)$ and $(0, 1, z_2)$ are lower in the MPE after the change. This reflects the more severe competitive effect of the rival on the incumbent at the larger market size $z_2$. Since the incumbent’s profits at $z_2$ are lower than before for those states when the rival is in the market, the incumbent enters the market less frequently at the larger market $z_2$. Overall, Table 3.4 shows that a stronger competitor in a larger market could give the incumbent a greater incentive to preempt entry before the market demand takes off. In this case, aggressive strategies could help the incumbent survive.

To calculate the survival rate for the preemptive MPE scenario, we obtain the incumbent’s expected number of outlets in the market in the new steady-state distribution; it is 0.9986, compared to 0.9977 previously. The survival rate for the incumbent’s outlets is

$$S_{In}^{MPE} = \frac{\mu_{In}}{\mu_{En}} = 1.001$$

Under this scenario, the incumbent in fact adds more outlets to the market after the rival becomes stronger.

To compare this survival rate to that in the non-preemptive counterfactual scenario, we calculate firms’ counterfactual CCPs under the new set of parameters and compare them to those before the change. Table 3.5 below illustrates these CCPs.

As shown in Table 3.5 under the counterfactual, the incumbent’s probabilities of being active increase slightly in most states. However, its exit rates at states $(1, 0, z_2)$ and $(1, 1, z_2)$ are much bigger than those before the change, 0.4009 (or 1 - 0.5991) vs. 0.1806 (or 1 - 0.8194). This larger exit rates reflect the greater competitive effect of the rival on the incumbent. With this larger exit
rates, the incumbent becomes less active in the market than before. In the steady-state distribution associated with this new counterfactual, the incumbent is active in a smaller proportion of the markets with 0.1787% instead of 0.2031%. The survival rate of the incumbent’s outlets under the non-preemptive counterfactual is then

\[
S_{Cft}^{\text{In}} = \frac{\mu_{Cft}^{\text{In}}}{\mu_{Cft}^{\text{En}}} = 0.8801
\] (3.3.5)

If the incumbent plays a non-preemptive strategy, it would need to close about 12% of its outlets after the rival becomes more competitive.

It should be noted that although we use a point in the parameter space in this numerical example, this example is not a special case. It can be shown through a numerical exercise using the homotopy method that for a range of parameters, aggression could help the incumbent survive a stronger competitor.

### 3.3.2.2 Scenario #2: Negative Relationship Between Aggression and Survival

In this example, the rival grows even more competitive than in the first one. Firms’ profit functions at \(z_2\) now take the values shown in Table 3.6.

As shown in Table 3.6, for those states \((0, 1, z_2)\) and \((1, 1, z_2)\) when the rival is in the market,

Table 3.6: Profit Functions of Firms at Various States (After, Case 2)
the incumbent’s profits decrease to 1.0 and -1.0 respectively from 2.0 and 1.0 as before. These decreases are also steeper than those shown in Numerical Example 1. The large decreases imply that to preempt the entry of its rival, the incumbent now needs to incur a much greater sacrifice. By excessively entering the market, the incumbent would face more head-to-head competition with the rival. The greater competitive effects from the rival means that the incumbent could lose a large amount in profits. Furthermore, because the rival is now more competitive, the incumbent’s competitive effect on the rival is smaller, -0.01 compared to -2.2 as before. This smaller competitive effect of the incumbent on the rival implies that the incumbent’s preemptive strategies would be much less effective in deterring entry. Firms’ equilibrium strategies show that this is indeed the case: the incumbent now acts much less aggressively compared to before. Table 3.7 compares firms’ CCPs in the MPEs before and after the change.

As shown in Table 3.7, the incumbent plays a more accommodative strategy in the equilibrium after the change. The incumbent chooses to be active much less frequently compared to its strategy before. Although this strategy appears very accommodative, the incumbent still acts on preemptive motives. As discussed in Fang (2016), firms almost always act on preemptive motives in an MPE. To remove the incumbent’s preemptive motives entirely, we again obtain firms’ CCPs for the counterfactual scenario and compare them with those before the change. Table 3.8 below show these CCPs.

As shown in Table 3.8, the incumbent’s probabilities of being active are lower in all states at the larger market size $z_2$ compared to those before the change. The lower probabilities at $z_2$ reflect the stronger competitive effects that the rival has on the incumbent, making it less willing to enter the market. The incumbent’s probabilities of being active at the smaller market size $z_1$, however, are slightly higher than those before the change. This is mainly due to the fact that the incumbent makes greater profits at market structure states $(0, 1)$ and $(1, 1)$ in the small market than in the large one, $2>1$ and $0.5>-1$. The incumbent is likely to enter slightly more often in the small market after the change.

Furthermore, if we compare the incumbent’s counterfactual strategy after the change to that

| State $s$ | $P_{MPE_a}^{In}(a = 1 | s, z)$ | $P_{MPE_a}^{En}(a = 1 | s, z)$ | $P_{MPE_b}^{In}(a = 1 | s, z)$ | $P_{MPE_b}^{En}(a = 1 | s, z)$ |
|---|---|---|---|---|
| $(0, 0, z_1)$ | 0.0005864 | 0.9195 | 0.1890 | 0.8745 |
| $(0, 1, z_1)$ | 0.0005864 | 0.9195 | 0.1890 | 0.8745 |
| $(1, 0, z_1)$ | 0.2492 | 0.8561 | 0.9952 | 0.4564 |
| $(1, 1, z_1)$ | 0.2492 | 0.8561 | 0.9952 | 0.4564 |
| $(0, 0, z_2)$ | 0.0003000 | 0.9756 | 0.8516 | 0.1158 |
| $(0, 1, z_2)$ | 0.0003000 | 0.9756 | 0.8516 | 0.1158 |
| $(1, 0, z_2)$ | 0.1445 | 0.9587 | 0.9997 | 0.05673 |
| $(1, 1, z_2)$ | 0.1445 | 0.9587 | 0.9997 | 0.05673 |
Table 3.8: Firms’ CCPs for Non-Preemptive Counterfactuals (Before-After, Case 2)

| State s | $P_{In}^{Cft}(a = 1|s,z)$ | $P_{En}^{Cft}(a = 1|s,z)$ | $P_{In}^{Cfb}(a = 1|s,z)$ | $P_{En}^{Cfb}(a = 1|s,z)$ |
|---------|----------------|----------------|----------------|----------------|
| $(0, 0, z_1)$ | 0.0004878 | 0.9195 | 0.0003759 | 0.9196 |
| $(0, 1, z_1)$ | 0.0004878 | 0.9195 | 0.0003759 | 0.9196 |
| $(1, 0, z_1)$ | 0.2141 | 0.8671 | 0.1712 | 0.8795 |
| $(1, 1, z_1)$ | 0.2141 | 0.8671 | 0.1712 | 0.8795 |
| $(0, 0, z_2)$ | 0.0002579 | 0.9756 | 0.0007458 | 0.9194 |
| $(0, 1, z_2)$ | 0.0002579 | 0.9756 | 0.0007458 | 0.9194 |
| $(1, 0, z_2)$ | 0.1254 | 0.9614 | 0.8194 | 0.1342 |
| $(1, 1, z_2)$ | 0.1254 | 0.9614 | 0.8194 | 0.1342 |

under the MPE scenario in Table 3.7, we can see that the incumbent’s counterfactual probabilities of being active are smaller in all states than its MPE CCPs. This confirms the statement previously that although the incumbent plays a relatively accommodative strategy in the MPE after the change, its preemptive motives are still at work. The counterfactual eliminates those preemptive motives and makes the incumbent choose to be active in the market even less frequently.

To compare the survival rates under the preemptive scenario and the non-preemptive scenario, we obtain the incumbent’s expected number of outlets in the steady-state distribution associated with the new MPE as well as that under the new counterfactual. The incumbent’s expected number of outlets is 0.0004495 in the MPE scenario and 0.0003709 in the counterfactual. The survival rates under the MPE scenario and the counterfactual one are thus

$$S_{In}^{MPE} = \frac{\mu_{In}^{MPE}}{\mu_{In}} = \frac{0.0004495}{0.9977} = 0.0004506 \quad (3.3.6)$$

$$S_{In}^{Cft} = \frac{\mu_{In}^{Cft}}{\mu_{In}} = \frac{0.0003709}{0.002031} = 0.1826 \quad (3.3.7)$$

As shown in equations (3.3.6) and (3.3.7), by playing the aggressive strategies, the incumbent would have to close almost 100% of its outlets in the new market, which is an enormous loss to the incumbent: all the sunk costs associated with opening an outlet would receive no return on the investment. By following non-preemptive strategies, however, the incumbent can keep about 18.3% of its original outlets in face of a stronger competitor. Furthermore, because the incumbent opens much fewer outlets under the non-preemptive scenario before the change, its loss in absolute size is also much smaller.

Again, although this numerical example represents a point in the parameter space of the structural primitives, the result associated with this example covers a range of parameters. Overall the numerical examples in this section indicate that the relationship between aggression and survival is not clear cut. Under certain conditions, aggression helps the incumbent firm survive a tougher competitor, but under other conditions, aggression harms the survival of the incumbent firm’s out-
3.4 Empirical Application

We now analyze the trade-off between aggression and survival using data from the fast casual restaurant industry. Before discussing the estimation strategy and parameter estimates, we first discuss the industry’s background, data source, and raw data patterns.

3.4.1 Industry Background

The industry under examination is the fast casual restaurant chains. In particular, we focus on two largest fast casual chains in Texas, Taco Cabana and Chipotle. Fast casual restaurants are thought to be healthier and more upscale than quick service restaurants. They are different from traditional fast food restaurants in that their food is of higher quality and dining atmosphere is more upscale. Their menu is also more expensive than fast food and is more likely to offer alcoholic beverages.\(^\text{10}\) Compared to full service restaurants, such as Applebee’s, fast casual restaurants offer limited to no table service.

This sector has been the fastest growing sector within the restaurant industry in the past decade. According to NPD group, in year 2010, visits to fast casual restaurant chains grew by 6% while those in the overall industry declined by 1%. This growth momentum continued to year 2015. Due to its modern and upscale characteristics, fast casual dining is particularly appealing to the millennial.\(^\text{11}\) Although fast casual sector currently still accounts for a small fraction of the total quick service industry, about 7% in terms of traffic, it is forecast to grow in the double digits through 2022 due to the change in demographics, and will become a major player in the overall food service industry.\(^\text{12}\)

Within the fast casual restaurant sector, we focus on the two largest fast casual chains in Texas, Taco Cabana and Chipotle. Taco Cabana is the industry leader and incumbent. It was founded in San Antonio, Texas in 1978, and mainly expanded in Texas. At its peak, it reached 162 outlets in 2014, and today still operates 160 stores in Texas. Chipotle is the industry follower. Although a national chain founded in 1993, Chipotle did not enter Texas until late 2000, but it has grown very rapidly since. At the end of 2015, it reached a total number of outlets of 136. A number of characteristics of the two chains make them an ideal case for studying the relationship between preemption and survival.

First, these two restaurant chains offer close substitutes of menu items, and both provide relatively pleasant dining environments.\(^\text{13}\) Taco Cabana offers burritos, tacos, fajita, nachos, quesadillas, burritos, tacos, fajita, nachos, quesadillas,
salad and combo plates, and Chipotle offers burritos, tacos, salads and combo bowls. Taco Cabana features semi-enclosed and patio dining areas, while Chipotle prides itself on ultra-modern interiors. Both chains offer alcoholic beverages.

Second, the restaurant outlets of the two chains operate at a relatively small scale and compete proximate geographic areas. This industry feature allows for the formation of many “isolated” markets, and provides sufficient market variation for our econometric estimation.

Third, entry to and exit from a market are the most important strategic decisions for these two chains. Both chains offer a uniform menu and pricing across markets. Entering a market at an opportune time thereby becomes the most important avenue for chain growth and competition.

Fourth, both chains are company owned. All entry and exit decisions in markets are made at the headquarter level. This ownership structure removes complications from franchise relationships that are found in most other chain restaurants.

3.4.2 Data Description

Our data is derived from a variety of sources. The main data-set that contains information regarding Chipotle’s and Taco Cabana’s outlets is part of the Mixed Beverage Tax Information Records kept by the Office of the Comptroller of Public Accounts in Texas. The State of Texas collects mixed beverage gross receipts tax and sales tax from mixed beverage permit holders. These taxes are leveraged on alcoholic drinks that are sold, prepared or served by businesses. In addition to alcohol, these taxes are also due on nonalcoholic beverages and ice that are mixed with an alcoholic beverage. This data-set started in 1993 when the responsibility of recording alcoholic gross receipts was transferred to the Office of the Comptroller.

The mixed beverage data-set contains detailed information on each establishment that holds a mixed beverage permit, including taxpayer’s name and address, business location name and address, permit number, permit issue date, out of business date, gross receipts reporting date, liquor sales, beer sales, wine sales, and the total gross receipts which add up all sales from the previous three categories. Based on taxpayers’ information and outlet names, we were able to identify with accuracy the outlets run by Chipotle and Taco Cabana. We geocoded each outlet’s address by using Google Maps geocoding API.

To complement this data-set with market level information, we collected data on demographics, income, infrastructure and traffic volume data from a number of sources. For demographics and income, our data come mostly from the 1990-2010 decennial censuses. To control for the census geographic boundary definition changes during the three decennial census periods, we use GeoLytics’s harmonized census data-set, the Neighborhood Change Database [NCDB] Tract Data from 1970-2010. This database adjusts earlier censuses to 2010 census geography, making feasible the

Taco Bell is a fast food, not fast casual restaurant. Its food items are of much lower quality, and its price per item is mostly kept below $5, while for fast casual restaurants, the item price typically ranges from $9-$13 (The Washington Post, February 2, 2015). Given that it does not fit the characteristics of a fast casual chain, Taco Bell is not included in the study as a major fast casual sector player. Its competitive effects on both Taco Cabana and Chipotle will be controlled through other means.
inter-temporal comparison of demographic changes in a given area. The smallest geography in this database is the census tract, which we choose as our market definition. For the intercensal and post-censal periods, we use US Census Bureau’s intercensal estimates and American Community Survey. The intercensal estimates and some American Community Survey data are available only at the county level; we distributed them to the census tract level based on historical trends.

For information on local infrastructure and traffic volume related to each outlet, we collected traffic volume data and Texas road network GIS shape-files from the Department of Transportation of Texas. For chain restaurants like Taco Cabana and Chipotle, traffic volumes are often one of the most direct demand indicators. The traffic volume data includes annual average daily traveler counts for about 31,400 traffic monitoring stations in Texas. It covers the years of 1999 to 2014. We extrapolated the traffic data for years outside of this range in our data-set. The Texas road network file is used to detect if an outlet is close to an interstate highway or primary and secondary roads in order to account for potential unobserved demand or cost factors.

Combining all these data-set, we obtain a panel that covers monthly mixed beverage sales for every outlet of Taco Cabana and Chipotle during the period of December 1993 to December 2015, demographic and income information for each census tract, and infrastructure and traffic information based on each outlet’s location. In total, the data-set includes 310 geographic markets (or census tracts), spanning over 96 cities and places in 39 Texas counties.

Table 3.9 summarizes the data. Notice that given the granularity of our market-time information, we have over 80,000 observations to draw inferences from. Among all time-location markets, Chipotle is active in 15.1% of the markets, and Taco Cabana 35.8%. This pattern is consistent with Taco Cabana being the industry leader in fast casual dining in Texas. On average, Chipotle makes about $720 from selling alcoholic drinks every month in each market, while Taco Cabana makes about $3,850, almost 5.4 times as high as that of Chipotle’s. This difference is in agreement with the casual observation that customers at Taco Cabana order more alcoholic drinks than those at Chipotle. Although we do not have the total monthly revenue for each restaurant, we assume in our analysis that sales from alcoholic beverages account for a constant proportion of a restaurant’s monthly total revenue. Under this assumption, we write a restaurant’s total revenue as

\[ \log(\text{total revenue}) = \log(\text{mixed beverage sales}) + \text{constant}. \]

Based on the most direct demand indicator, the daily traffic faced by Taco Cabana is higher than that by Chipotle. Taco Cabana has a weighted daily traffic of 232 persons/meter while Chipotle’s is 172 persons/meter. This difference in daily traffic explains partially the higher revenues of Taco Cabana.

For the geographic markets under consideration, the average total population is about 5,280 with a mean population density at 1,510 persons per square kilometers. Population from the 15-64 age cohort accounts for 70% of the total population, and in terms of race, white and Hispanic populations account for over 80% of the total population. The average income per household is


15The weighted daily traffic is constructed from dividing the daily traffic counts at a traffic monitoring station that each restaurant is closest to by the distance in meters between the station and that restaurant.
about $59,000. Most markets are within the five largest urban areas, including Austin, Fort Worth, Dallas, Houston and San Antonio. Over one third of the markets are crossed by an interstate highway, and over 80% of the markets contain primary or secondary roads. This is not surprising considering that interstate highways and primary and secondary roads are the main locations for the outlets of these two restaurant chains.

Table 3.9: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Entry decisions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active status of Chipotle</td>
<td>0.151</td>
<td>0.358</td>
</tr>
<tr>
<td>Active status of Taco Cabana</td>
<td>0.358</td>
<td>0.479</td>
</tr>
<tr>
<td><strong>Revenue</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revenue for Chipotle</td>
<td>721.138</td>
<td>417.378</td>
</tr>
<tr>
<td>Revenue for Taco Cabana</td>
<td>3849.046</td>
<td>2619.112</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total population</td>
<td>5280.023</td>
<td>3008.664</td>
</tr>
<tr>
<td>Population density (persons per km²)</td>
<td>1510.105</td>
<td>1096.258</td>
</tr>
<tr>
<td>Share of population (15-34)</td>
<td>0.341</td>
<td>0.131</td>
</tr>
<tr>
<td>Share of population (35-64)</td>
<td>0.362</td>
<td>0.0814</td>
</tr>
<tr>
<td>Share of population (65 and up)</td>
<td>0.0986</td>
<td>0.0615</td>
</tr>
<tr>
<td>Share of Hispanic population</td>
<td>0.252</td>
<td>0.191</td>
</tr>
<tr>
<td>Share of white population</td>
<td>0.563</td>
<td>0.139</td>
</tr>
<tr>
<td>Share of black population</td>
<td>0.105</td>
<td>0.122</td>
</tr>
<tr>
<td>Share of Asian population</td>
<td>0.0638</td>
<td>0.0875</td>
</tr>
<tr>
<td>Inflation adjusted income</td>
<td>59424.396</td>
<td>29912.966</td>
</tr>
<tr>
<td><strong>Infrastructure and traffic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily traffic for Chipotle</td>
<td>172.096</td>
<td>244.169</td>
</tr>
<tr>
<td>Daily traffic for Taco Cabana</td>
<td>231.594</td>
<td>374.489</td>
</tr>
<tr>
<td>In big urban area</td>
<td>0.826</td>
<td>0.379</td>
</tr>
<tr>
<td>Interstate highway</td>
<td>0.387</td>
<td>0.487</td>
</tr>
<tr>
<td>Primary or secondary road</td>
<td>0.819</td>
<td>0.385</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td></td>
<td>82150</td>
</tr>
</tbody>
</table>

3.4.3 Entry and Revenue Patterns

We now discuss the underlying entry patterns found in the data. Table 3.10 provides a tabulation of all possible market configurations in our estimation sample. The market appears to be competitive in the sense that the two chains avoid co-locating in the same markets. A large number of configurations has the chains acting as sole incumbents in the geographic markets. Furthermore, Taco Cabana exhibits more sensitivity to competition relative to Chipotle’s sensitivity. Finally, we see that Taco Cabana in general has a greater presence across the markets, as it is the incumbent firm in this industry.

Figure 1 illustrate the dynamics in entry and exit in this industry. Recall that our data is not
only granular with respect to geographic markets, but also with respect to time as our observations are at the monthly level, so entry and exit varies at a high frequency. Taco Cabana is the first entrant in this industry, so it entered markets before Chipotle. In terms of exit, Taco Cabana exits markets on a more frequent basis than Chipotle. In some months, Chipotle can enter as many as 8 markets while Taco Cabana can enter as many as 12 markets. Finally, total exit is capped at around 2 per time period for both of the chains.

Figure 2 displays the evolution of revenue for the two chains. We see that Chipotle’s total revenue has been steadily rising, while Taco Cabana’s revenue decreased initially, but rises gradually in the latter years. If we look at the industry’s total revenue, the general trend has been upward over time; this growth in revenue is consistent with the analytical setting we use to assess the aggression-survival trade-off.

Tables 3.11 and 3.12 provide results from linear probability regressions for Chipotle and Taco Cabana respectively. All of the specifications contain market fixed effects, as well as time trends. These regressions reveal that both chains are sensitive to competition. Reduced form evidence of sensitivity to competition is important, as competition is a prerequisite to preemptive motives. Also note that there are some differences in the sensitivity to competition across the chains. In particular, Chipotle seems less sensitive to competition than Taco Cabana. As an aside, we also point out that firm-specific daily traffic appears to be the strongest indicator of market size. This pattern is important as it confirms the existence of a reliable firm-specific exclusion restriction that affects entry decisions. Finally, we see that Chipotle is drawn to big urban areas, while there exists no evidence of the same preference for Taco Cabana.

Taken together, these raw data patterns would suggest to us that Chipotle is a strong entrant. As highlighted in the analytical section, aggressive strategies may be less effective for the incumbent when the entrant is strong.

3.5 Empirical Strategy

This section describes the empirical strategy we use to estimate the model. We first discuss the two-step approach we use for estimating the dynamic game, followed by heuristic arguments for model identification.
Figure 1: Entry and Exit Dynamics

(a) Chipotle active

(b) Taco Cabana active

(c) Chipotle entry

(d) Taco Cabana entry

(e) Chipotle exit

(f) Taco Cabana exit
Table 3.11: Linear Probability Regression for Chipotle’s Decision to be Active

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active status of Taco Cabana</td>
<td>-0.102</td>
<td>-0.102</td>
<td>-0.102</td>
<td>-0.0740</td>
</tr>
<tr>
<td></td>
<td>(0.0265)</td>
<td>(0.0265)</td>
<td>(0.0265)</td>
<td>(0.0230)</td>
</tr>
<tr>
<td>Total population</td>
<td>-0.00000865</td>
<td>-0.0000100</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000702)</td>
<td>(0.00000650)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td>-0.235</td>
<td>0.586</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.435)</td>
<td>(4.433)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (15-34)</td>
<td>0.0000143</td>
<td>0.0000160</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000945)</td>
<td>(0.00000864)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (35-64)</td>
<td>0.0000122</td>
<td>0.0000108</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000103)</td>
<td>(0.00000929)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (65 and up)</td>
<td>-0.0000162</td>
<td>-0.00000368</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000137)</td>
<td>(0.0000103)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic population</td>
<td>-0.000000563</td>
<td>-0.000000772</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000236)</td>
<td>(0.00000188)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White population</td>
<td>0.00000250</td>
<td>0.00000259</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000195)</td>
<td>(0.00000188)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black population</td>
<td>0.000000350</td>
<td>0.00000146</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000305)</td>
<td>(0.00000248)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian population</td>
<td>0.000000508</td>
<td>-0.00000317</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000422)</td>
<td>(0.00000355)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflation adjusted income</td>
<td>-0.000000171</td>
<td>9.36e-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000000152)</td>
<td>(0.000000128)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily traffic for Chipotle</td>
<td>0.00128**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000468)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In big urban area</td>
<td>0.0955***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0190)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interstate highway</td>
<td>-0.0198</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0206)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary or secondary road</td>
<td>-0.0337</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0299)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0699***</td>
<td>-0.0699***</td>
<td>-0.0637**</td>
<td>-0.106**</td>
</tr>
<tr>
<td></td>
<td>(0.0167)</td>
<td>(0.0167)</td>
<td>(0.0213)</td>
<td>(0.0335)</td>
</tr>
<tr>
<td>Observations</td>
<td>82150</td>
<td>82150</td>
<td>82150</td>
<td>82150</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.2437</td>
<td>0.2437</td>
<td>0.2438</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Table 3.12: Linear Probability Regression for Taco Cabana’s Decision to be Active

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active status of Chipotle</td>
<td>-0.119***</td>
<td>-0.119***</td>
<td>-0.119***</td>
<td>-0.0723***</td>
</tr>
<tr>
<td></td>
<td>(0.0296)</td>
<td>(0.0296)</td>
<td>(0.0296)</td>
<td>(0.0261)</td>
</tr>
<tr>
<td>Total population</td>
<td>-0.00000419</td>
<td>-0.00000594</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000104)</td>
<td>(0.00000874)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td>3.905</td>
<td>0.222</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.005)</td>
<td>(4.369)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (15-34)</td>
<td>0.00000551</td>
<td>0.0000106</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000108)</td>
<td>(0.00000923)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (35-64)</td>
<td>-0.00000734</td>
<td>-0.00000264</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000134)</td>
<td>(0.0000115)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (65 and up)</td>
<td>0.0000120</td>
<td>0.0000108</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000132)</td>
<td>(0.0000112)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic population</td>
<td>0.00000135</td>
<td>0.00000114</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000330)</td>
<td>(0.00000266)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White population</td>
<td>-3.61e-09</td>
<td>0.000000961</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000267)</td>
<td>(0.0000204)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black population</td>
<td>-1.78e-08</td>
<td>0.000000566</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000419)</td>
<td>(0.00000394)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian population</td>
<td>0.00000780</td>
<td>0.00000494</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00000484)</td>
<td>(0.00000442)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflation adjusted income</td>
<td>5.60e-08</td>
<td>-7.26e-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000000158)</td>
<td>(0.000000186)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily traffic for Taco Cabana</td>
<td>0.00126***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000328)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In big urban area</td>
<td>0.0272</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0444)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interstate highway</td>
<td>0.0572</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0402)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary or secondary road</td>
<td>-0.0787</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0544)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.173***</td>
<td>0.173***</td>
<td>0.180***</td>
<td>0.155*</td>
</tr>
<tr>
<td></td>
<td>(0.0202)</td>
<td>(0.0202)</td>
<td>(0.0243)</td>
<td>(0.0645)</td>
</tr>
<tr>
<td>Observations</td>
<td>82150</td>
<td>82150</td>
<td>82150</td>
<td>82150</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1322</td>
<td>0.1322</td>
<td>0.1322</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
3.5.1 Estimation Procedure

We will follow the Bajari, Benkard, and Levin (2007) approach to estimate the model, which is a two-step method. The first stage of estimation is used to approximate the policy function, while the second stage of estimation is used to recover the underlying sunk cost parameters, \( \gamma = (EC_i, EC_j, SV_i, SV_j) \). As in Bajari, Benkard, and Levin (2007), we assume that the data observed are generated by a single MPE strategy profile, which then allows us to pool markets during estimation. As pointed out in Jeziorski (2014), this assumption is weaker than an equilibrium selection assumption that is often imposed in full-solution (i.e., nested fixed point) estimation methods.

3.5.1.1 First Stage Estimation

To approximate the policy function \( \sigma_i \), we employ a flexible probit model. That is, we use the following reduced form specification

\[
\tilde{\sigma}_i = \hat{P}_i^\sigma = \Phi(g(s, z)),
\]

(3.5.1)

where \( \Phi(\cdot) \) is the c.d.f. for a Normal distribution, and \( g(s, z) \) is a flexible link function of the market structure states \( s \) and exogenous market states \( z \).

Before pursuing the second stage of estimation, we also need to establish the relationship between observed sales, and the payoff relevant states \( (s, z) \). We use the following regression to map revenue (in markets that \( i \) is active in) onto the payoff relevant states associated with those markets

\[
R_i = \theta_{i1}\hat{\sigma}_j + \theta_{i2}z_i + \nu_i,
\]

(3.5.2)

where \( \nu_i \) is an i.i.d. shock. We allow for potential business stealing effects to operate through \( \theta_{i1} \).
With the revenue regression estimated, we can then obtain a predicted value for revenue, \( \hat{R}_i \), for any given state \( z \) and rival \( j \)'s decision \( a_j \).

### 3.5.1.2 Second Stage Estimation

With the approximated policy function and fitted revenue function in hand, we then proceed with the second step of estimation. For any given initial state \((X_1)\), we can then forward simulate the following for firm \( i \) in market \( m \):

\[
\bar{V}_{i,m}(X_1; \sigma, \gamma) = \mathbb{E} \left[ \sum_{\tau=1}^{\infty} \beta^{\tau-1} \Pi_{i,m}(\sigma(X_\tau)), X_\tau; \gamma \right | X_1, \sigma] \tag{3.5.3}
\]

\[
\approx \frac{1}{K} \sum_{k=1}^{K} \sum_{\tau=1}^{T} \beta^{\tau-1} \Pi_{i,m}(\sigma(X^k_\tau)), X^k_\tau; \gamma). \tag{3.5.4}
\]

Subscript \( k \) represents each forward simulation, where \( K \) paths of length \( T \) are simulated in the second stage. The term \( \sigma(X^k_\tau) \) denotes a vector of simulated actions based on the approximated policy profile \( \hat{\sigma}_i \) from the first stage estimation. Furthermore, we forward simulate the exogenous market states, \( z \), by using predicted values based on an autoregressive process we use to fit the evolution of \( z \).

With this construction of forward simulated actions and payoffs, we can then consider perturbations of the policy function to generate \( B \) alternative policies. With each alternative policy, we can obtain the forward simulated profit stream using the previous two steps. We let \( b \) index the individual inequalities, with each inequality consisting of an initial market structure and state \( X^b_1 = (s^b_1, z^b_1, \varepsilon^b_1) \), an index for the deviating firm \( i \), and an alternative policy \( \tilde{\sigma}_i \) for firm \( i \). The difference in valuations for firm \( i \) in market \( m \) using inequality \( b \) is denoted by

\[
h_{i,b,m}(\hat{\sigma}, \gamma) = \bar{V}_{i,m}(X^b_1; \hat{\sigma}, \gamma) - \bar{V}_{i,m}(X^b_1; \tilde{\sigma}_i, \hat{\sigma} - \hat{\sigma}_i, \gamma), \tag{3.5.5}
\]

This difference should be positive in equilibrium, since off-equilibrium values has to be lower than discounted profits under equilibrium play. Therefore, this criterion listed below identifies a \( \hat{\gamma} \) to minimize the violations of the equilibrium requirement:

\[
Q(\gamma) = \frac{1}{B} \sum_{m} \sum_{i} \sum_{b} (\min\{h_{i,b,m}(\hat{\sigma}, \gamma), 0\})^2. \tag{3.5.6}
\]

### 3.5.2 Model Identification

The model of entry we estimate is fairly standard, so we will provide a heuristic description about identification of our model. As per the discussion in Bajari, Chernozhukov, Hong, and Nekipelov (2009), nonparametric identification of a dynamic game requires firm-specific states that affect the payoffs \( \Pi_i(a_{it}, a_{jt}, s_i, z_i, \varepsilon_i) \). By construction, our model has such states. First, \( s_i \) has a direct impact on \( i \) via \( i \)'s entry cost and scrap value, while \( s_{i,t} \) has no direct impact on rival \( j \). Furthermore,
we also allow for potential exogenous revenue shifters, $z_{it}$, that uniquely affect $i$.

### 3.6 Empirical Results

In our second-stage estimation procedure, we make use of $K = 1000$ random inequalities. In each inequality, we compare the simulated value functions in equilibrium versus alternative strategies. Random perturbations of the approximated policy function from first-stage estimation are used to generate alternative strategies.

#### 3.6.1 Summary of Model Estimates

Table 3.13 contains the main estimates in the revenue and cost functions. As also reflected in the reduced form patterns of entry, we see that operating in big urban areas with highway/road infrastructure are beneficial for both retail chains. Interestingly, we find that the daily traffic counts affect Chipotle negatively while they affect Taco Cabana positively. Furthermore, the estimates confirm that rival presence will certainly lead to business-stealing effects, especially so for the incumbent Taco Cabana. This asymmetry in the magnitude of competition sensitivity is worth noting as it may affect the value of preempting markets differently among the two chains.

There also appear to be some asymmetries in terms of the type of markets the two chains do well in. For example, Taco Cabana appears to appeal more to the Hispanic population, while Chipotle appeals to all the ethnic groups except the Hispanic population. Also, both chains appear to fair poorly in high income areas, though Taco Cabana benefits more from low income areas than Chipotle. Finally, there is a monotonic relationship between age and sales for Taco Cabana, while Chipotle appeals to the youngest and oldest age groups.

The estimates also reveal the entry costs and scrap values for the two chains. Taco Cabana has noticeably higher entry costs than Chipotle. Also worth noting is that the scrap values have a positive sign, which suggests that the chains may be able to recover some of the sunk costs upon leaving a market, say by selling the real estate where their outlets are situated. The fact that the sunk costs are not fully recovered reiterate the fact that exit from a market is costly for both firms.

#### 3.6.2 Assessing the Aggression-Survival Trade-Off

The structural estimates can help us understand better the link between aggressive entry and survival. Recall from the earlier section about the theoretical framework behind aggression-survival, we illustrate that a trade-off will exist when the incumbent faces a strong potential entrant. In the event that a trade-off exists, then managers should account for potentially hidden future risks of aggressive entry, while the lack of a trade-off would suggest that managers should account for potentially hidden future benefits of aggressive entry.

Our results confirm a scenario in which the incumbent, Taco Cabana faces entry by a strong competitor, Chipotle. We infer that Chipotle is a strong competitor based on the following observations. Namely, Chipotle’s revenue seems to fall less to the presence of Taco Cabana than
## Table 3.13: Estimates of Revenue and Cost Function

<table>
<thead>
<tr>
<th></th>
<th>Chipotle</th>
<th></th>
<th>Taco Cabana</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
<td>Std. Error</td>
</tr>
<tr>
<td><strong>Revenue function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rival presence</td>
<td>-139.3824</td>
<td>1.4825</td>
<td>-935.6182</td>
<td>2.0045</td>
</tr>
<tr>
<td>Population density</td>
<td>-22.6639</td>
<td>0.0770</td>
<td>-52.4649</td>
<td>0.2420</td>
</tr>
<tr>
<td>Share population (15-34)</td>
<td>547.9075</td>
<td>1.5888</td>
<td>3019.5347</td>
<td>2.2572</td>
</tr>
<tr>
<td>Share population (35-65)</td>
<td>298.5320</td>
<td>1.6924</td>
<td>4339.2912</td>
<td>3.0134</td>
</tr>
<tr>
<td>Share population (65 and up)</td>
<td>802.5380</td>
<td>2.0074</td>
<td>5449.9861</td>
<td>4.3239</td>
</tr>
<tr>
<td>Share Hispanic population</td>
<td>-74.3329</td>
<td>0.3919</td>
<td>3561.1890</td>
<td>0.8115</td>
</tr>
<tr>
<td>Share white population</td>
<td>113.0916</td>
<td>0.2703</td>
<td>-526.3449</td>
<td>1.5871</td>
</tr>
<tr>
<td>Share black population</td>
<td>477.2624</td>
<td>0.5158</td>
<td>1121.0640</td>
<td>2.0654</td>
</tr>
<tr>
<td>Share asian population</td>
<td>263.4737</td>
<td>0.8195</td>
<td>-1490.0793</td>
<td>3.2202</td>
</tr>
<tr>
<td>Inflation adjusted income</td>
<td>-2.4975</td>
<td>0.1068</td>
<td>-51.4154</td>
<td>0.4372</td>
</tr>
<tr>
<td>In big urban area</td>
<td>205.7293</td>
<td>0.2141</td>
<td>29.1413</td>
<td>0.9698</td>
</tr>
<tr>
<td>Interstate highway</td>
<td>49.1806</td>
<td>0.1890</td>
<td>345.1453</td>
<td>0.5041</td>
</tr>
<tr>
<td>Primary or secondary road</td>
<td>85.0734</td>
<td>0.1115</td>
<td>-290.8736</td>
<td>0.8493</td>
</tr>
<tr>
<td>Daily traffic</td>
<td>-4.3237</td>
<td>0.0259</td>
<td>35.0658</td>
<td>0.1060</td>
</tr>
<tr>
<td><strong>Cost function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry cost</td>
<td>-5609.1117</td>
<td>8.6219</td>
<td>-33387.0907</td>
<td>1845.6809</td>
</tr>
<tr>
<td>Scrap value</td>
<td>5437.8742</td>
<td>8.7566</td>
<td>33345.8660</td>
<td>4035434.0298</td>
</tr>
</tbody>
</table>
Taco Cabana’s fall in revenue under the presence of Chipotle. A strong entrant suggests that the fast casual restaurant industry is closer to Scenario #2 in our numerical analysis. Recall that in this scenario, incumbent survival suffers from an aggressive entry strategy and would be better off following an accommodative strategy from the beginning.

3.7 Conclusion

Under the context of retail chain entry, our research provides analytical and empirical insights about the relationship between aggressive entry and incumbent survival. We first present the analytical framework of dynamic entry for studying this relationship. With this model in place, we highlight via numerical analysis that whether or not there is a trade-off between aggressiveness and survival depends heavily on the strength of the entrant. That is, there is a positive relationship between aggressiveness and survival when the entrant is not too strong, and a negative relationship when the entrant is too strong. In our empirical application that uses data from the fast casual restaurant industry, we infer that there will likely be a negative relationship between aggressiveness and survival, as the entrant (i.e., Chipotle) appears to be strong given that it has a stronger business-stealing effect on the incumbent (i.e., Taco Cabana) than the business-stealing effect it suffers from in the presence of its rival. Under this scenario, our analytical insights would suggest that a more accommodative strategy may be ideal if \textit{ex post} survival is an objective of the incumbent. Ultimately, we hope that our analytical and empirical insights can help managers select an equilibrium strategy (among multiple possibilities).

Our analytical findings could be applied to other retail settings, such as coffee stores, convenience stores, or fast food. Provided that the retail sector is concentrated, not too differentiated across firms, and relies on brick-and-mortar store expansion for revenue growth, we believe our conditions for when the aggression-survival trade-off is most relevant can make potential risks associated with aggressive entry more salient to retail managers. For example, one may conjecture that preemption is less risky of a strategy for the incumbent of Canada’s fast food industry (i.e., McDonald’s), as the potential entrants do not appear to be stronger in comparison with the incumbent (Igami and Yang, 2016).

The current paper abstracts away from other potential risks associated with aggressive (and early) entry. For example, aggressive entry may introduce the following risks as well, above and beyond a failure to deter entry. First, channel conflicts between franchisors and franchisees are more likely to occur (Blair and Lafontaine, 2005), as strategies aimed to saturate local markets will likely lead to own-brand cannibalization (e.g., Kalnins, 2004; Pancras, Sriram and Kumar, 2012). Second, aggressive entry may preclude a firm from observational learning about the market’s potential from past entry decisions if there is uncertainty about market size (e.g., Shen and Xiao, 2014; Yang, 2016). Finally, aggressive entry leads to greater antitrust scrutiny if such strategies lead to monopolistic markets (e.g., Schmalensee, 1978). Taken together, we believe future work could investigate how the aggression-survival relationship compares with these risks.
Bibliography


Blair, Roger and Francine Lafontaine (2005), The Economics of Franchising. Cambridge University Press.


