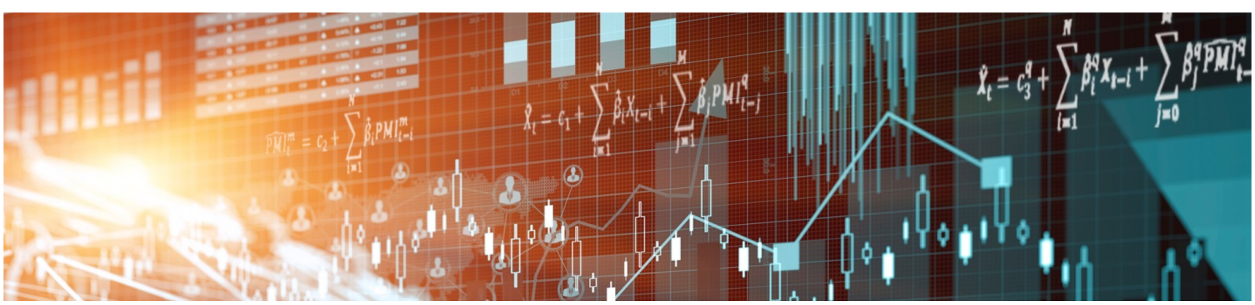
BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2018-60

Inference in Games Without Nash Equilibrium: An Application to Restaurants' Competition in Opening Hours



by Erhao Xie

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2018-60

December 2018

Inference in Games Without Nash Equilibrium: An Application to Restaurants' Competition in Opening Hours

by

Erhao Xie

Funds Management and Banking Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
exie@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

I am indebted to my supervisor, Victor Aguirregabiria, and thesis committee members, Yao Luo and Ryan Web, for their invaluable supervision and continuous encouragement. I thank Andres Aradillas-Lopez, Heski Bar-Isaac, Matthew Mitchell, Lorenzo Magnolfi, Ismael Mourifié, Eduardo Souza-Rodrigues, Colin Stewart, Yuanyuan Wan, Ping Xiao, and seminar participants at the Bank of Canada, Concordia, McGill, Toronto, Tsinghua, Victoria, CEA Annual Conference 2017, Doctoral Workshop for Applied Econometrics 2016, EARIE 2017, IAEE 2018, IIOC 2018, Jornadas de Economia Industrial 2017, and Young Economist Symposium (Yale, 2017) for constructive comments. I also thank Qinggang Meng, a provincial manager of McDonald's in China, who provided detailed industrial insights.

Abstract

This paper relaxes the Bayesian Nash equilibrium (BNE) assumption commonly imposed in empirical discrete choice games with incomplete information. Instead of assuming that players have unbiased/correct expectations, my model treats a player's belief about the behavior of other players as an unrestricted unknown function. I study the joint identification of belief and payoff functions. I show that in games where one player has more actions than the other player, the payoff function is partially identified with neither equilibrium restrictions nor the usual exclusion restrictions. Furthermore, if the cardinality of players' action sets varies across games, then the payoff and belief functions are point identified up to scale normalizations and the restriction of equilibrium beliefs is testable. For games where action sets are constant across players and observations, I obtain very similar identification results without imposing restrictions on beliefs, as long as the payoff function satisfies a condition of multiplicative separability. I apply this model and its identification results to study the store hours competition between McDonald's and Kentucky Fried Chicken (KFC) in China. The null hypothesis that KFC has unbiased beliefs is rejected. Failing to account for KFC's biased beliefs generates an attenuation bias on estimated strategic effects. Finally, the estimation results of the payoff functions indicate that the decision about store hours is a type of vertical differentiation. By operating through the night, a firm not only attracts night-time consumers but also can steal competitors' day-time customers. This result has implications on the optimal regulation of stores' opening hours.

Bank topics: Econometric and statistical methods; Market structure and pricing
JEL codes: C57, L13, L85

Résumé

Cette étude lève le postulat d'équilibre de Nash bayésien habituellement imposé dans les jeux empiriques avec choix discrets en information incomplète. Au lieu de supposer que les joueurs ont des attentes non biaisées (parfaites), mon modèle traite les croyances d'un joueur à l'égard du comportement des autres joueurs comme une fonction inconnue sans contrainte. Je cherche à identifier conjointement les fonctions de croyances et de gains. Je montre que les jeux dans lesquels un joueur a davantage de possibilités d'action que l'autre, la fonction de gains est partiellement identifiée sans introduire d'hypothèses d'équilibre ou les restrictions habituelles d'exclusion. De plus, si la cardinalité des ensembles d'actions des joueurs varie d'un jeu à l'autre, les fonctions de gains et de croyances sont alors identifiées ponctuellement jusqu'à une normalisation scalaire et l'hypothèse des équilibres des croyances est vérifiable par des tests. Pour les jeux dans lesquels les ensembles d'actions sont constants d'un joueur et d'une observation à l'autre, j'obtiens des résultats d'identification très semblables sans imposer de restrictions aux croyances tant que la fonction de gains satisfait à la condition de séparabilité multiplicative. J'applique ce modèle et les résultats d'identification obtenus à l'étude de la concurrence au chapitre des heures d'ouverture entre McDonald's et Kentucky Fried Chicken (KFC) en Chine. L'hypothèse nulle selon laquelle les croyances de KFC ne sont pas biaisées est rejetée. La non-prise en compte des croyances biaisées de KFC génère un biais d'atténuation des effets stratégiques estimés. Enfin, les résultats de l'estimation des fonctions de gains montrent que le choix des heures d'ouverture des établissements constitue une forme de différenciation verticale. En restant ouverte la nuit, une entreprise peut non seulement attirer une clientèle nocturne, mais aussi ravir aux entreprises concurrentes des clients qui fréquentent leurs établissements le jour. Ce résultat a des conséquences pour l'adoption d'une réglementation optimale des heures d'ouverture des commerces.

Sujets : Méthodes économétriques et statistiques; Structure de marché et fixation des prix
Codes JEL : C57, L13, L85

Non-Technical Summary

Game theory is one of the central topics in economics. It studies the situation where an individual's payoff depends not only on his/her own actions but also on other individuals' choices. This dependence is very common in the real world. Competition among firms, social behaviors of people and interactions of financial intermediations are all examples of game theory. In studies of game theory, researchers normally rely on the concept of Nash equilibrium. It assumes that each player of the game has a correct belief about other players' behaviors and chooses an optimal action accordingly. However, the game is often complex and changing. This makes it substantially difficult to prove that players can correctly predict the behaviors of other players. If players have incorrect beliefs, relying on Nash equilibrium would lead to incorrect conclusions. For instance, researchers will falsely quantify how a player's decision depends on other players' choices. This mistake could lead to an incorrect prediction on players' choices after a policy intervention.

In this paper, I develop a method to address the potential failure of Nash equilibrium. First, I show that the assumption of correct beliefs is testable. Second, when such an assumption is rejected, I further show that researchers can still correctly infer players' payoffs. The implementation of above test and inference can be done using standard estimation techniques.

Applying the above method, I study the competition between McDonald's and Kentucky Fried Chicken (KFC) restaurants in China. Specifically, I focus on their competition for business hours. My results show that KFC under-predicts McDonald's business hours in areas where consumers have high income. It therefore rejects the null hypothesis that KFC has correct beliefs. Moreover, KFC tends to expand its business hours when it expects McDonald's to do so. Such a dependence becomes insignificant when researchers incorrectly assume Nash equilibrium. These results highlight the importance of relaxing Nash equilibrium in many empirical applications.

1 Introduction

Over the past decade, game theoretic models with incomplete information have been actively applied to study oligopolistic competition and individuals' social interactive behaviors.[1] In this stream of literature, researchers commonly assume that players' observed choices are consistent with Bayesian Nash Equilibrium (BNE). Under this powerful solution concept, researchers then estimate players' utility/payoff functions and predict their behaviors in counterfactual environments.

Despite its power and usefulness in applied empirical work, BNE places a strong restriction on players' expectations such that each player has equilibrium/unbiased beliefs about other players' behaviors (i.e. a player's beliefs are other players' *actual* choice probabilities given the available information). In reality, economic agents could have limited ability to process information and predict other players' strategies. In addition, many empirical games have multiple equilibria, which further complicates the construction of unbiased belief. In these games, a player could be uncertain about which equilibrium strategy is chosen by other players.[2] Furthermore, market conditions and government policies often vary dramatically. This poses difficulties in learning other players' behaviors through past experience. Finally, recent empirical work has shown the failure of Nash Equilibrium in different types of games using both field and experimental data. A partial list includes Goeree and Holt (2001), Goldfarb and Xiao (2011), Asker et al. (2016), Doraszelski et al. (2018), Kashaev (2016), Aguirregabiria and Magesan (2017), Aguirregabiria and Xie (2017), and Jeon (2018).

If players have biased beliefs in games, falsely imposing the equilibrium condition would bias the estimates of payoff functions and counterfactual predictions. To address this issue, this paper relaxes the unbiased belief assumption. Specifically, each player maximizes her expected utility given her subjective belief, which can be any probability distribution over the other player's action set. This nests BNE as a special case when each player forms an equilibrium belief. It also allows players' behaviors to be off-

[1]A partial list of work includes firm entry studied by Gowrisankaran and Krainer (2011) and Aradillas-Lopez and Gandhi (2016); product differentiation studied by Seim (2006), Augereau et al. (2006), and Sweeting (2009); social interactive effect analyzed by Brock and Durlauf (2001, 2007) and Bajari et al. (2010); network structure by Vitorino (2009) and methodological contributions by Aradillas-Lopez (2010, 2012).

[2]This type of strategic uncertainty is defined by Van Huyck et al. (1990) and Crawford and Haller (1990), and studied by Morris and Shin (2002, 2004), and Heinemann et al. (2009), among others. As argued by Besanko et al. (2010), it could be very common in oligopolistic competition. Moreover, the existence of multiple equilibria can facilitate the identification of players' payoffs as shown by Sweeting (2009), De Paula and Tang (2012), and Aradillas-Lopez and Gandhi (2016); also see Aguirregabiria and Mira (2018) and Xiao (2018).

equilibrium due to biased beliefs. In the estimation of the model, both utility/payoff and belief are treated as ***unknown unrestricted non-parametric*** functions, in contrast to the standard approach in the literature of imposing equilibrium restrictions.

The principle of revealed preference implies that, under general conditions, researchers can infer the expected utility function using data on players' choices. However, since expected utility is a composite function of the utility and belief, it is challenging to separately identify these two functions. In this paper, I show that if the cardinality of the action set (i.e. the number of choices a player has) varies across players, the payoff function is partially identified without imposing equilibrium restrictions. I then characterize the identified set. There are many examples and applications of games where players have a different number of possible choices. For instance, in models of competition in multi-product or multi-store markets (e.g. in price, quantity or quality choice), firms with different numbers of products or stores face a different number of actions when choice is at the product or store level.

Suppose further that the cardinality of each player's action set has variation across games (e.g. a retailer has a different number of stores in different markets). Then, the base return (e.g. monopoly profit in an entry game) is identified, and each player's interactive effect is identified up to her belief at ***only one*** realization of the state variable. It consequently identifies the sign and a lower bound of the interactive effect. These results shed light on the nature of the game (e.g. strategic substitutes and strategic complements). Furthermore, researchers can infer how a player adjusts her beliefs across different games. It naturally yields a testable restriction of unbiased beliefs and provides information on how players form expectations in games.

In some empirical applications, a player's choice set is endogenously determined, causing the standard selection problem. I show how to apply recent techniques in finite mixture literature to deal with such a problem. Moreover, in many other empirical applications, each player has the same number of actions and the cardinality of the action set remains constant across observations. In these applications, I show that when the payoff function satisfies a condition of multiplicative separability, researchers can establish a similar asymmetric feature as in the games with asymmetric action sets. As a result, the payoff function is also partially identified. Additionally, I show that usual exclusion restrictions can further shrink the identified set of payoff functions. Therefore, the identification results studied in this paper apply to a broad class of games. Finally, all proofs are constructive and naturally imply a two-step estimator.

I apply these identification and estimation results to study the decisions about operating hours of Kentucky Fried Chicken (KFC) and McDonald's (McD) in China. Their competition is modeled as an incomplete information game such that each chain simultaneously chooses how many of its existing outlets to open at night. Consequently, different numbers of outlets owned by the two chains in various markets naturally construct the asymmetry and variation of action sets required to identify both payoff and belief functions. Without imposing the BNE assumption, I find that KFC tends to extend its store hours when it expects McD to do so. This provides the first empirical evidence that suggests store hours are strategic complements. This result also contributes to the theoretical literature by Ferris (1990) and Klemperer and Padilla (1997), who view store hours as strategic complements, and Inderst and Irmen (2005), Shy and Stenbacka (2006, 2008), and Wenzel (2010, 2011), who treat them as strategic substitutes. As described in Section 4, distinguishing such a strategic nature is the key to evaluating the deregulation policy that lifts restrictions on store hours in many industries. Finally, the estimation results reject the null hypothesis that KFC has unbiased beliefs. Incorrectly imposing the equilibrium condition generates an attenuation bias on the strategic effects.

In literature that studies empirical games, the identification of players' payoff functions typically relies on an exclusion restriction that affects one player's payoff without affecting other players. This paper shows that, even in the absence of this usual exclusion restriction, the variation of players' action sets also provides identification power.[3] This result is similar to Orhun (2013), who shows that the variation of potential entrants can identify an entry game under equilibrium assumption. In this paper, I further investigate the identification results without equilibrium constraint. Consequently, it also contributes to recent literature on players' non-equilibrium behaviors in games. Some important contributions are Aradillas-Lopez and Tamer (2008), Goldfarb and Xiao (2011), Fershtman and Pakes (2012), Kline and Tamer (2012), Uetake and Watanabe (2013), An (2017), Gillen (2010), Asker et al. (2016), Doraszelski et al. (2018), and Kashaev (2016). In this literature, my paper is closely related to Aguirregabiria and Magesan (2017) and Aguirregabiria and Xie (2017), who show that the equilibrium assumption is testable with the usual exclusion restriction. However, to identify the payoff, they need to assume that players

[3] Admittedly, the cardinality of a player's action set is also an exclusion restriction. It affects each player's belief without affecting ***any*** player's payoff. However, to compare with existing literature and avoid confusion, the action set and its cardinality are not referred to as exclusion restrictions in this paper.

form equilibrium beliefs at several realizations of the exclusion restriction. In contrast, this paper exploits the asymmetry between the cardinality of players' action sets and potential variation in cardinality across observations. Through this, the identification of payoff and belief is achieved without any restrictions on players' beliefs. Importantly, such a result does not rely on the existence of the usual exclusion restrictions.

The rest of this paper is organized as follows. Section 2 describes the model, and Section 3 presents the identification results. The empirical application is shown in Section 4. I conclude in Section 5. Some generalizations of the model and identification results are left to the Appendix.

2 Model

Consider a two-player static game. Players are indexed by $i \in \{1,2\}$, and $-i$ represents the other player. Appendix A.4 shows how to generalize the identification results to a game with multiple players. Each player i simultaneously chooses an action denoted by a_i from her action set $A_i = \{0,1,\cdots,J_i\}$. Players are allowed to have different action sets and different numbers of actions. The Cartesian product $A = A_1 \times A_2$ represents the space of action profiles in this game. Let $\mathbf{a} = (a_1,a_2) \in A$ be an action profile or realized outcome of this game. Player i's payoff for the action profile $\mathbf{a}$ is

$$\Pi_i(\mathbf{x},\varepsilon_i,\mathbf{a}) = \tilde{\Pi}_i(\mathbf{x},a_i,a_{-i}) + \varepsilon_i(a_i),$$

where $\mathbf{x} \in \mathbb{R}^{L_x}$ denotes a vector of state variables that affect players' payoffs and is public information. The term $\varepsilon_i(a_i)$ represents a variable that affects player i's payoff of action a_i. It is private information observed only by player i and unobserved by player $-i$. Therefore, it is a game of incomplete information. The payoffs $\tilde{\Pi}_i(\mathbf{x},a_i,a_{-i})$ are non-parametrically specified.

Define $\pi_i(\mathbf{x},a_i) = \tilde{\Pi}_i(\mathbf{x},a_i,a_{-i}=0)$ and $\delta_i(\mathbf{x},a_i,a_{-i}) = \tilde{\Pi}_i(\mathbf{x},a_i,a_{-i}) - \tilde{\Pi}_i(\mathbf{x},a_i,a_{-i}=0)$. By construction, $\delta_i(\mathbf{x},a_i,a_{-i}=0)=0$. Without loss of generality, the payoff function can be written as

$$\Pi_i(\mathbf{x},\varepsilon_i,\mathbf{a}) = \pi_i(\mathbf{x},a_i) + \delta_i(\mathbf{x},a_i,a_{-i}) \cdot \mathbb{1}(a_{-i} \neq 0) + \varepsilon_i(a_i). \tag{1}$$

Throughout the paper, I consider the payoff function specified by equation (1) for exposition purposes. Note that it is non-parametrically specified. Following the language of De Paula and Tang (2012), $\pi_i(\cdot)$ is referred to as the base return, and it represents player i's payoff when the other player chooses action 0. Term $\delta_i(\cdot)$ is referred to as the interactive effect/payoff, and it measures how player i's payoff is affected by player $-i$'s behavior.

Assumption 1 states an independence restriction imposed on each player's private information.

Assumption 1. *(a) For each $i = 1,2$, $\varepsilon_i = (\varepsilon_i(0), \varepsilon_i(1), \cdots, \varepsilon_i(J_i))'$ follows a CDF $G_i(\cdot)$ that is absolutely continuous with respect to Lebesgue measure in $\mathbb{R}^{J_i+1}$.*

(b) ε_i is independent across players and independent of public information $\mathbf{x}$.

Unlike existing literature in games with incomplete information, Assumption 1 does not require $G_i(\cdot)$ to be common information among players. Suppose a player does not know the true distribution of other players' preference shocks: she would form a biased belief about the distribution of other players' behaviors even though she can correctly solve other players' optimal strategies. Such a source of biased belief is allowed in Assumption 1. Moreover, the model and identification results are generalizable to the case that $G_i(\cdot)$ depends on a vector of finite-dimensional unknown parameters. This extended model and its identification results are presented in Appendix A.2.[4]

Assumption 2. *(a) Each player's belief about the other player's behavior depends only on public information* $\mathbf{x}$.

(b) Each player chooses an action that maximizes her expected payoff given her belief.

Assumption 2 (a) assumes that player i's belief about the other player's behavior does not depend on her private information. Given Assumption 1 (b) such that players have independent private information, ε_i has no predictive power about player $-i$'s payoff and behavior; consequently, ε_i does not affect player i's belief. Specifically, define $b_i^j(\mathbf{x})$ as player i's belief about the probability that player $-i$ will choose action j. Moreover, let $\mathbf{b}_i(\mathbf{x}) = (b_i^0(\mathbf{x}), \cdots, b_i^{J_{-i}}(\mathbf{x}))'$ be a vector of belief functions that are

[4] Aradillas-Lopez (2010), Wan and Xu (2014), and Xu (2014) extend the independence assumption of ε_i and allow it to be correlated across players even conditional on observed state variables. Despite their power, their applicability is very limited in my framework. Note that player i's belief will depend on ε_i if it is correlated with ε_{-i}. From an econometrician's perspective, the belief is an unknown function that depends on an unknown variable ε_i. With the BNE assumption, we can solve for this belief using a fixed-point algorithm; however, it will not work if BNE fails, as is allowed in my framework.

non-parametrically specified. The only assumption I impose is that the belief functions must be valid probability distributions over the other player's action set (i.e. $0 \leq b_i^j(\mathbf{x}) \leq 1 \; \forall \; j$ and $\sum_{j=0}^{J_{-i}} b_i^j(\mathbf{x}) = 1$).

Given payoff and belief functions, player i's expected payoff of action a_i is

$$E\left[\Pi_i(\mathbf{x},\varepsilon_i,a_i)\right] = \pi_i(\mathbf{x},a_i) + \sum_{j=1}^{J_{-i}} \delta_i(\mathbf{x},a_i,a_{-i}=j)\cdot b_i^j(\mathbf{x}) + \varepsilon_i(a_i). \tag{2}$$

Assumption 2 (b) states that player i chooses an action that maximizes this expected payoff. Define $a_i^*(\mathbf{x},\varepsilon_i)$ as player i's strategy function, which can be characterized as

$$a_i^*(\mathbf{x},\varepsilon_i) = \underset{a_i\in A_i}{\text{argmax}} \left\{\pi_i(\mathbf{x},a_i) + \sum_{j=1}^{J_{-i}} \delta_i(\mathbf{x},a_i,a_{-i}=j)\cdot b_i^j(\mathbf{x}) + \varepsilon_i(a_i)\right\}. \tag{3}$$

As player $-i$ does not observe ε_i, her unbiased expectation of player i's behavior is player i's best response probability function or *conditional choice probability* (CCP). Let $\mathbf{p}_i(\mathbf{x}) = \left(p_i^0(\mathbf{x}),\cdots,p_i^{J_i}(\mathbf{x})\right)'$ denote a vector of player i's CCPs, where $p_i^j(\mathbf{x})$ is her choice probability of action j conditional on state variable $\mathbf{x}$. Given the best response function $a_i^*(\mathbf{x},\varepsilon_i)$ defined above, the conditional choice probability takes the following form:

$$p_i^j(\mathbf{x}) = \int \mathbb{1}\left\{a_i^*(\mathbf{x},\varepsilon_i) = j\right\} dG_i(\varepsilon_i). \tag{4}$$

For instance, if $\varepsilon_i(a_i)$ is type 1 extreme value distributed and independent across actions, the conditional choice probability $p_i^k(\mathbf{x})$ is

$$p_i^k(\mathbf{x}) = \frac{\exp\left\{\pi_i(\mathbf{x},a_i=k) + \sum_{j=1}^{J_{-i}} \delta_i(\mathbf{x},a_i=k,a_{-i}=j)\cdot b_i^j(\mathbf{x})\right\}}{\sum_{l=0}^{J_i}\exp\left\{\pi_i(\mathbf{x},a_i=l) + \sum_{j=1}^{J_{-i}} \delta_i(\mathbf{x},a_i=l,a_{-i}=j)\cdot b_i^j(\mathbf{x})\right\}}.$$

The conditional choice probability defined by equation (4) only assumes that a player maximizes expected payoff given her belief. Such a belief is allowed to be any probability distribution over the other player's action set. In contrast, Bayesian Nash Equilibrium restricts players to be perfectly rational in the sense that a player's belief is the other player's true choice probability conditional on the available information. My framework therefore nests BNE as a special case and is summarized by Definition 1.

Definition 1. *Players' behaviors are consistent with Bayesian Nash Equilibrium if each player's belief is* *the other player's actual conditional choice probability, i.e.* $b_i^j(\mathbf{x}) = p_{-i}^j(\mathbf{x}) \;\forall\; 0 \leq j \leq J_i$ *and* $i = 1, 2$.

Remark. *Some identification results in this paper exploit variation in players' action sets across observations. To understand those results, it is important to keep in mind that player i's choice probabilities and beliefs depend on* J_1 *and* J_2 *as they affect the dimensions of vectors* $\mathbf{p}_i$ *and* $\mathbf{b}_i$*. A notation for choice probabilities and beliefs that emphasizes this dependence is* $\mathbf{p}_i(\mathbf{x}, J_i, J_{-i})$ *and* $\mathbf{b}_i(\mathbf{x}, J_i, J_{-i})$*. For the sake of notation simplicity, I maintain the expression of* $\mathbf{p}_i(\mathbf{x})$ *and* $\mathbf{b}_i(\mathbf{x})$ *and only include* J_i *and* J_{-i} *as explicit arguments when necessary.*

Example 1. Operating hours game in the empirical application in Section 4.

There are two fast food chains, KFC and McD, competing through decisions on business hours. Suppose chain i owns J_i *number of outlets in a given market and chooses some number of its existing stores to operate during the night. Therefore, the competition in business hours can be seen as an entry game such that each chain simultaneously chooses how many stores to enter into the night market. Moreover, two chains could own different numbers of outlets and consequently have heterogeneous action sets. Term* $\pi_i(\mathbf{x}, a_i = k)$ *represents chain i's profit of opening k non-stop service stores if the other chain closes all stores at night: for instance, monopoly profit in the night market. Vector* $\mathbf{x}$ *represents variables that affect each chain's profit, such as income per capita and population at the local market. When chain* $-i$ *decides to operate j stores at night, it will have some impact on chain i's night profit; this is captured by* $\delta_i(\mathbf{x}, a_i = k, a_{-i} = j)$*. Finally,* $\varepsilon_i(a_i)$ *represents chain i's private information of its own profitability, such as managerial skill and staff's coordination efficiency.*

3 Identification

In this section, I first present conditions on the data generating process. In Subsection 3.2, I show how to exploit the asymmetry and variation of action sets to identify each player's payoff and belief functions, without requiring the usual exclusion restrictions. Subsection 3.3 considers another type of game in which players have the same action set that remains constant across observations, but the payoff function satisfies a multiplicative separability condition. I show that a similar asymmetric feature can be constructed in this type of game such that the payoff function is partially identified. Furthermore, the introduction of usual exclusion restrictions provides additional identification power and sharpens the identified set. Therefore, my identification results hold in a broad class of games. Finally, Subsection 3.4 illustrates how to apply recent techniques in finite mixture literature to deal with the endogeneity of players' action sets. All proofs are left to the Appendix.

3.1 Conditions on the Data Generating Process

Suppose researchers have access to a data set about the same two players that play M independent games (e.g. one game in each of M isolated markets). In each game/observation indexed by m, both players and the econometrician observe realizations of the state variables $\mathbf{x}_m$. Moreover, each player i observes her own payoff shock $\varepsilon_{i,m}$. Researchers cannot observe player i's private information but know its probability distribution $G_i(\cdot)$. Given the observed state variables, each player forms a belief and chooses her optimal action.

Assumption 3. *A player forms the same beliefs for any two observations with the same public information. That is, for $m \neq m'$ but $\mathbf{x}_m = \mathbf{x}_{m'} = \mathbf{x}$, we have $\mathbf{b}_{i,m}(\mathbf{x}) = \mathbf{b}_{i,m'}(\mathbf{x}) = \mathbf{b}_i(\mathbf{x})$.*

Assumption 3 states that each player has a unique belief conditional on public information. In models that impose equilibrium restrictions, an analogous assumption would be that each player employs the same equilibrium strategy when multiple equilibria exist. Therefore, even though the model allows for multiple equilibria, there is a unique equilibrium observed in the data. Such an assumption is commonly imposed in the literature.[5] As discussed in the Conclusion, the identification results by Aguirregabiria and Mira (2018) can be applied to relax this assumption of belief uniqueness.

The asymptotic consistency comes from M going to infinity. In this situation, $\hat{\mathbf{p}}_i(\mathbf{x}_m)$ is consistently estimated.[6] Consequently, for identification results, $\mathbf{p}_i(\mathbf{x})$ is assumed to be known by researchers for every realization of $\mathbf{x}$. Researchers' objective is to identify player i's base return $\pi_i(\mathbf{x}, a_i)$, interactive effect $\delta_i(\mathbf{x}, a_i, a_{-i})$ and belief function $\mathbf{b}_i(\mathbf{x})$ using the data described above.

[5] A partial list includes Aguirregabiria and Mira (2007), Bajari et al. (2007), Pakes et al. (2007), Pesendorfer and Schmidt-Dengler (2008), Bajari et al. (2010), and Aradillas-Lopez (2012).

[6] If $\mathbf{x}$ were discrete variables with finite support, or continuous variables with a smooth choice probability function, then the choice probability can be consistently estimated by the standard kernel estimator. If $\mathbf{x}$ were continuous variables and $\mathbf{p}_i(\mathbf{x})$ had some points of discontinuity that are not known by researchers ex ante, then some variants of the standard kernel method developed by Müller (1992) and Delgado and Hidalgo (2000) can still establish the consistency and asymptotic normality for the choice probability estimator.

It is known in the discrete choice literature that only differences in payoffs are identified. Therefore, a normalization is required to achieve identification, as summarized in Assumption 4.

Assumption 4. *For player $i = 1,2$, the payoff for action* 0 *is normalized to zero. That is,* $\pi_i(\mathbf{x}, a_i = 0) = 0$ *and* $\delta_i(\mathbf{x}, a_i = 0, a_{-i}) = 0 \; \forall \; \mathbf{x}, a_{-i}$.

Even though normalizing the payoff of one action to zero is an innocuous assumption in single-agent discrete choice models, it imposes some restrictions on a player's payoff when players interact with each other. Specifically, this normalization restricts player i's payoff of action 0 to be unaffected by the other player's action. Such normalization is plausible if action 0 is modeled as an outside option. For instance, it represents that firm i does not enter a particular market in a standard entry game. Consequently, its profit is independent of other firms' behaviors in such a market.[7]

I investigate identification *conditional* on any $\mathbf{x} \in \mathbb{R}^{L_x}$, though it is suppressed for notation simplicity throughout this section. Following Hotz and Miller (1993), Assumptions 1 and 4 imply that there is a one-to-one mapping $F_i(\cdot) : \mathbb{R}^{J_i+1} \rightarrow \mathbb{R}^{J_i+1}$ from player i's conditional choice probability to her expected payoff. Specifically, let $F_i(\cdot)$ be the inverse of the integral function defined by equation (4), so we then have

$$\pi_i(a_i = k) + \sum_{j=1}^{J_{-i}} \delta_i(a_i = k, a_{-i} = j) \cdot b_i^j = F_i^k(\mathbf{p}_i) \;\; \forall \, 0 \leq k \leq J_i, \tag{5}$$

where $F_i^k(\cdot)$ denotes the k^{th} element of the inverse function. Note that $F_i^0(\mathbf{p}_i) = 0$ based on the normalization stated in Assumption 4. Given that $G_i(\cdot)$ is known by researchers, $F_i(\cdot)$ is also known. For instance, if $\varepsilon_i(a_i)$ is independently type 1 extreme value distributed, we have the mapping $F_i^k(\mathbf{p}_i) = \log\big(\frac{p_i^k}{p_i^0}\big)$.

In empirical games with incomplete information, researchers commonly assume that players play an equilibrium strategy and the conditions described in Definition 1 hold. Consequently, we can replace b_i^j with its counterpart p_{-i}^j in equation (5). As p_{-i}^j can be consistently estimated, this equilibrium assumption provides identification power.[8]

For the rest of this paper, I drop the BNE conditions as described in Definition 1. In the econometric model, each player is an expected utility maximizer, and her belief is allowed to be any probability

[7] In situations where it is implausible to assume player i's payoff of $a_i = 0$ is invariant to the other player's choices, all the identification results still hold if researchers' interest is the difference of payoff rather than payoff. For instance, the interested parameters are $\tilde{\pi}_i(\mathbf{x}, a_i) = \pi_i(\mathbf{x}, a_i) - \pi_i(\mathbf{x}, a_i = 0)$ and $\tilde{\delta}_i(\mathbf{x}, a_i, a_{-i}) = \delta_i(\mathbf{x}, a_i, a_{-i}) - \delta(\mathbf{x}, a_i = 0, a_{-i})$.

[8] For estimation techniques, see Su (2014) for an excellent discussion and comparison of different methods.

distribution over the other player's action set. Under this framework, I investigate the identification of payoffs and beliefs.

3.2 Identification with Asymmetric Action Spaces

This subsection exploits the asymmetry and variation in players' action sets to identify each player's payoffs and beliefs. Let $\mathcal{J} = \{0, 1, \cdots \bar{J}\}$ denote the support of J_i for each player i. In Example 1, each chain could have a different number of actions that varies across markets.

Intuitively, observational data reveal player i's expected payoff of every action since p_i^j is consistently estimated. These J_i restrictions depend on J_{-i} unknown belief parameters (i.e. the number of player $-i$'s actions minus one). When $J_i > J_{-i}$, I show that it is possible to obtain a transformation of equation (5) such that beliefs are differenced out and what remains is a relationship between unknown payoffs and known choice probabilities. Such a relationship then characterizes the identified set for payoff functions. Furthermore, if J_{-i} varies across the M markets and takes on a value of zero with positive probability, then we would have a player i's single agent problem and her base return $\pi_i(\cdot)$ would be point identified. Finally, comparing markets with $J_{-i} = 0$ with ones where $J_{-i} > 0$ provides information on player i's interactive payoffs and beliefs.

Proposition 1. *(a) Suppose Assumptions 1 to 4 hold and the data contain observations with $J_i > J_{-i} = 1$, then for any two choice alternatives j and k, the identified set of player i's payoff parameters $\pi_i(\cdot)$ and $\delta_i(\cdot)$ is given by the set of values that satisfies the following restriction:*

$$\frac{F_i^j(\mathbf{p}_i) - \pi_i(a_i = j)}{F_i^k(\mathbf{p}_i) - \pi_i(a_i = k)} = \frac{\delta_i(a_i = j, a_{-i} = 1)}{\delta_i(a_i = k, a_{-i} = 1)}.$$

(b) Further, suppose that the data also contain observations with $J_{-i} = 0$, then player i's base return $\pi_i(a_i = k)$ is point identified as $F_i^k[\mathbf{p}_i(J_{-i} = 0)]\ \forall\, k$. Furthermore, the identified set of player i's interactive effect and belief is given by the set of values that satisfies the following restriction:

$$\delta_i(a_i = k, a_{-i} = 1) b_i^1 = F_i^k[\mathbf{p}_i(J_{-i} = 1)] - F_i^k[\mathbf{p}_i(J_{-i} = 0)]\ \forall\, 0 \le k \le J_i.$$

Proposition 1 (a) characterizes the identified set of payoff functions when player i has more actions than the other player. This result holds true even when each player's action set is fixed across observations. Furthermore, Proposition 1 (b) states that the base return $\pi_i(a_i)$ is point identified when players' action spaces vary across games. In Example 1, this term represents chain i's monopoly profit in the night market because day-time profits are normalized to zero. This paper also characterizes the identified set of player i's interactive effect and beliefs. In this subsection, I focus on the case where J_{-i} takes on values of zero or one. The results for $J_{-i} > 1$ are very similar and presented in appendix A.3.

Since $0 \leq b_i^1 \leq 1$, the sign and lower bound of player i's interactive payoff are identified. Furthermore, researchers can infer how player i adjusts her beliefs across different observations; it naturally implies a testable restriction of unbiased belief.

Proposition 2. *Under the conditions met in Proposition 1 such that $\delta_i(a_i, a_{-i} = 1)b_i^1$ is point identified, if such a term is non-zero, it follows that*

(a) The interactive effect ratio $\frac{\delta_i(a_i=j,a_{-i}=1)}{\delta_i(a_i=k,a_{-i}=1)}$ is point identified for every two actions j and k.

(b) The sign of $\delta_i(a_i, a_{-i} = 1)$ and lower bound of $\left|\delta_i(a_i, a_{-i} = 1)\right|$ are identified for every a_i.

(c) Suppose the data contain observations with $J_i^{'}$, $J_i^{''} \geq 1$, then $\frac{b_i^1(J_i')}{b_i^1(J_i'')}$ is identified and naturally implies a testable restriction of unbiased belief: $\frac{b_i^1(J_i')}{b_i^1(J_i'')} = \frac{p_{-i}^1(J_i')}{p_{-i}^1(J_i'')}$.

The interactive effect ratio $\frac{\delta_i(a_i=j,a_{-i}=1)}{\delta_i(a_i=k,a_{-i}=1)}$ sheds light on a player's choice incentive. It concludes which of player i's action is more sensitive to the other player's behavior. In competitive games, a player has the incentive to choose an action that is insensitive to the other player's action. For instance, in an entry/expansion game, an incentive for a firm to open an additional store is to alleviate the negative impact of other firms; such an incentive can be measured by the interactive effect ratio, as it represents how the negative impact is attenuated when a firm opens one additional store. Similarly, an incentive for cooperation is also quantified by the interactive effect ratio in coordination games. For instance, a player has an incentive to choose a sensitive action to exploit positive spillover effects. Finally, in the context of a product choice game, the interactive effect ratio provides information about which product offered by firm i is a close substitute for the other firm's product.

The sign and ratio of the interactive effect also determine the strategic nature of the game. Suppose we have estimated that $\frac{\delta_i(a_i=j,a_{-i}=1)}{\delta_i(a_i=k,a_{-i}=1)} > 1$ for all $j > k$, players' actions are strategic substitutes if the sign of the interactive effect is negative and are strategic complements if it is positive.[9] Determining the strategic nature is one of the central questions in my empirical application; if firms' competition in operating hours shows strategic complementarity, this would imply that this strategic choice is related to vertical product differentiation and would have implications on the optimal (de)regulation of opening hours. Furthermore, inferring the sign of this interactive component is the main empirical question in many papers, such as Sweeting (2009) and De Paula and Tang (2012).

Table 1 summarizes the identification results in this section and compares them with existing literature. The first row presents the result in Aguirregabiria and Magesan (2017). With the usual exclusion restriction that affects only one player's payoff without effect on other players' preferences, they prove that a function of belief is identified. Such a function can be used to test the hypothesis of unbiased belief. However, there is no identification results for payoff functions except that researchers are willing to assume equilibrium behaviors at several realizations. In contrast, this paper shows that the asymmetry between players' action sets provides additional identification power. Without the usual exclusion restrictions, it partially identifies the payoff function. Additionally, variation of players' action sets across observations would allow researchers to almost point identify each player's payoff and belief. It naturally implies a testable restriction of unbiased belief. Finally, as footnote 3 states, a player's action set can be seen as another type of exclusion restriction. It affects each player's belief without affecting anyone's payoff. To avoid possible confusion, the action set is not referred to as exclusion restriction in this paper.

Table 1: Summary of Identification Results

Model Restrictions	Identified Set Payoff	$\frac{\delta_i(a_i=j,a_{-i}=1)}{\delta_i(a_i=k,a_{-i}=1)}$	$\pi_i(a_i,a_{-i})$	Sign, L.B. of $\delta_i(a_i,a_{-i})$	Identified Set Belief	Unbiased Belief Test
Exclusion Restriction: Existing Literature					✓	✓
Asymmetric Action Sets	✓					
Variation in Action Sets	✓	✓	✓	✓	✓	✓

[9]To see this clearly, recall that $\delta_i(a_i=k,a_{-i}=1)=\Pi_i(a_i=k,a_{-i}=1)-\Pi_i(a_i=k,a_{-i}=0)$, where $\Pi_i(a_i,a_{-i})$ represents player i's payoff of outcome (a_i,a_{-i}). Consequently, $\frac{\delta_i(a_i=j,a_{-i}=1)}{\delta_i(a_i=k,a_{-i}=1)}>1\ \forall\ j>k$ is equivalent to $\left|\Pi_i(a_i=k,a_{-i}=1)-\Pi_i(a_i=k,a_{-i}=0)\right|<\left|\Pi_i(a_i=j,a_{-i}=1)-\Pi_i(a_i=j,a_{-i}=0)\right|$, for all $j>k$. With $\delta_i(a_i,a_{-i}=1)<0$, it is precisely the condition of strategic substitutes since player i's incentive to choose a higher action decreases as player $-i$ increases her action.

3.3 Identification with Multiplicative Separability

In many empirical applications, players have the same number of possible choices; moreover, the action space may remain constant across observations. In those cases, the identification results from previous subsections cannot be directly applied. However, in this subsection, I show that a similar asymmetric feature can be constructed through a conventional restriction on a player's payoff function. For instance, suppose the interactive effect $\delta_i(a_i, a_{-i})$ is multiplicative separable between a player's own action and her opponent's action; then each player's belief is summarized by a sufficient statistic that is interpreted as the player's subjective expectation. This result implies a reduction in the dimension of player i's beliefs from J_{-i} to one (i.e. the dimension of the sufficient statistic). As a result, each player's identification problem mimics the structure of the asymmetric game described in Subsection 3.2 regardless of the number of actions available to the opponent. Therefore, the payoff function is partially identified. Moreover, when researchers can observe usual exclusion restrictions, we can sharpen the identified set and achieve similar results as when there is variation in players' action sets.

In this subsection, suppose $J_1 = J_2 = \bar{J} > 1$. Assumption 5 states a conventional payoff function such that the interactive effect is multiplicative separable between the two players' actions.

Assumption 5. $\delta_i(a_i, a_{-i}) = \delta_i(a_i, a_{-i} = 1) \cdot \eta_i(a_{-i})$ *with* $\eta_i(a_{-i} = 1) = 1$.

In the empirical application described by Example 1, $\delta_i(a_i = j, a_{-i} = 1)$ measures the impact of player $-i$'s first 24-hour store on player i's profit of opening j non-stop service stores. Consequently, $\eta_i(a_{-i} = k)$ captures the proportional change of the interactive effect when player $-i$ opens k stores at night. Assumption 5 restricts this proportional rate of change to be constant across player i's different actions.[10] This restriction is commonly imposed in the estimation of games with multiple players and multiple actions. For instance, Aradillas-Lopez and Gandhi (2016) exploit the same restriction; they refer to $\eta_i(\cdot)$ as the strategic index and to $\delta_i(\cdot)$ as the overall scale of the strategic effect.[11] Note that I allow $\eta_i(a_{-i})$ to depend on public information $\mathbf{x}$ even though it is suppressed for notational simplicity. Moreover, most current literature restricts $\eta_i(a_{-i})$ to be a particular functional form that only depends on

[10]For instance, $\frac{\delta_i(a_i=k, a_{-i}=j)}{\delta_i(a_i=k, a_{-i}=j')} = \frac{\delta_i(a_i=k', a_{-i}=j)}{\delta_i(a_i=k', a_{-i}=j')}$ $\forall\ j' \neq j,\ k' \neq k$.

[11]Aradillas-Lopez and Gandhi (2016) allow ε_i to enter $\delta_i(a_i, a_{-i} = 1)$ in a non-linear fashion, while it is additive separable in my framework. However, their interest is in the strategic index $\eta_i(a_{-i})$ and needs to assume that $\delta_i(a_i, a_{-i} = 1)$ is non-decreasing in a_i to establish the identification result. In contrast, my interest is $\delta_i(a_i, a_{-i} = 1)$ and so does not require any more restrictions on payoff functions.

the other player's action; some examples include $\eta_i(a_{-i}) = \log(1+a_{-i})$, i.e. each additional store opened by player $-i$ affects player i's profit at a diminishing rate as specified in Nishida (2014) or $\eta_i(a_{-i}) = a_{-i}$, i.e. each additional store opened by player $-i$ affects player i's profit at a constant rate as specified in Augereau et al. (2006). In contrast, this paper allows $\eta_i(\cdot)$ to be a non-parametric function in both a_{-i} and $\mathbf{x}$. Under Assumption 5, player i's expected payoff for action a_i defined in equation (2) becomes

$$\begin{aligned} E\big[\Pi_i(\varepsilon_i, a_i)\big] &= \pi_i(a_i) + \sum_{j=1}^{J_{-i}} \delta_i(a_i, a_{-i}=1)\cdot \eta_i(a_{-i}=j)\cdot b_i^j + \varepsilon_i(a_i) \\ &= \pi_i(a_i) + \delta_i(a_i, a_{-i}=1)\cdot \big[\sum_{j=1}^{J_{-i}} \eta_i(a_{-i}=j)\cdot b_i^j\big] + \varepsilon_i(a_i) \qquad (6) \\ &= \pi_i(a_i) + \delta_i(a_i, a_{-i}=1)\cdot g_i + \varepsilon_i(a_i), \end{aligned}$$

where $g_i = \sum_{j=1}^{J_{-i}} \eta_i(a_{-i}=j)\cdot b_i^j$ represents player i's subjective expectation of the value $\eta_i(a_{-i}=j)$. In the previous subsection, if $J_{-i}=1$, then player i's expected payoff of action a_i is

$$E\big[\Pi_i(\varepsilon_i, a_i)\big] = \pi_i(a_i) + \delta_i(a_i, a_{-i}=1)\cdot b_i^1 + \varepsilon_i(a_i).$$

It is easy to see that if we treat g_i in equation (6) as being analogous to b_i^1 in the above equation, then these two equations would share the same structure. As a result, the identification result in Proposition 1 (a) holds trivially. This is summarized in the following corollary.

Corollary 1. *Under Assumptions 1 to 5, for any two actions j and k, player i's payoff parameters $\pi_i(\cdot)$ and $\delta_i(\cdot)$ are given by the set of values that satisfies the following restriction:*

$$\frac{F_i^j(\mathbf{p}_i) - \pi_i(a_i=j)}{F_i^k(\mathbf{p}_i) - \pi_i(a_i=k)} = \frac{\delta_i(a_i=j, a_{-i}=1)}{\delta_i(a_i=k, a_{-i}=1)}.$$

Existing literature on the estimation of empirical games assumes the existence of an exclusion restriction, a variable that affects one player's payoff but not that of other player(s). Without such a restriction, the payoff function is non-identified, even with the BNE assumption (Bajari et al. (2010) and Aradillas-Lopez (2010)). In this subsection, I show that with the usual exclusion restrictions, researchers can sharpen the identified set in Proposition 1 and achieve similar results as games with variation in action spaces. Assumption 6 states the conditions on exclusion restrictions.

Assumption 6. *(a) For each player i, there exists a variable $z_i \in \mathbb{R}$ that affects only player i's payoff; moreover, z_i has exogenous variation over its support.*

(b) There exists a variable $s \in \mathbb{R}$ that affects each player's interactive effect $\delta_i(\cdot)$ but not the base return $\pi_i(\cdot)$; moreover, s has exogenous variation over its support.

As explained above, the existence of z_i is commonly assumed in literature. In the empirical application described in Example 1, a plausible candidate for z_i would be the market's distance to chain i's nearest distribution center. While this distance could substantially affect chain i's delivery costs and profit, it has no direct impact on chain $-i$, though it may indirectly affect chain $-i$ through its impact on chain i.

Assumption 6 (b) requires that s does not affect player i's payoff if the other player chooses action 0. Even though such an exclusion restriction is ignored in literature on identification of games, it is usually specified in many existing empirical applications. For instance, in Example 1, a plausible candidate for s could be the distance between KFC and McD in a single market. If McD does not operate through the night, such a distance would have no impact on KFC's night profit. In contrast, the interactive effect would be affected by this distance, since an opponent of closer proximity may have a larger impact than one that is further away. This type of horizontal differentiation created by distance has been studied in empirical games by Seim (2006), Zhu and Singh (2009), and Rennhoff and Owens (2012). Moreover, consider another context of entry game in the airline industry studied by Ciliberto and Tamer (2009). When airline i's competitor enters into a single market (e.g. city-to-city pair), company i's profit depends on its competitor's characteristics in surrounding markets because consumers would prefer a convenient connecting flight. In contrast, these characteristics have no impact on airline i if its competitor is absent. This type of network effect is very prevalent in many industries, such as retail, fast food and banking. All previous papers introduce variable s as a plausible model specification instead of an instrument/exclusion restriction to facilitate identification and estimation. In contrast, this paper formally discusses the role of such a variable in identification without imposing the equilibrium assumption.

Proposition 3. *(a) Under Assumptions 1 to 5 and Assumption 6 (a), $\frac{\delta_i(z_i, a_i=j, a_{-i})}{\delta_i(z_i, a_i=k, a_{-i})}$ is identified for any two actions j and k if there exist at least two realizations of z_{-i}, say z_{-i}^1 and z_{-i}^2, such that $\mathbf{p}_i(z_i, z_{-i}^1) \neq \mathbf{p}_i(z_i, z_{-i}^2)$.*

(b) Suppose further Assumption 6 (b) holds and there exist at least two realizations of s, say s^1 and s^2, such that $\frac{\delta_i(z_i,s^1,a_i=j,a_{-i})}{\delta_i(z_i,s^1,a_i=k,a_{-i})} \neq \frac{\delta_i(z_i,s^2,a_i=j,a_{-i})}{\delta_i(z_i,s^2,a_i=k,a_{-i})}$, then $\pi_i(z_i,a_i)$ and $\delta_i(z_i,s,a_i,a_{-i}=1)g_i(z_i,z_{-i},s)$ are point identified for every $a_i \in A_i$ and every z_i, z_{-i}, s.

Existing empirical applications usually restrict $\eta_i(\cdot)$ to be a positive function, and consequently $g_i(\cdot)$ is positive. Under such a restriction, the sign of the interactive payoff is identified given the identification of $\delta_i(\cdot)g_i(\cdot)$.[12] Moreover, since the multiplicative separability mimics the feature in games with asymmetric actions, the exclusion restrictions studied in this subsection can be used to sharpen the identified set in Proposition 1 (a) when action spaces are asymmetric across players but are constant across observations. In addition, the method proposed by Aguirregabiria and Magesan (2017) can be used to construct a testable restriction of unbiased belief. Finally, this result can be generalized to a game with more than two players as shown in Appendix A.4.

3.4 Endogenous Action Sets

The identification results in Subsection 3.2 assume that players' action sets are exogenous. However, in many empirical applications, each player's possible choices (e.g. number of stores or products in a single market) are endogenously determined. When both action sets and actual choice depend on some common variables which are unobserved to researchers, we encounter a selection problem. For instance, firms tend to open more stores in markets with better unobserved heterogeneity; consequently, comparing outcomes of markets with a different number of outlets does not only reflect strategic effect and belief but also the difference of unobserved heterogeneity. Therefore, the results in Subsection 3.2 may not hold. This subsection addresses the selection problem under a finite mixture assumption. I establish identification results when the number of support for the unobservable is two and leave the results for more components in future research.

To facilitate the illustration, I describe the model and identification results based on Example 1. The competition between KFC and McD can be seen as a two-stage game. In the first stage, denoted by t_1, each player i simultaneously chooses $J_i \in \{0,1,\cdots,\bar{J}\}$. As described above, decision variable J_i represents the number of stores that chain i builds in a single market. In the second stage t_2 and conditional

[12] Since $g_i(\cdot)$ represents subjective expectation instead of belief, it is not a valid probability distribution. Therefore, the lower bound of the interactive payoff is not identified.

on market structure (J_1, J_2) chosen in first stage, player i decides $a_i \in \{0, 1, \cdots, J_i\}$ where a_i represents how many stores to operate for 24 hours as described in Section 2.

Player i's payoff function at each stage is defined as

$$\Pi_{i,t_1}(\mathbf{x}, \tilde{x}, \omega, \varepsilon_{i,t_1}, J_i, J_{-i}) = \tilde{\Pi}_{i,t_1}(\mathbf{x}, \tilde{x}, \omega, J_i, J_{-i}) + \varepsilon_{i,t_1}(J_i), \tag{7}$$

$$\Pi_{i,t_2}(\mathbf{x}, \omega, \varepsilon_{i,t_2}, a_i, a_{-i}) = \pi_{i,t_2}(\mathbf{x}, \omega, a_i) + \delta_{i,t_2}(\mathbf{x}, \omega, a_i, a_{-i}) \cdot \mathbb{1}(a_{-i} \neq 0) + \varepsilon_{i,t_2}(a_i). \tag{8}$$

Term ω represents the variable at market level that affects each player's payoffs in both stages. It is observed by both players but unobserved by econometricians. It is the existence of such unobservable that causes the selection problem. Moreover, the second-stage payoff function $\Pi_{i,t_2}(\cdot)$ defined by equation (8) is an adjusted version of equation (1) that takes the market-level unobservable into account. In addition, $\tilde{\Pi}_{i,t_1}(\cdot, J_i, J_{-i})$ represents player i's first-stage expected payoff when the market structure is (J_i, J_{-i}). Such a payoff depends on player i's expectation of her own and the other player's behaviors conditional on first-stage choice (J_i, J_{-i}); for instance, decisions on pricing, quality and store hours. If each player has unbiased belief in both stages, the equilibrium payoff in the second stage or its discounted value would enter linearly into the first-stage payoff function. As this paper allows players to have off-equilibrium behaviors, I non-parametrically specify the first-stage payoff function and abstract its relationship with the second-stage payoff. Finally, public information $\mathbf{x}$ is constant between two periods for the sake of brevity. The model and identification results easily generalize to allow $\mathbf{x}$ to vary across time.

The payoff functions defined by equations (7) and (8) place two restrictions. First, there exists a variable $\tilde{x}$, observed by both players and researchers, that only enters into the first-stage payoff. As the decision to construct a store or enter a market requires considerable one-time investment, any variable that affects players' entry cost could be a plausible candidate for $\tilde{x}$. The second restriction assumes that the unobservable ω is constant across two stages. Equivalently, there exists no market-level shock on the unobservable at period t_2.[13] Such a restriction can be a good approximation if t_1 and t_2 are sufficiently close.

The following assumption states the restrictions on unobservable ω and private information $\varepsilon_i(\cdot)$.

Assumption 7. *(a) ω is a discrete variable with finite support $\{\omega^1, \omega^2, \cdots, \omega^L\}$; in addition, $h(\omega^l|\mathbf{x}, \tilde{x})$*

[13] In other words, any common knowledge market-level shock at period t_2 can be perfectly predicted by ω.

represents the probability density function conditional on $(\mathbf{x}, \tilde{x})$.

(b) $\varepsilon_{i,t} = \big(\varepsilon_{i,t}(0), \cdots, \varepsilon_{i,t}(\bar{J})\big)'$ *follows a CDF* $G_{i,t}(\cdot)$ *that is absolutely continuous in Lebesque measure in* $\mathbb{R}^{\bar{J}+1}$. *Moreover, it is independent across players, time and public information* $(\mathbf{x}, \tilde{x})$.

This paper studies a game with discrete choice. It is impossible to non-parametrically infer how a continuous unobservable affects each player's choice probability (e.g. Hu (2017)). Consequently, as stated in Assumption 7 (a), this paper requires unobservable ω to have finite support. Moreover, Assumption 7 (b) extends the independence restriction of Assumption 1 to a two-stage game. In addition to independence across players, it also requires each player's private information to be independent across time. As I will describe below, if researchers relax such an assumption and allow private information to be correlated across time, a function of belief is not identified. Consequently, it is impossible to construct an unbiased belief test. However, all identification results for payoff function still hold. Moreover, if researchers observe the exclusion restriction studied in Subsection 3.3, an unbiased belief test can be constructed using the method developed by Aguirregabiria and Magesan (2017).

Given Independence Assumption 7, player i's private information $\varepsilon_{i,t}(\cdot)$ has no predictive power for player $-i$'s payoff. Therefore, following Assumption 2, each player's beliefs in both stages depend only on public information that is known by both players. Define $b^j_{i,t_1}(\mathbf{x}, \tilde{x}, \omega)$ as player i's first-stage belief about the probability that player $-i$ builds j stores in a single market. Similarly, $b^j_{i,t_2}(\mathbf{x}, \omega, J_i, J_{-i})$ represents the second-stage belief about the probability that player $-i$ operates j stores during the night. Given these beliefs, each player chooses an action that maximizes her expected payoff in each stage. Consequently, the optimal strategy takes the following form:

$$J_i^*(\mathbf{x}, \tilde{x}, \omega, \varepsilon_{i,t_1}) = \underset{J_i \in \{0, \cdots \bar{J}\}}{\operatorname{argmax}} \left\{ \sum_{j=0}^{\bar{J}} \tilde{\Pi}_{i,t_1}(\mathbf{x}, \tilde{x}, \omega, J_i, j) \cdot b^j_{i,t_1}(\mathbf{x}, \tilde{x}, \omega) + \varepsilon_{i,t_1}(J_i) \right\},$$

$$a_i^*(\mathbf{x}, \omega, \varepsilon_{i,t_2}, J_i, J_{-i}) = \underset{a_i \in A_i}{\operatorname{argmax}} \left\{ \pi_{i,t_2}(\mathbf{x}, \omega, a_i) + \sum_{j=1}^{J_{-i}} \delta_{i,t_2}(\mathbf{x}, \omega, a_i, a_{-i} = j) \cdot b^j_{i,t_2}(\mathbf{x}, \omega, J_i, J_{-i}) + \varepsilon_{i,t_2}(a_i) \right\}.$$

It is easy to see the selection problem through optimal strategy $a_i^*(\cdot)$. The identification of payoff and belief replies on the variation of (J_i, J_{-i}), which depends on unobservable ω. This causes an endogeneity problem. Intuitively, the variable $\tilde{x}$, which enters only in the first-stage payoff, serves as an instrument to exogenously shift (J_i, J_{-i}) and identifies the model.

Similar to Section 2, define $p^j_{i,t_1}(\mathbf{x},\tilde{x},\omega)$ and $p^j_{i,t_2}(\mathbf{x},\tilde{x},\omega,J_i,J_{-i})$ as player i's CCPs in each stage with the following expression:

$$p^j_{i,t_1}(\mathbf{x},\tilde{x},\omega) = \int \mathbb{1}\left\{J^*_i(\mathbf{x},\tilde{x},\omega,\varepsilon_{i,t_1}) = j\right\} dG_{i,t_1}(\varepsilon_{i,t_1}),$$
$$p^j_{i,t_2}(\mathbf{x},\omega,J_i,J_{-i}) = \int \mathbb{1}\left\{a^*_i(\mathbf{x},\omega,\varepsilon_{i,t_2},J_i,J_{-i}) = j\right\} dG_{i,t_2}(\varepsilon_{i,t_2}).$$

Unlike in Section 3, these CCPs cannot be directly estimated from the data since ω is unobserved to researchers. Instead, for any outcome (J_1,J_2,a_1,a_2), the conditional joint distribution is defined as

$$Pr(J_1,J_2,a_1,a_2|\mathbf{x},\tilde{x}) = \sum_{l=1}^{L}\sum_{i=1}^{2} h(\omega^l|\mathbf{x},\tilde{x}) \cdot p^{J_i}_{i,t_1}(\mathbf{x},\tilde{x},\omega^l) \cdot p^{a_i}_{i,t_2}(\mathbf{x},\tilde{x},\omega^l,J_i,J_{-i}). \tag{9}$$

Extending the data generating process described in Subsection 3.1 to a two-stage game, then $Pr(J_1,J_2,a_1,a_2|\mathbf{x},\tilde{x})$ can be consistently estimated from the data. The following proposition states the main result:

Proposition 4. *Under Assumption 7, suppose $L = 2$, $\bar{J} \geq 1$, $Pr(J_1,J_2,a_1,a_2|\mathbf{x},\tilde{x})$ is known by researchers and $\tilde{x}$ has exogenous variation; then each player's conditional choice probabilities at both stages $\mathbf{p}_{i,t_1}(\cdot)$, $\mathbf{p}_{i,t_2}(\cdot)$ and the probability density function of unobservable $h(\omega|\mathbf{x},\tilde{x})$ are non-parametrically identified up to relabel.*

Proposition 4 establishes the identification of conditional choice probability under the assumption that unobservable ω has two possible values. When $\bar{J} > 1$, such a requirement is testable given the identification and estimation results proposed by Kasahara and Shimotsu (2014). In addition, the identification result requires $\tilde{x}$ to have enough variation such that the conditional joint distribution $Pr(\cdot)$ satisfies a rank condition. As $Pr(\cdot)$ is observed or directly estimated, such a requirement is also testable. Details of the rank condition are given in Appendix A.1.

Proposition 4 also contributes to finite mixture literature studied by Hall and Zhou (2003), Hall et al. (2005), Henry et al. (2014), Compiani and Kitamura (2016), Hu (2017), and Aguirregabiria and Mira (2018). It is commonly known that, to non-parametrically identify the finite mixture model, researchers need to observe at least three random variables that are independent conditional on unobservable ω. In this paper, even though researchers can observe four random decision variables (i.e. each player's choice at both stages), they are not conditional independent as a player's second-stage choice depends on her first-stage decision. Therefore, the identification results in the finite mixture model do not directly apply. To see the identification intuition, consider the variation of $\tilde{x}$. It exogenously shifts each player's decision in the first stage; consequently, the distribution of ω conditional on the first-stage decision (J_1, J_2) depends on $\tilde{x}$.[14] Moreover, $\tilde{x}$ does not affect each player's second-stage conditional choice probability. Therefore, the identification problem in the second stage mimics the model studied by Henry et al. (2014) in which an exclusion restriction only affects the mixture weights. They show that with the observation of just one random variable, such a model is identified up to mixture weights at only two realizations of the exclusion restriction. In this paper, one important difference from Henry et al. (2014) is that researchers are able to observe two random decision variables (i.e. each player's choice) in the second stage; the observation of this additional random variable ensures the point identification. This result can be particularly useful in empirical applications of a two-stage game with market-level unobserved heterogeneity, either with or without equilibrium constraint. At last, it is important to note that the proof of Proposition 4 in Appendix A.1 does not exploit all restrictions the model imposes; therefore, it seems that we can achieve at least partial identification with a larger number of support in ω. Moreover, Appendix A.4 generalizes the game to more than two players; in this situation, Aguirregabiria and Mira (2018) show that each player's CCPs with a larger support of ω is point identified.

The identification objective is each player's payoff and belief in the second stage. As the proof of Propositions 1 and 2 suggests, under Assumptions 1 to 4, the identification of payoff and belief only requires that each player's CCPs can be estimated by researchers. Consequently, given Proposition 4, the results in Propositions 1 and 2 trivially hold true under the existence of unobservable ω. It is summarized by following corollary.

Corollary 2. *Under conditions met by Proposition 4 and Assumption 3, 4, for each player i and* $(\mathbf{x}, \tilde{x})$*, it follows that*

(a) The base return $\pi_{i,t_2}(\cdot)$ *is point identified.*

(b) The subjective expectation of other players' impact $\sum_{j=1}^{J_{-i}} \delta_{i,t_2}(\mathbf{x}, \omega, a_i, a_{-i} = j) \cdot b^{j}_{i,t_2}(\mathbf{x}, \omega, J_i, J_{-i})$ *is identified* $\forall\ J_i,\ J_{-i} \leq \bar{J}$.

[14]This property holds true regardless of whether $h(\omega)$ depends on $\tilde{x}$ or not.

(c) The interactive effect ratio $\frac{\delta_{i,t_2}(\mathbf{x},\omega,a_i=j,a_{-i}=1)}{\delta_{i,t_2}(\mathbf{x},\omega,a_i=k,a_{-i}=1)}$ *is identified for each of two actions j and k.*

(d) For any two choices in first stage $J_i', J_i'' \geq 1$, $\frac{b^1_{i,t_2}(\cdot,J_i',J_{-i}=1)}{b^1_{i,t_2}(\cdot,J_i'',J_{-i}=1)}$ *is identified and naturally implies a testable restriction of unbiased belief in the second stage:* $\frac{b^1_{i,t_2}(\cdot,J_i',J_{-i}=1)}{b^1_{i,t_2}(\cdot,J_i'',J_{-i}=1)} = \frac{p^1_{-i,t_2}(\cdot,J_i',J_{-i}=1)}{p^1_{-i,t_2}(\cdot,J_i'',J_{-i}=1)}$.

Suppose researchers relax Assumption 7 such that private information is correlated across time but still independent across players, then the distribution of player i's private information in the second stage ε_{i,t_2} is independent of the other player's decision J_{-i}. Consequently, the results in Corollary 2 (a), (b) and (c) also hold true as these results rely on the variation of J_{-i} in the second stage. However, the result in Corollary 2 (d) no longer holds as such a result relies on the variation of J_i; player i's decision J_i affects the conditional distribution of ε_{i,t_2} as private information is correlated across time.

4 Empirical Application

4.1 Motivation and Industry Background

China's Western-style fast food industry is characterized as a duopoly of KFC and McD.[15] KFC opened its first Chinese outlet in Beijing in 1987. Three years later, McD opened its first outlet in Shenzhen, Guangdong Province. After that, these two Western-style fast food chains expanded their business in the world's largest emerging economy. Despite its leading role in the world, McD expanded its business at a slower rate and its outlet stores are outnumbered by KFC by more than 2:1. Specifically, at the beginning of 2016, KFC operated 4,952 outlets across China, while McD only owned 2,231. In terms of geographic distribution, KFC operated in all 31 provinces of mainland China, while McD only entered into 27 provinces.[16] Unlike the case in Western countries, KFC and McD are considered to be close substitutes in China. For instance, China's KFC serves hamburger beef together with chicken products, and McD also sells a variety of fried chicken. Moreover, both chains offer traditional Chinese food

[15] Other giant Western-style fast food companies have relatively small market shares in China. By 2014, Burger King had more than 300 stores while Subway owned about 600 stores. Most of these stores were located in the more developed areas of China. Some consider another brand called Dicos as a third player in the Chinese market. It serves very similar products to KFC and McD and has a comparable size with McD. By the end of 2016, Dicos operated 2,092 stores, distributed across all 31 provinces in mainland China. However, only 2% of Dicos stores operated at night. Therefore, I do not model Dicos as a third player in the night market.

[16] In April 2016, the four provinces that McD had not yet entered were Xinjiang, Ningxia, Tibet, and Qinghai. These four provinces are excluded from my sample.

(e.g. congee) to attract Chinese consumers. Another important feature in the operation of these chains in China, in contrast to Western countries, is that the companies directly own and operate most of their outlets. At the end of 2014, about 15% of McD's stores were franchised. Similarly, KFC had less than 10% franchised stores.[17] On the logistics side, HAVI Logistics provides distribution services for McD and has seven distribution centers across China. The holding company of KFC, YumChina, provides logistic services itself and owns 16 distribution centers. For both chains, the distribution center supplies key raw materials (e.g. raw meat) to each store. Consequently, a market's distance to a chain's nearest distribution center substantially affects the logistic costs for that chain. In my analysis, this distance is used as an exclusion restriction on a single player's payoff. It facilitates the identification and estimation of players' payoffs and beliefs.

KFC and McD compete with each other through many dimensions such as entry/expansion, menu selection, pricing, and location choice. In this paper, I study their strategic interactions through store hours decisions.[18] It aims to answer two empirical questions. First, is the decision of store hours a strategic substitute or strategic complement? Second, do firms have biased expectation of their competitors' behaviors? If so, what are the features of biased beliefs and how do they affect the estimated payoff functions?

Many countries have strong regulations on stores' opening hours in the retail industry and there is an ongoing debate on the deregulation of such restrictions. The strategic nature of opening hours is then the key element in the evaluation of these deregulation policies. In the theoretical literature, Inderst and Irmen (2005), Shy and Stenbacka (2006, 2008), and Wenzel (2010, 2011) view the decision of store hours as a strategic device that softens price competition. In addition, it is horizontally differentiated because consumers may prefer to consume at different times of the day. In contrast, my discussions with industry experts suggest that extending operating hours enables the staff to better prepare for busy breakfast service, in addition to building brand value. Specifically, if a chain successfully attracts a consumer at night, it is more likely to also attract this consumer during the day-time due to lower search costs or switching costs. Under this view, the choice of store hours has a component of vertical differentiation similar to

[17]For information on McD, see Sina News, retrieved from http://finance.sina.com.cn/chanjing/gsnews/2016-05-23/doc-ifxsktkr5912661.shtml. For KFC's information, see People News, retrieved from http://finance.people.com.cn/n/2012/0927/c70846-19130602.html.

[18]At earlier stages of their business in China, neither chain operated at night. As China's economy grew, McD started to operate 24-hour stores in 2005 and KFC began in 2009.

quality choice. This view is taken by Ferris (1990) and Klemperer and Padilla (1997). Failing to consider this quality aspect would under-evaluate not only consumer welfare gains but also stores' business hours from policy deregulation of opening hours. Therefore, it is essential to quantify the extent of the vertical differentiation component in the choice of store hours. Unfortunately, there is little empirical evidence on this question. To the best of my knowledge, only Kügler and Weiss (2016) empirically study store hours competition in the Austrian gasoline market and they find insignificant strategic effect.[19]

Although KFC and McD have competed in China for almost 30 years and are familiar with each other, drastic changes in the economic environment and market conditions could preclude perfect prediction of a competitor's behaviors. At the macroeconomic level, China maintains its economic growth at a miracle rate. At the industrial level, several news agencies in China exposed a series of KFC food safety scandals.[20] These scandals not only depreciated KFC's brand value in China, but also raised concerns about Western-style fast food restaurants in general. Euromonitor International reported that KFC's sales decreased after 2012 while McD's sales remained constant,[21] amid the rapid growth of China's catering industry in recent years.[22] As reported by China Market Research Group, consumers have started to lose trust in both chains. These dramatic changes in economic conditions and consumer preferences represent structural transformations. They pose difficulty on each firm's strategic reasoning and are sources of firms' biased beliefs.

From a practical standpoint, if players have biased beliefs in empirical games, incorrectly imposing the equilibrium condition will bias the estimated magnitude or even the sign of the interactive effects. Additionally, investigating whether firms have unbiased expectations and how they form beliefs has its own interest, especially when the economic environment varies dramatically.

The identification results studied in this paper enable researchers to quantify the extent of store hours' quality aspect and it is robust to potential biased beliefs. Moreover, we can infer how a chain adjusts its

[19] Kügler and Weiss (2016) employ a reduced form estimation. In contrast, this paper estimates a structural model and has several advantages. First, researchers can take potential biased beliefs into account, so the estimates are robust. Second, the estimates of interactive effects have structural interpretations and can be compared with the base return (e.g. monopoly profit). Third, the estimates can be used to construct counterfactual predictions.

[20] See "Chinese Consumers Are Losing Trust in McDonald's and KFC," retrieved from http://www.businessinsider.com/r-food-scares-strip-mcdonalds-kfc-of-treat-status-in-china-2015-8.

[21] See "China Starts to Loses Its Taste for McDonald's and KFC," retrieved from https://www.agweb.com/article/china-starts-to-loses-its-taste-for-mcdonalds-and-kfc-blmg.

[22] As reported by the Ministry of Commerce of the P.R.C., the annual growth rate of sales in China's catering industry was 10.6% over the period from 2010 to 2015.

expectations and test the hypothesis of unbiased beliefs.

4.2 Data

I gather information on every KFC and McD outlet store in China through their respective official websites.[23] The information for each store includes its brand name, address, telephone number and other store characteristics, such as 24-hour service, breakfast service, and drive-through. Due to consumers' travel costs, distances between stores within the same chain and between different chains can have an impact on the store hours decision. Such a network structure is very challenging to control when the number of stores is large, especially in Beijing or Shanghai, where there are about 500 stores.[24] Therefore, instead of studying strategic interactions in large cities, I focus on small counties/districts in which the number of total stores of KFC and McD is considerably smaller.[25]

Consumers are typically reluctant to travel long distances for fast food, especially at night. Therefore, I define market at a finer level. For every county where KFC or McD exists, I obtain a list of cinemas through Baidu Maps. I then derive the centroid of each cluster of cinemas and define a market as an area that lies within a 5-km radius around the centroid.[26] This market definition does not necessarily assume that night consumers for fast food are customers in cinemas. Typically, cinemas are located in entertainment districts or densely populated residential areas in China; therefore, the centroid of cinemas cluster serves as an approximation for the location where night life is present. I believe that people who reside or go there are potential consumers of fast food. Based on this market definition, 95% of the stores in the sample are included in one of these local markets. Moreover, 35 out of 1171 counties/districts encompass multiple markets while the rest of the counties have a unique market.[27] In this study, I focus on markets where the number of KFC stores is less than or equal to 4 and McD has no more than 1 store.

[23]The websites are http://www.kfc.com.cn/kfccda/storelist for KFC and http://www.mcdonalds.com.cn/top/map for McD. The data were gathered on 28 April, 2016.

[24]Specifically, such a network structure is completely characterized by a vector of distances between any two stores in the market. It then consists of $\frac{n(n-1)}{2}$ elements in a market with n stores. A complete characterization of large markets would yield very imprecise estimates due to high dimensionality.

[25]The political hierarchy in China is Nation→Province→City→County/District→Town. County is the smallest unit at which demographic data are richly available. Moreover, district is at the same level as county and is typically located in the central area of a city. Throughout this paper, I refer to districts as counties for the sake of brevity.

[26]See Appendix A.5 of how I construct the clusters. Since KFC and McD are typically located close to each other, a radius of 5 km is a very conservative definition; market rarely changes using a radius of 3 km as another definition.

[27]Given these statistics, observations rarely change when markets are defined at the county level; therefore, the estimation results are robust to different market definitions.

These markets comprise almost 90% of the total sample. Including markets with more stores expands McD's action spaces and raises KFC's payoff and belief dimensions. However, given a few additional observations, the parameters for these extra dimensions and KFC's impact on McD would be imprecisely estimated. Therefore, I only study how KFC's action is affected by McD in this paper. Table 2 presents the joint distribution of the number of stores per market for these two chains. These numbers vary across markets; therefore, the identification results in Subsection 3.2 can be applied to identify KFC's payoff and belief.

Table 2: Distribution of the Number of Markets by the Number of KFC and McD Stores

	McD Stores		
KFC Stores	0	1	Total
0	506	20	526
1	390	34	424
2	107	42	149
3	36	32	68
4	23	17	40
Total	1,062	145	1,207

Note: KFC and McD act as potential entrants in markets without any chain. These markets are selected based on the existence of Dicos. Equivalently, KFC and McD are treated as potential entrants when Dicos exists.

Demographic variables such as population and GDP are obtained through China Data Online, provided by the University of Michigan. I also collect night light data from the NOAA National Geophysical Data Center, provided by the U.S. Department of Commerce. Night light data measure the development level and population density of a local market. Furthermore, I collect information on the geographical location of every distribution center for both McD and KFC and calculate their distances to each market. I also calculate the average minimum distance of stores of the same chain and of different chains within the same local market. Table 3 presents summary statistics.

4.3 Reduced Form Estimation

As informed by fast food industry experts, it is the local/regional manager who decides the operating hours for every outlet in a market. Therefore, I model this situation as a static game, such that each chain simultaneously chooses how many of its current outlets to keep open at night. Specifically, given the

Table 3: Summary Statistics on Local Markets

Variable	Definition	Mean	Std. Dev.	Min	Max
Income	GDP per capita, 10,000 RMB	4.13	3.52	0.51	45.94
Pop	Population, 100,000	6.52	3.78	0.28	28.5
Center	Dummy, =1 if market located at city center	0.11	0.32	0	1
Light	Night light density	54.56	7.93	19	63
KFCDist	Average distance between KFC Stores, km	0.51	0.86	0	6.10
z_{KFC}	Distance to nearest KFC's distribution center, 100 km	2.12	1.36	0.09	10.16
z_{McD}	Distance to nearest McD's distribution center, 100 km	2.76	1.75	0.31	9.91
Cinema	Number of cinemas	1.65	1.98	0	14
s	Distance between two chains' centroids	0.80	0.75	0.01	3.73
KFCStores	Number of KFC stores	0.90	1.04	0	4
McDStores	Number of McD stores	0.12	0.33	0	1
KFC24h	Number of KFC 24-hour stores	0.56	0.71	0	3
McD24h	Number of McD 24-hour stores	0.38	0.49	0	1
GDPGrowth	Average annual growth rate of GDP per capita from 2000 to 2014	15.23	3.80	1.22	36.97
South	Regional dummy, =1 if in Guangdong or Hainan Prov.	0.05	0.21	0	1
NE	Regional dummy, =1 if in Liaoning, Jilin or Heilongjiang Prov.	0.10	0.30	0	1
Observations	1207				

Note: Statistics for z_{KFC} and $KFC24h$ are calculated conditional on the existence of KFC. Statistics for z_{McD} and $McD24h$ are calculated conditional on the existence of McD. Statistics for s are calculated conditional on the existence of both chains.

market structure shown in Table 2, KFC has three choices when it owns more than one outlet, opening zero/one or two stores at night.[28] In contrast, McD only chooses between opening zero/one 24-hour store. Moreover, both chains' action spaces vary across markets. Therefore, the identification results in Subsection 3.2 can be applied to identify KFC's profit function and beliefs. In addition to variation in action space, my data set also contains valid exclusion restrictions that facilitate identification and estimation. Specifically, these exclusion restrictions are z_{MCD} (i.e. distance to nearest McD's distribution center) and s (i.e. distance between the two chains' centroids).

To address the endogeneity of players' action sets, I consider the two-stage game described in Subsection 3.4. I assume that each chain's decisions at both stages depend on a market-level unobservable. Such an unobservable has two possible values. Specifically, at the first stage t_1, each chain decides whether to enter into a local market. Such a decision depends on the economic condition at t_1. Moreover, as entry requires dynamic consideration, it is also affected by the firm's expectation of future profitability. In the

[28] I treat the operation of two and more than two 24-hour stores as the same action because KFC operates more than two stores at night in only 1.5% of the markets. Furthermore, when KFC only owns one store in a market, it cannot operate more than one 24-hour store. In this case, opening more than one 24-hour store is treated as a strictly dominated action and will not be chosen.

second stage t_2, each chain makes a static decision on opening hours for every store it owns. Such a choice depends on the market environment at t_2 and is independent of past economic condition. In my sample, *GDPGrowth* is the geometric average of GDP growth rate during the past 15 years. Conditional on *Income* at t_2, the growth rate measures the past *Income* and the firm's expectation of market expansion speed at stage t_1.[29] Consequently, it affects each chain's first-stage decision but has no impact on the second-stage choice. As described in Subsection 3.4, the existence of such an exclusion restriction identifies each player's CCPs at both stages and weighting mixtures.

Table 4 presents a reduced form Multinomial Logit model of KFC's decision concerning how many outlets to operate overnight. The reduced form estimates for both chains' entry decision and McD's store hours choice are presented in Appendix A.5. As the first column suggests, demographic variables such as *Income* and *Population* are statistically insignificant while *Light* explains a large fraction of variation in KFC's decisions. Because market is defined as substantially smaller than the unit for which *Income* and *Population* data are gathered, night light data consequently provide additional information at a finer level and substantially increases in-sample fitness. This is consistent with a growing literature that measures economic activity using outer-space data (e.g. Henderson et al. (2012)). Moreover, the number of cinemas is added as proxies for local demand of fast food, and its estimates are significantly positive, as anticipated.

Given the market structure in Table 2, *MCDStores* is a dummy that equals one if McD is present in the market. The estimated coefficient on $z_{McD} \times MCDStores$ suggests a significant negative impact of z_{McD} on KFC's decision in markets where McD owns an outlet. Intuitively, z_{McD} has a negative effect on McD's profit as it increases delivery cost. Moreover, it only affects KFC's profit indirectly through its impact on McD's decision because the distance to McD's distribution center is irrelevant to KFC's own costs or revenues. Table 4 suggests that the negative impact of z_{McD} on McD is transformed to a negative influence on KFC, which suggests that these chains' store hours decisions are strategic complements if KFC correctly predicts McD's behaviors.

As Subsection 3.3 describes, if z_{McD} affects KFC's decision only through its impact on McD, it provides additional identification power to infer how KFC adjusts its belief about McD's choice. With the help of variation of McD's action sets, I can formally test whether z_{McD} is a valid exclusion restriction.

[29] It is not a perfect measure of *Income* at t_1 since I do not have information on the entry time.

Table 4: Multinomial Logit Regression: Number of KFC's 24-Hour Stores

	One 24-H	Two 24-H	One 24-H	Two 24-H
Income	0.0118	0.0007	0.0089	0.0015
	(0.0420)	(0.0895)	(0.0420)	(0.0908)
Pop	-0.0633	-0.1067	-0.0735^{*}	-0.1194
	(0.0387)	(0.0701)	(0.0400)	(0.0750)
z_{KFC}	-0.1768^{*}	-0.7275^{***}	-0.1187	-0.7817^{***}
	(0.0980)	(0.2539)	(0.1153)	(0.3041)
$\log(1+KFCDist)$	-0.2906	0.0596	-0.3062	-0.0430
	(0.5249)	(0.8240)	(0.5408)	(0.8629)
KFCStores	1.3846^{***}	2.8163^{***}	1.4129^{***}	2.9081^{***}
	(0.3886)	(0.5026)	(0.4215)	(0.5780)
Center	0.2533	-0.7767	0.3182	-0.8396
	(0.4099)	(0.7876)	(0.4166)	(0.8217)
Light	0.0708^{***}	0.1718^{***}	0.0764^{***}	0.1671^{**}
	(0.0220)	(0.0653)	(0.0232)	(0.0675)
Cinema	0.2050^{**}	0.2599^{*}	0.2050^{**}	0.2580^{*}
	(0.0942)	(0.1406)	(0.0946)	(0.1431)
McDStores	4.5179^{*}	4.1495	4.2024^{*}	4.1104^{*}
	(2.6878)	(2.7854)	(2.2869)	(2.4505)
$z_{McD} \cdot McDStores$	-0.3927^{**}	-0.3252	-0.4045^{**}	-0.2834
	(0.1805)	(0.2909)	(0.1915)	(0.3160)
$\log(1+s)$	1.0450	2.5363	1.0495	2.5670
	(1.0437)	(1.6632)	(1.1680)	(1.8435)
$Center \cdot McDStores$	2.0198	5.1759^{***}	1.6433	5.0406^{**}
	(1.4130)	(1.9751)	(1.5434)	(2.1493)
$\log(1+s) \times Center$	-3.5737^{*}	-6.0019^{**}	-4.0362^{*}	-6.6121^{**}
	(2.1388)	(2.6768)	(2.4003)	(3.0238)
$z_{McD} \times (1-McDStores)$			-0.0765	0.0688
			(0.0831)	(0.2086)
McD24h			1.2410	1.4169
			(1.0063)	(1.6220)
$Pr(LowType)$	0.7067^{***}		0.7169^{***}	
	(0.0452)		(0.0427)	
Regional Dummies (3 Regions)	Yes		Yes	
# of Unobserved Types	2		2	
log-likelihood	-1399.5		-1398.0	
Observations	1207			

Note: *, **, *** represent significant at significance level of 10%, 5% and 1% respectively.

However, such a test would be substantially difficult in other empirical games. As shown in the second column, coefficients on $z_{McD} \times (1 - McDStores)$ indicate that z_{McD} has a highly insignificant impact on KFC's decision when McD is absent in a market. This is consistent with the assumption that z_{McD} is a valid exclusion restriction. When McD owns zero outlets in a market, z_{McD} does not affect McD's store hours decision and consequently is independent of KFC's opening hours.

Unobserved market heterogeneity, if not adequately controlled, will invalidate my identification results and bias the estimates of payoff and belief functions. In this situation, McD's actual decision of store hours would reveal some information about unobserved market heterogeneity. Therefore, the number of McD's 24-hour stores should significantly affect KFC's decision after controlling for all other variables. However, the coefficients are highly insignificant, as shown in the second column. This indicates that assuming market-level unobservable to take two possible values is a good approximation.[30] In this section, these two values of unobserved heterogeneity are referred to as high- and low-type markets, respectively. High-type markets represents the ones where McD is more likely to open a 24-hour store.

Variable s is another exclusion restriction that facilitates the identification of KFC's base return and interactive effects. Its impact on KFC's opening hours decision differs by the locations of markets. For more details, recall that $Center$ is a dummy variable that equals 1 if the market is located at the center of a city (i.e. downtown area).[31] It influences KFC's decision in markets where McD is present but has limited impact in markets where McD is absent. It is the only demographic variable that has this feature. One plausible explanation for this result is that McD's decision has a heterogeneous impact on KFC that differs across market locations. The econometric model estimated in the next subsection formally investigates this conjecture.

4.4 Structural Estimation of Empirical Games

Even though the reduced form estimates shed light on KFC's choice incentive, they quantify neither the competitive effect nor KFC's belief. In order to capture these latter two effects, I estimate an econometric

[30] A model without market-level unobservable is highly rejected.

[31] In this paper, "city" refers to administrative areas in China. Typically, a city contains a core urban area and satellite towns. The core area is usually much more developed than the satellite towns, and there is generally rural area between them. Moreover, the distance between core area and satellite town is quite far, so it is implausible for a consumer to travel such a distance just for fast food. $Center$ equals 1 if the market lies in a core area.

model of games. Consider the second stage when each chain makes store hours decisions. KFC's base return is given by

$$\pi_{KFC}(\mathbf{x}, z_{KFC}, \omega, a_{KFC} = j) = \left(\mathbf{x}', \log(z_{KFC})\right)\alpha^{j} + \omega. \tag{10}$$

Recall that $a_{KFC} = j$ represents that KFC operates j stores at night and ω is the market-level unobserved heterogeneity. Vector $\mathbf{x}$ contains control variables in Table 3. The interactive effect is specified as

$$\delta_{KFC}(\mathbf{x}, s, a_{KFC} = j, a_{MCD} = 1) = \theta_1^j + \theta_2^j \log(1+s) + \theta_3^j Center + \theta_4^j \log(1+s) \times Center. \tag{11}$$

Intuitively, the interactive effect depends on s, the distance between the two chains' centroids. Furthermore, as shown in Table 4, *Center* has a significant impact on KFC only when McD is present in a market. Therefore, this factor may influence the interactive effect and is controlled for in equation (11). Other demographic variables are excluded in this interactive payoff to avoid imprecise estimates. As belief is an unknown, it multiplies by the interactive payoff. In this non-linear model, adding an additional parameter substantially affects estimation precision, compared to linear models. Moreover, other demographic variables are shown to have an insignificant impact on the interactive payoff under the equilibrium condition.

As specified in equations (10) and (11), the econometric models put no restrictions on coefficients across KFC's different actions a_{KFC}. This therefore captures economies of scale and cannibalization effects in a flexible way. Moreover, some cautions should be exercised in the interpretation of payoff functions. When night-time market and day-time market are independent, equations (10) and (11) are interpreted as KFC's profit function in the night market. In contrast, when operating at night has positive spillover effects on the day-time profit, these payoff functions represent the net increase of KFC's *total* profit (for both night-time and day-time) by operating 24-hour stores.

In this paper, I assume the private information $\varepsilon_i(a_i)$ follows type 1 extreme value distribution and is independent across actions and players. Therefore, I estimate a Logit model. Moreover, as I focus on KFC, McD's CCP of opening one 24-hour store is specified to take the following reduced form:

$$p^1_{MCD}(\mathbf{x}, z_{KFC}, z_{MCD}, s, \omega) = \frac{\exp\left[(\mathbf{x}', z_{KFC}, z_{MCD}, s, \mathbf{x}' \cdot z_{MCD}, \omega)\gamma\right]}{1 + \exp\left[(\mathbf{x}', z_{KFC}, z_{MCD}, s, \mathbf{x}' \cdot z_{MCD}, \omega)\gamma\right]}. \tag{12}$$

In the literature that studies empirical games under BNE, $p^1_{MCD}(\cdot)$ would be used to approximate KFC's belief. In contrast, this paper allows KFC to have a biased belief and treats KFC's belief as an unknown to be estimated. It also takes a Logit form as equation (12) but does not need to have the same parameters:

$$b^1_{KFC}(\mathbf{x}, z_{KFC}, z_{MCD}, s, \omega) = \frac{\exp\left[(\mathbf{x}', z_{KFC}, z_{MCD}, s, \mathbf{x}' \cdot z_{MCD}, \omega)\lambda\right]}{1 + \exp\left[(\mathbf{x}', z_{KFC}, z_{MCD}, s, \mathbf{x}' \cdot z_{MCD}, \omega)\lambda\right]}. \tag{13}$$

If $\gamma = \lambda$, this implies that KFC has unbiased expectations. In contrast, if $\gamma \neq \lambda$, this suggests KFC has biased beliefs. Given the identification results in Section 3 and the Logit form of belief, λ is identified. However, it would be imprecisely estimated due to small sample size. To reduce the estimation burden, I restrict equations (12) and (13) to share the same coefficients on $(\mathbf{x}, z_{KFC}, s, \omega)$. Such a restriction is equivalent to assuming that KFC has unbiased belief if $z_{MCD} = 0$. Therefore, imposing such a restriction would not over-reject the null hypothesis of KFC's unbiased beliefs. Finally, each chain's entry probabilities at the first stage also take a reduced Logit form.

Given the empirical model described above, a log-likelihood function of both chains' joint decisions on entry and store hours is well defined. I estimate the model using the Expectation Maximization Algorithm studied in Arcidiacono and Jones (2003).[32] In addition, as described above, my empirical model nests the restriction that KFC has an unbiased belief about McD's store hours decision. Therefore, it naturally yields a likelihood ratio test of KFC's correct belief.

Table 5 presents the estimated coefficients and marginal effects of the belief function. The column titled "McD's Choice" represents equation (12), estimated under equilibrium assumption. With such a constraint, it serves as an approximation of KFC's belief. The column titled "KFC's Belief" represents equation (13), estimated without unbiased belief constraint. Market characteristics in the interaction terms are taken as deviations from the sample means or median; therefore, the estimated effect of z_{MCD} represents its impact on an average market. Finally, a likelihood ratio test is conducted to test the null hypothesis of KFC's unbiased belief. This hypothesis is rejected at the 1% significance level, as indicated in the last row. Specifically, in an average market, KFC slightly under-predicts the probability that McD will open one 24-hour store (i.e. 0.5556 vs. 0.6325).

[32]In my empirical model, the parameters of KFC's payoff depend on McD's CCPs. Without unobserved heterogeneity, the log-likelihood function is additive separable between KFC's and McD's choice. Such a feature naturally implies a two-step estimator. However, the existence of unobserved heterogeneity destroys such additive separability. Finding the global maximum of the log-likelihood function is computationally burdensome. Arcidiacono and Jones (2003) proposes a sequential version of EM algorithm that is computationally efficient.

Table 5: Estimates of KFC's Belief (Logit Formula)

	McD's Choice		KFC's Belief	
	Coefficient	Marginal Effect	Coefficient	Marginal Effect
z_{McD}	-1.6108^{**}	-0.3744^{**}	-0.1912	-0.0472
	(0.6978)	(0.1841)	(0.8259)	(0.2055)
$z_{McD} \times (Income - \overline{Income})$	0.1826^{**}	0.0425^{*}	0.6076^{***}	0.1500^{***}
	(0.0914)	(0.0244)	(0.1443)	(0.0373)
$z_{McD} \times (KFCStores - Med[KFCStores])$	0.3214	0.0747	-1.1334^{*}	-0.2798^{**}
	(0.3236)	(0.0796)	(0.5876)	(0.1317)
Control Variables	Yes		Yes	
# of Unobserved Types	2		2	
Choice Probability at Average Market	0.6325		0.5556	
log-likelihood	-1397.8		-1389.6	
Unbiased Belief Test (p-value)	p=0.0009			
Observations	1207			

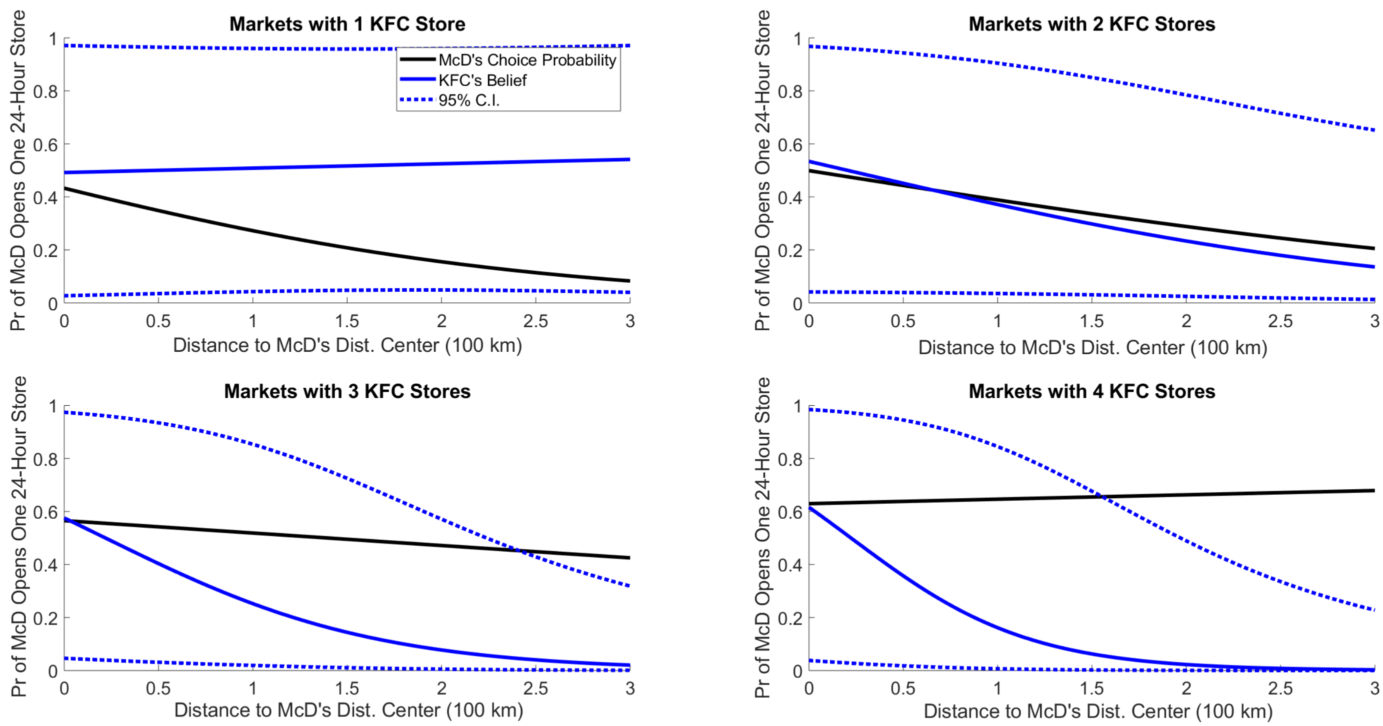
Note: $Med(\cdot)$ represents the sample median of corresponding variable. $Med(KFCStores) = 1$. Choice probability is calculated assuming unobserved heterogeneity is the high-type market. The choice probability is almost zero for the low-type market.

As shown in the column "McD's Choice", z_{McD} significantly decreases McD's probability of operating at night. Moreover, such a negative impact is alleviated when a market has higher average income. A comparison with the "KFC's Belief" column suggests that KFC over-evaluates the attenuation effect caused by higher *Income* as the interaction term exhibits a significantly larger magnitude.

To better understand how KFC's belief is affected by its network structure in a market, Figure 1 shows the impact of z_{McD} on McD's choice and KFC's belief for markets with different numbers of KFC stores. When KFC owns more outlets, higher delivery costs have a lesser impact on McD's decision as reflected by the flatter slope of the choice probability lines (i.e. black lines), moving from the top left graph to the bottom right one. As McD faces more competitive pressure, as measured by the number of KFC's outlets, it responds less to its own delivery cost and is likely to remain open at night. In contrast, KFC is over-optimistic in the sense that it believes a denser network structure is likely to kick McD out of the night market. This is reflected by the steeper slope of KFC's belief lines (i.e. blue lines), moving from the top left graph to the bottom right one. It leads KFC to substantially under-predict McD's opening hours when KFC owns more than two outlets.

Table 6 shows the estimates of KFC's payoff function. Consistent with Aguirregabiria and Magesan

Figure 1: KFC's Belief about McD's Probability of Operation at Night



(2017), these results suggest that incorrectly imposing the equilibrium assumption generates an attenuation bias on the interactive effect. In addition, Figure 2 plots the estimate of the interactive payoff as a function of s. It is clear that the magnitudes of the estimated interactive payoffs under an unrestricted belief specification are larger than the ones under the equilibrium specification. Even when such a bias is insignificant, falsely imposing unbiased expectations can still lead to incorrect conclusions. As shown in Figure 7 in Appendix A.5, the estimates of the interactive payoffs are insignificant under the equilibrium condition.[33] In contrast, Figure 2 shows that store hours are strategic complements in markets located in city centers, and they are strategic substitutes in markets belonging to satellite towns. For an average market located in a city center, the interactive effect is equivalent to a reduction in z_{KFC} from 130 km to 10 km.[34] Finally, the strategic complement nature of store hours suggests that extending operations to overnight service is a quality measure. Intuitively, the best response to a competitor's quality improvement is to increase one's own quality.[35] In contrast, the decision of store hours will exhibit strategic

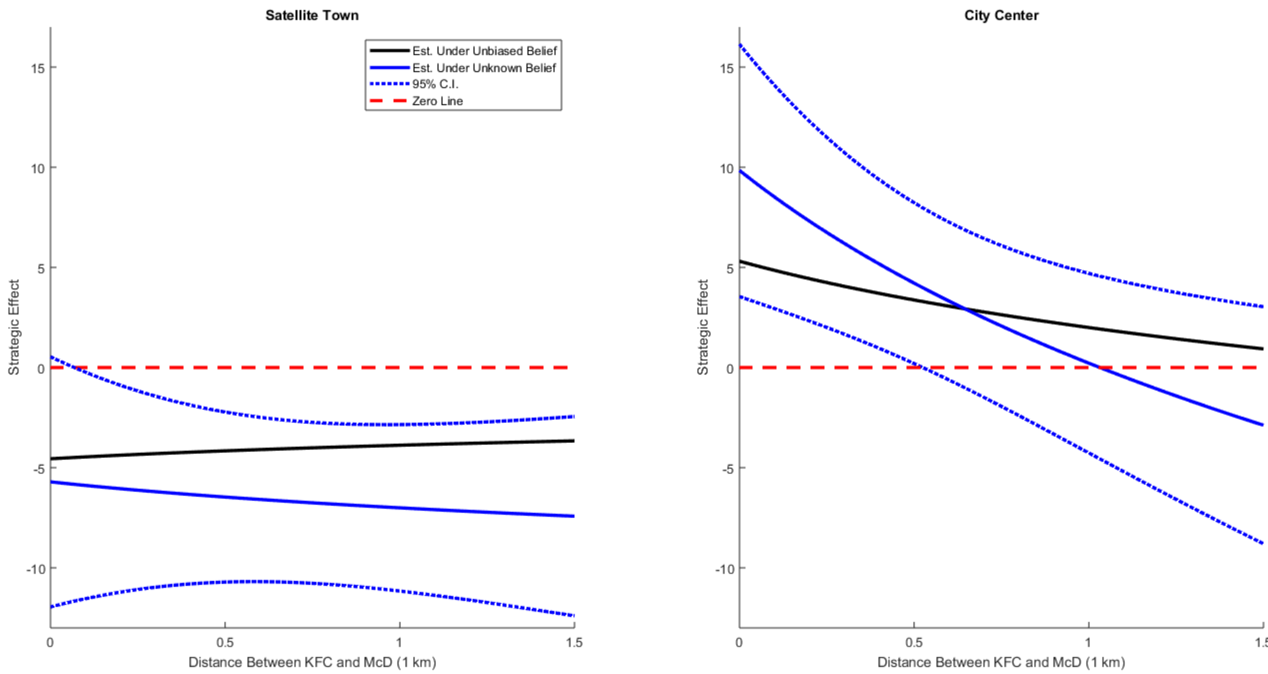
[33]The estimates under unbiased belief suggest that the interactive payoffs differ by the markets' locations. This is reflected by the highly significant coefficient on *Center*. However, it does not provide enough evidence to suggest which location (satellite town vs. city center) has significant non-zero strategic effect.

[34]According to the sample distribution, this reduction moves z_{KFC} from 30th percentile to its lowest value.

[35]In general, quality choice can exhibit either strategic substitute or complement depending on the model primitives. This can be seen in Brekke et al. (2010), who consider a general class of demand and cost functions to study firms' decisions on

substitute if it is horizontally differentiated.[36]

Figure 2: KFC's Interactive Payoff of Two 24-H Stores



In a recent analysis, Shen and Xiao (2014) study KFC and McD entry and expansion decisions in China from 1987 to 2007. They find that a chain's market presence has a spillover effect on the other chain's entry/expansion decision. They offer and quantify two explanations, namely demand expansion and market learning.[37] Though these factors play important roles in entry/expansion, I believe that their impact on store hours decisions is limited. Suppose KFC and McD have equal market shares. For a pure demand expansion factor to generate positive interactive payoff, McD's decision to operate overnight would need to more than double its market size, and it is very implausible for the night market to have such a large effect. Moreover, industry expert opinion suggests that chains can gather sufficient knowledge about market characteristics through their outlets, which suggests that the informational spillover effect from the other chain's decision is negligible. Consequently, despite the potential existence of spillover effect, its limited impact in the decision of store hours is unlikely to drive a positive strategic effect found in this paper. A more plausible interpretation of such a positive interactive payoff is the indirect business-stealing effect studied by Klemperer and Padilla (1997). Specifically, longer operation times by

price and quality under spatial competition.

[36]For an example, see Table 1 and 2 in Shy and Stenbacka (2008) and Figure 3 in Wenzel (2011).

[37]Demand expansion refers to the effect of a firm's market presence on boosting the market size. Market learning refers to the situation where a firm can infer market conditions by observing the other firm's decision.

Table 6: Estimates of KFC's Payoff Function

	Unbiased Belief		Unknown Belief	
	One 24-H	Two 24-H	One 24-H	Two 24-H
Interactive Payoff				
Constant	-2.0460	-4.5548	-0.7736	-5.7075^{*}
	(2.1108)	(2.9587)	(2.1132)	(3.1865)
$\log(1+s)$	3.4186^{*}	0.9755	1.1954	-1.8707
	(2.0366)	(3.2459)	(2.0570)	(4.2968)
$Center$	4.6724^{*}	9.8636^{***}	5.7414^{**}	15.5499^{***}
	(2.6650)	(3.8220)	(2.6451)	(4.6707)
$\log(1+s)\times Center$	-6.2695^{*}	-5.7532	-5.9877^{*}	-12.0136^{*}
	(3.3906)	(4.7331)	(3.5265)	(6.4380)
Base Return				
$\log(Income)$	0.3631	0.4969	0.3509	0.8213
	(0.2693)	(0.5099)	(0.2687)	(0.5826)
$\log(Pop)$	-0.0445	-0.0535	-0.0584	0.2009
	(0.3257)	(0.7757)	(0.3014)	(0.6394)
$\log(z_{KFC})$	-0.4094	-1.4867^{***}	-0.4120^{*}	-1.5280^{***}
	(0.2849)	(0.4817)	(0.2350)	(0.4767)
$\log(1+KFCDist)$	-0.2890	-0.1616	-0.1961	-0.0123
	(0.6523)	(1.2631)	(0.5654)	(0.9570)
$Center$	0.2491	-0.8952	0.4106	-0.9374
	(0.4093)	(1.1922)	(0.4076)	(0.8612)
$Light$	0.0609^{***}	0.1771^{**}	0.0624^{***}	0.2254^{***}
	(0.0223)	(0.0726)	(0.0225)	(0.0646)
$Cinema$	0.1397	0.2089	0.1287	0.2635
	(0.0950)	(0.2049)	(0.0970)	(0.2158)
Regional Dummies (3 Regions)	Yes		Yes	
# of Unobserved Types	2		2	
log-likelihood	-1397.8		-1389.6	
Observations	1207			

McD hurts KFC as it steals KFC's customers (i.e. both day-time and night-time consumers). However, this negative impact is alleviated if KFC also operates overnight. As shown by Klemperer and Padilla (1997), with this indirect business-stealing effect, firms have strong incentives to expand their operation times upon the deregulation of store hours. This could be harmful, however, for some firms, especially for small ones that are not able to easily extend their business hours. It is also anticipated to increase consumer welfare. However, from a social planner's perspective, the deregulation of store hours could generate social loss overall.

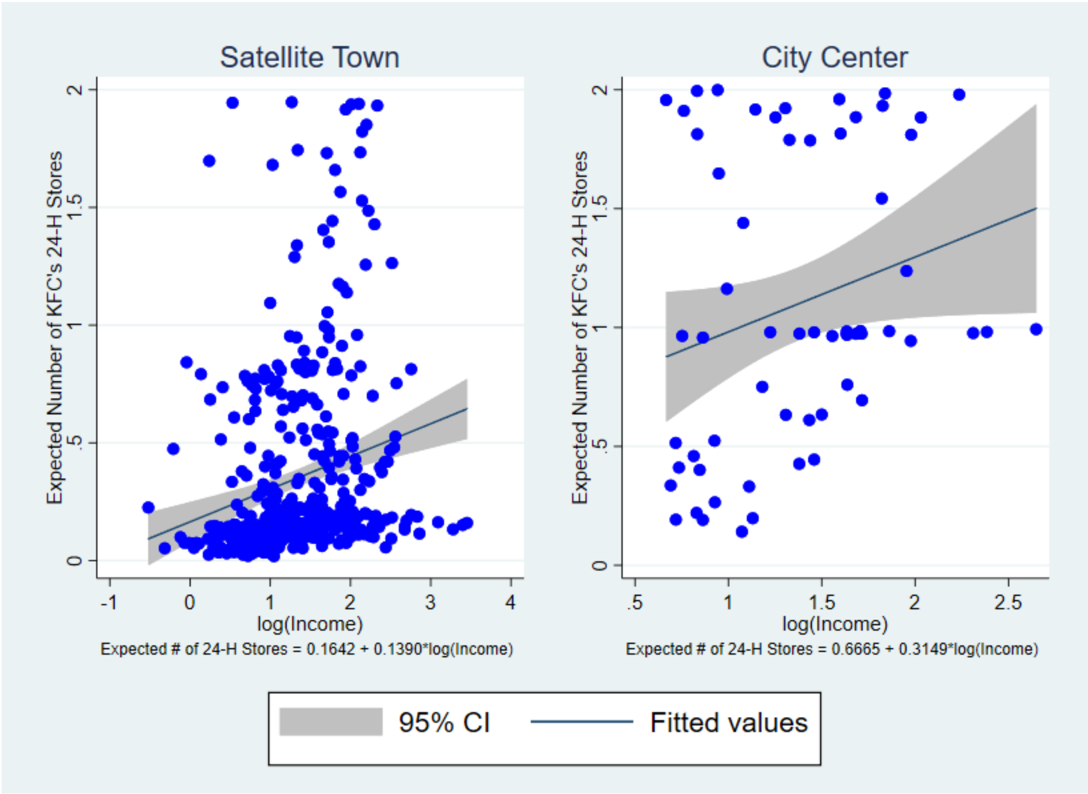
4.5 Counterfactual Analysis

To better understand how the evaluation of deregulation policy depends on the strategic nature of store hours, this subsection performs a counterfactual analysis. It also illustrates that incorrectly assuming unbiased belief would generate considerable bias for the counterfactual prediction. As shown in Table 2, there are 556 markets with KFC stores but without McD. Given my focus on the operating hours decision, these markets can be seen to have a regulation policy such that McD is forbidden to enter into the night market. The counterfactual analysis studies KFC's response if there were one McD store in those 556 markets. Analogously, it can be seen as a deregulation policy that lifts the restriction on McD's store hours.

In these 556 counterfactual markets, I assume that KFC's managers make the same mistakes as the markets where McD already exists. Therefore, the counterfactual analysis assumes KFC's belief function is equation (13) with estimates shown in Table 5.

Figure 3 shows the counterfactual prediction of the expected number of KFC's 24-hour stores from the above deregulation policy. On average, KFC will operate 1.10 non-stop service stores in the city center, while only 0.35 stores remain open at night in the satellite town. This suggests that the deregulation policy has considerable heterogeneous effects for different areas. This is because the strategic nature of store hours differs by the locations of markets. Specifically, store hours are strategic complements in the city center. Consequently, when McD is present and has some likelihood of operating a 24-hour store, KFC will have an extra incentive to open at night. In contrast, store hours are strategic substitutes in the satellite town; therefore, KFC is unwilling to operate longer hours when McD exists. Moreover, when

Figure 3: Expected Number of KFC's 24-Hour Stores



a market has higher average income, it increases McD's operation hours. In the city center where store hours are strategic complements, it introduces an extra incentive for KFC to operate longer. This explains why KFC is more sensitive to average income in the city center (i.e. the slope in the right plot is more than two times the slope in the left plot). Finally, the heterogeneous effects on different locations cannot be explained by the attractiveness of the city center, as coefficients of *Center* are highly insignificant in KFC's base return shown in Table 6.

Assuming McD has the same CCPs in counterfactual markets, Figure 4 shows the magnitude of KFC's belief bias, defined as KFC's belief minus McD's CCPs. On average, KFC over-predicts the probability that McD operates 24 hours by 14 percentage points. Moreover, the bias substantially depends on the market structure and characteristics. Consistent with Table 5, KFC over-evaluates the impact of higher income on McD and over-predicts McD's operation hours when markets are richer (left graph). In contrast, KFC mistakenly believes that a denser network structure would kick McD out of the night market; therefore, it underestimates McD's business hours when KFC owns more than two outlets (right graph).

Figure 4: Belief Bias (KFC's Belief minus McD's CCP)

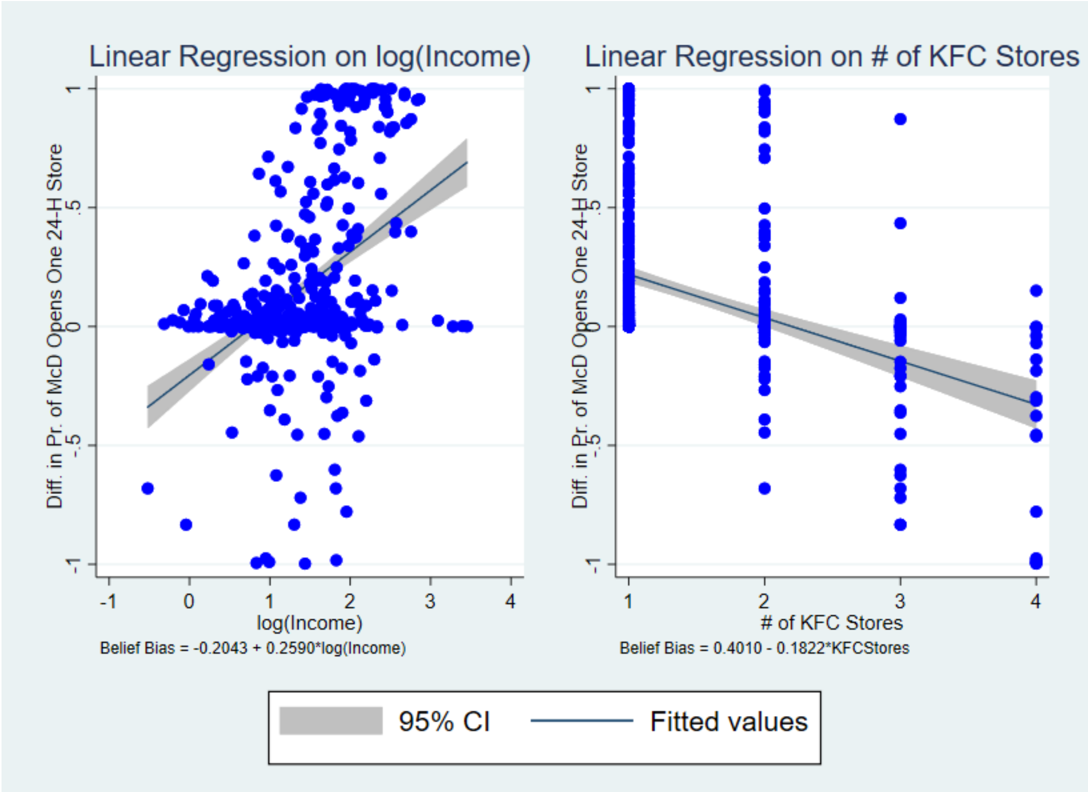


Figure 5: Regression of Prediction Bias on log(Income)

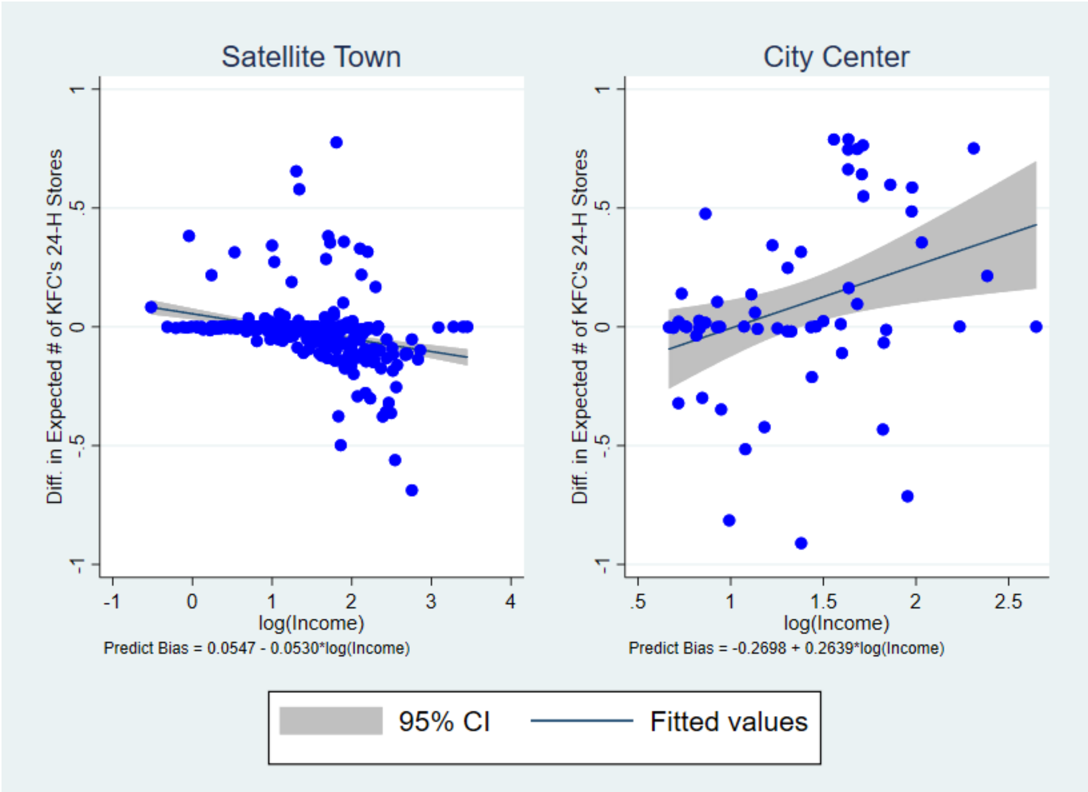
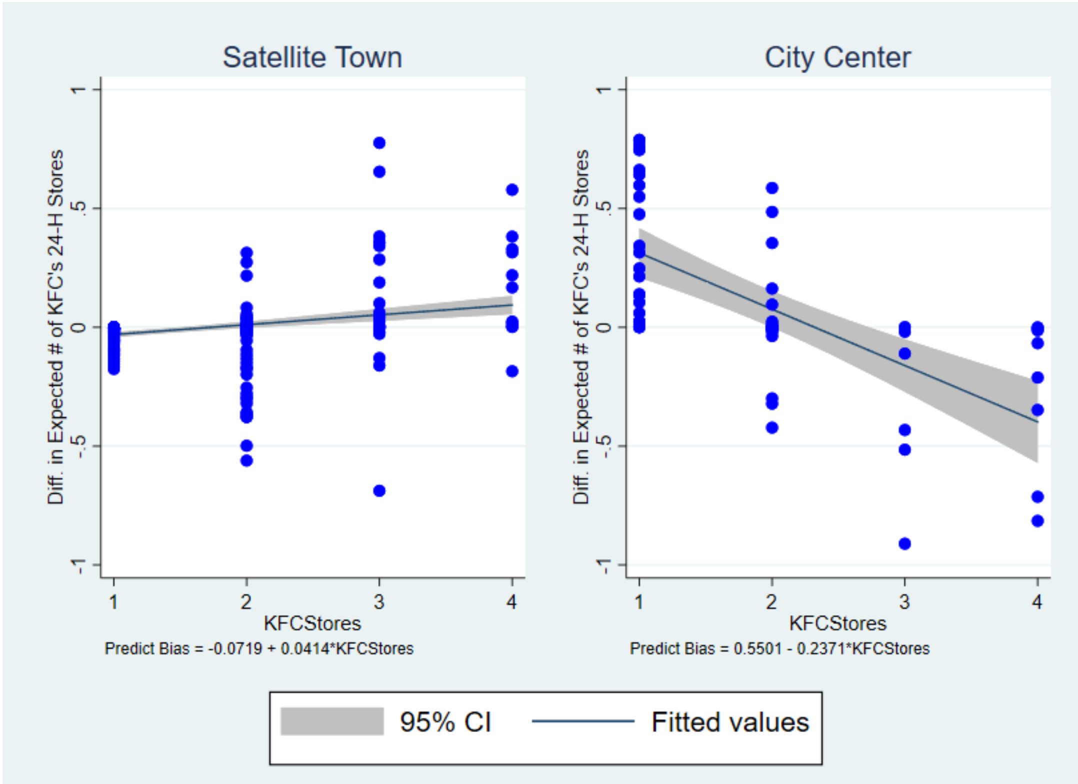


Figure 6: Regression of Prediction Bias on # of KFC Stores



As there is considerable bias in KFC's belief, assuming it has an unbiased expectation of McD's behaviors would yield an incorrect counterfactual prediction. Figures 5 and 6 present the prediction bias, defined as the prediction under unrestricted KFC's belief minus the prediction under unbiased expectation. As shown in Figure 4, KFC tends to over-predict McD's opening hours in markets with higher average income. In the satellite town where store hours are strategic substitutes, it implies that KFC would overestimate the negative impact caused by McD's operation at night. Consequently, compared with the prediction under unbiased belief, KFC would choose shorter business hours when markets are richer. This is reflected by a negative slope in the left plot of Figure 5. Conversely, as store hours are strategic complements in city center, KFC tends to operate longer in markets with higher average income, as reflected by a positive slope in the right graph of Figure 5. In addition, since the strategic effect has a larger magnitude in the city center, KFC responds to average income more sensitively (i.e. slope is larger in magnitude in the right graph). Finally, with the same logic, when KFC has more outlets in a market, it will under-predict McD's business hours. Consequently, KFC tends to over-operate in the satellite town and under-operate in the city center, as shown in Figure 6.

5 Conclusion

This paper studies the identification of an incomplete information game without imposing the Bayesian Nash Equilibrium. The econometric model imposes only weak assumptions on players' behaviors in the sense that each player's belief can be any probability distribution over the other player's action sets. With variation of action sets across players or/and across markets, I show that a player's payoff function is identified up to her belief at only one realization of the state variable. Furthermore, we can identify how a player adjusts her belief across different games; it naturally yields a testable restriction of players' unbiased beliefs.

This paper assumes that a player forms the same belief in two markets with same observables. It fails to include BNE when multiple equilibria are observed in the data. However, suppose a player's belief depends on a variable that is common to each player but unobserved to researchers. This unobservable can be seen as a sunspot variable that indexes different equilibria. Therefore, such a modification renders the nesting of multiple BNEs. In this extended model, my identification results hold trivially if players' choice probabilities conditional on the sunspot unobservable are identified. This identification result is established by Aguirregabiria and Mira (2018).[38]

Applying these identification results, this paper empirically studies the competition between KFC and McD on business hours. The hypothesis of KFC's unbiased belief is rejected. Moreover, the decision of store hours acts as strategic complement in the city center. It implies that the store hours decision is vertically differentiated and this feature is ignored in many of existing papers. When researchers aim to evaluate the deregulation policy that lifts restrictions on store hours, the ignorance of vertical differentiation would lead to at least two consequences. First, it will underestimate consumer welfare gains. Second, it will under-predict stores' business hours after the deregulation.

[38] Aguirregabiria and Mira (2018) require a game with at least three players. Even though the main context of this paper focuses on a two-player game, Appendix A.4 generalizes the results to multi-player games. Therefore, the conditions required in Aguirregabiria and Mira (2018) hold.

References

Aguirregabiria, V. and Magesan, A. (2017). Identification and estimation of dynamic discrete games when players' beliefs are not in equilibrium. *Working Paper*.

Aguirregabiria, V. and Mira, P. (2007). Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53.

Aguirregabiria, V. and Mira, P. (2018). Identification of games of incomplete information with multiple equilibria and unobserved heterogeneity. *Working Paper*.

Aguirregabiria, V. and Xie, E. (2017). Identification of biased beliefs in games of incomplete information using experimental data. *Working Paper*.

An, Y. (2017). Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. *Journal of Econometrics*, 200:326–343.

Aradillas-Lopez, A. (2010). Semiparametric estimation of a simultaneous game with incomplete information. *Journal of Econometrics*, 157(2):409–431.

Aradillas-Lopez, A. (2012). Pairwise-difference estimation of incomplete information games. *Journal of Econometrics*, 168(1):120–140.

Aradillas-Lopez, A. and Gandhi, A. (2016). Estimation of games with ordered actions: An application to chain-store entry. *Quantitative Economics*, 7(3):727–780.

Aradillas-Lopez, A. and Tamer, E. (2008). The identification power of equilibrium in simple games. *Journal of Business & Economic Statistics*, 26(3):261–283.

Arcidiacono, P. and Jones, J. B. (2003). Finite mixture distributions, sequential likelihood and the EM algorithm. *Econometrica*, 71(3):933–946.

Asker, J., Fershtman, C., Jeon, J., and Pakes, A. (2016). The competitive effects of information sharing. *Working Papers*.

Augereau, A., Greenstein, S., and Rysman, M. (2006). Coordination versus differentiation in a standard war: 56K modems. *Rand Journal of Economics*, 37(4):887–909.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Bajari, P., Hong, H., Krainer, J., and Nekipelov, D. (2010). Estimating static models of strategic interactions. *Journal of Business & Economic Statistics*, 28(4):469–482.

Besanko, D., Doraszelski, U., Lu, L. X., and Satterthwaite, M. (2010). On the role of demand and strategic uncertainty in capacity investment and disinvestment dynamics. *International Journal of Industrial Organization*, 28(4):383–389.

Brekke, K. R., Siciliani, L., and Straume, O. R. (2010). Price and quality in spatial competition. *Regional Science and Urban Economics*, 40(6):471–480.

Brock, W. A. and Durlauf, S. N. (2001). Discrete choice with social interactions. *Review of Economic Studies*, 68(2):235–260.

Brock, W. A. and Durlauf, S. N. (2007). Identification of binary choice models with social interactions. *Journal of Econometrics*, 140(1):52–75.

Ciliberto, F. and Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828.

Compiani, G. and Kitamura, Y. (2016). Using mixtures in econometric models: A brief review and some new results. *The Econometrics Journal*, 19(3):C95–C127.

Crawford, V. P. and Haller, H. (1990). Learning how to cooperate: Optimal play in repeated coordination games. *Econometrica*, 58(3):571–595.

De Paula, A. and Tang, X. (2012). Inference of signs of interaction effects in simultaneous games with incomplete information. *Econometrica*, 80(1):143–172.

Delgado, M. A. and Hidalgo, J. (2000). Nonparametric inference on structural breaks. *Journal of Econometrics*, 96(1):113–144.

Doraszelski, U., Lewis, G., and Pakes, A. (2018). Just starting out: Learning and equilibrium in a new market. *American Economic Review*, 108(3):565–615.

Ferris, J. S. (1990). Time, space, and shopping: The regulation of shopping hours. *Journal of Law Economics & Organization*, 6(1):171–187.

Fershtman, C. and Pakes, A. (2012). Dynamic games with asymmetric information: A framework for empirical work. *Quarterly Journal of Economics*, 127(4):1611–1661.

Gillen, B. (2010). Identification and estimation of level-k auctions. *Working Paper*.

Goeree, J. K. and Holt, C. A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*, 91(5):1402–1422.

Goldfarb, A. and Xiao, M. (2011). Who thinks about the competition? Managerial ability and strategic entry in US local telephone markets. *American Economic Review*, 101(7):3130–3161.

Gowrisankaran, G. and Krainer, J. (2011). Entry and pricing in a differentiated products industry: Evidence from the ATM market. *Rand Journal of Economics*, 42(1):1–22.

Hall, P., Neeman, A., Pakyari, R., and Elmore, R. (2005). Nonparametric inference in multivariate mixtures. *Biometrika*, 92(3):667–678.

Hall, P. and Zhou, X.-H. (2003). Nonparametric estimation of component distributions in a multivariate mixture. *The Annals of Statistics*, 31(1):201–224.

Heinemann, F., Nagel, R., and Ockenfels, P. (2009). Measuring strategic uncertainty in coordination games. *The Review of Economic Studies*, 76(1):181–221.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Henry, M., Kitamura, Y., and Salanié, B. (2014). Partial identification of finite mixtures in econometric models. *Quantitative Economics*, 5(1):123–144.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *Review of Economic Studies*, 60(3):497–529.

Hu, Y. (2017). The econometrics of unobservables: Applications of measurement error models in empirical industrial organization and labor economics. *Journal of Econometrics*, 200(2):154–168D.

Inderst, R. and Irmen, A. (2005). Shopping hours and price competition. *European Economic Review*, 49(5):1105–1124.

Jeon, J. (2018). Learning and investment under demand uncertainty in container shipping. *Working Papers*.

Kasahara, H. and Shimotsu, K. (2014). Non-parametric identification and estimation of the number of components in multivariate mixtures. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1):97–111.

Kashaev, N. (2016). Testing for Nash behavior in entry games with complete information. *Working Papers*.

Klemperer, P. and Padilla, A. J. (1997). Do firms' product lines include too many varieties? *Rand Journal of Economics*, 28(3):472–488.

Kline, B. and Tamer, E. (2012). Bounds for best response functions in binary games. *Journal of Econometrics*, 166(1):92–105.

Kügler, A. and Weiss, C. (2016). Time as a strategic variable: Business hours in the gasoline market. *Applied Economics Letters*, 23(15):1051–1056.

Lewbel, A. (2000). Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics*, 97(1):145–177.

Lewbel, A. and Tang, X. (2015). Identification and estimation of games with incomplete information using excluded regressors. *Journal of Econometrics*, 189(1):229–244.

Matzkin, R. L. (1992). Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica*, 60(2):239–270.

Morris, S. and Shin, H. S. (2002). Social value of public information. *American Economic Review*, 92(5):1521–1534.

Morris, S. and Shin, H. S. (2004). Coordination risk and the price of debt. *European Economic Review*, 48(1):133–153.

Müller, H. G. (1992). Change-points in nonparametric regression analysis. *Annals of Statistics*, 20(2):737–761.

Nishida, M. (2014). Estimating a model of strategic network choice: The convenience-store industry in okinawa. *Marketing Science*, 34(1):20–38.

Orhun, A. (2013). Spatial differentiation in the supermarket industry: The role of common information. *Quantitative Marketing and Economics*, 11(1):3–37.

Pakes, A., Ostrovsky, M., and Berry, S. (2007). Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *Rand Journal of Economics*, 38(2):373–399.

Pesendorfer, M. and Schmidt-Dengler, P. (2008). Asymptotic least squares estimators for dynamic games. *Review of Economic Studies*, 75(3):901–928.

Rennhoff, A. D. and Owens, M. F. (2012). Competition and the strategic choices of churches. *American Economic Journal Microeconomics*, 4(3):152–170.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *Rand Journal of Economics*, 37(3):619–640.

Shen, Q. and Xiao, P. (2014). McDonald's and KFC in China: Competitors or companions? *Marketing Science*, 33(2):287–307.

Shy, O. and Stenbacka, R. (2006). Service hours with asymmetric distributions of ideal service time. *International Journal of Industrial Organization*, 24(4):763–771.

Shy, O. and Stenbacka, R. (2008). Price competition, business hours and shopping time flexibility. *Economic Journal*, 118(531):1171–1195.

Su, C. L. (2014). Estimating discrete-choice games of incomplete information: Simple static examples. *Quantitative Marketing and Economics*, 12(2):167–207.

Sweeting, A. (2009). The strategic timing incentives of commercial radio stations: An empirical analysis using multiple equilibria. *Rand Journal of Economics*, 40(4):710–742.

Uetake, K. and Watanabe, Y. (2013). Estimating supermodular games using rationalizable strategies. *Advances in Econometrics*, 31:233–247.

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, 80(1):234–248.

Vitorino, M. A. (2009). Empirical entry games with complementarities: An application to the shopping center industry. *Journal of Marketing Research*, 49(2):175–191.

Wan, Y. and Xu, H. (2014). Semiparametric identification of binary decision games of incomplete information with correlated private signals. *Journal of Econometrics*, 182(2):235–246.

Wenzel, T. (2010). Liberalization of opening hours with free entry. *German Economic Review*, 11(4):511–526.

Wenzel, T. (2011). Deregulation of shopping hours: The impact on independent retailers and chain stores. *Scandinavian Journal of Economics*, 113(1):145–166.

Xiao, R. (2018). Identification and estimation of incomplete information games with multiple equilibria. *Journal of Econometrics*, 203(2):328–343.

Xu, H. (2014). Estimation of discrete games with correlated types. *Econometrics Journal*, 17(3):241–270.

Zhu, T. and Singh, V. (2009). Spatial competition with endogenous location choices: An application to discount retailing. *Quantitative Marketing and Economics*, 7(1):1–35.

A Appendix

A.1 Proofs

Proof of Proposition 1: According to equation (5), we then have the following equations for any two actions j and k of player i:

$$\pi_i(a_i = j) + \delta_i(a_i = j, a_{-i} = 1) \cdot b_i^1 = F_i^j(\mathbf{p}_i),$$
$$\pi_i(a_i = k) + \delta_i(a_i = k, a_{-i} = 1) \cdot b_i^1 = F_i^k(\mathbf{p}_i).$$

It is easy to see that we can cancel b_i^1 using the previous two equations. It yields Proposition 1 (a).

Now, suppose $J_{-i} = 0$ (i.e. player $-i$ has only one choice), then $b_i^1(J_{-i} = 0) = 0$ and equation (5) turns into

$$\pi_i(a_i = k) = F_i^k[\mathbf{p}_i(J_{-i} = 0)].$$

It yields identification of $\pi_i(a_i)$. Furthermore, for $J_{-i} = 1$, combining equation (5) and identification results of $\pi_i(a_i)$. It yields

$$\delta_i(a_i = k, a_{-i} = 1)b_i^1 = F_i^k[\mathbf{p}_i(J_{-i} = 1)] - F_i^k[\mathbf{p}_i(J_{-i} = 0)] \ \forall\ 0 \leq k \leq J_i.$$

Therefore, the perceived interactive effect $\delta_i(a_i = k, a_{-i} = 1)b_i^1$ is identified. The above equation characterizes the identified set for the interactive effect and each player's belief. □

Proof of Proposition 2: Given Proposition 1, we have the following equations for any two alternatives j and k:

$$\delta_i(a_i = j, a_{-i} = 1)b_i^1 = F_i^j[\mathbf{p}_i(J_{-i} = 1)] - F_i^j[\mathbf{p}_i(J_{-i} = 0)],$$
$$\delta_i(a_i = k, a_{-i} = 1)b_i^1 = F_i^k[\mathbf{p}_i(J_{-i} = 1)] - F_i^k[\mathbf{p}_i(J_{-i} = 0)].$$

Assuming $\delta_i(a_i, a_{-i}=1)b_i^1 \neq 0$, we can divide above equations and get

$$\frac{\delta_i(a_i=j, a_{-i}=1)}{\delta_i(a_i=k, a_{-i}=1)} = \frac{F_i^j[\mathbf{p}_i(J_{-i}=1)] - F_i^j[\mathbf{p}_i(J_{-i}=0)]}{F_i^k[\mathbf{p}_i(J_{-i}=1)] - F_i^k[\mathbf{p}_i(J_{-i}=0)]}.$$

Furthermore, as player i's belief is a valid probability distribution, we have $0 \leq b_i^1 \leq 1$. It consequently yields

$$\mathrm{sign}\{\delta_i(a_i=j, a_{-i}=1)\} = \mathrm{sign}\{F_i^j[\mathbf{p}_i(J_{-i}=1)] - F_i^j[\mathbf{p}_i(J_{-i}=0)]\},$$
$$\left|\delta_i(a_i=j, a_{-i}=1)\right| \geq \left|F_i^j[\mathbf{p}_i(J_{-i}=1)] - F_i^j[\mathbf{p}_i(J_{-i}=0)]\right|.$$

Finally, for any $J_i^{'}\ J_i^{''} \geq 1$, we have

$$\delta_i(a_i=j, a_{-i}=1)b_i^1(J_i') = F_i^j[\mathbf{p}_i(J_i', J_{-i}=1)] - F_i^j[\mathbf{p}_i(J_i', J_{-i}=0)],$$
$$\delta_i(a_i=j, a_{-i}=1)b_i^1(J_i'') = F_i^j[\mathbf{p}_i(J_i'', J_{-i}=1)] - F_i^j[\mathbf{p}_i(J_i'', J_{-i}=0)].$$

Dividing these two equations yields

$$\frac{b_i^1(J_i')}{b_i^1(J_i'')} = \frac{F_i^j[\mathbf{p}_i(J_i', J_{-i}=1)] - F_i^j[\mathbf{p}_i(J_i', J_{-i}=0)]}{F_i^j[\mathbf{p}_i(J_i'', J_{-i}=1)] - F_i^j[\mathbf{p}_i(J_i'', J_{-i}=0)]}.$$

This completes the proof. □

Proof of Proposition 3: Given player i's expected payoff in equation (6), we then have the following:

$$\pi_i(z_i, a_i=j) + \delta_i(z_i, a_i=j, a_{-i}=1)g_i(z_i, z_{-i}) = F_i^j(\mathbf{p}_i). \tag{14}$$

Given Assumption 6 (a), we can find at least two realizations of z_{-i}, say z_{-i}^1 and z_{-i}^2. Plug them separately into the above equation and subtract to cancel $\pi_i(z_i, a_i=k)$, and it becomes

$$\delta_i(z_i, a_i=j, a_{-i}=1)\cdot\left[g_i(z_i, z_{-i}^1) - g_i(z_i, z_{-i}^2)\right] = F_i^j\left[\mathbf{p}_i(z_i, z_{-i}^1)\right] - F_i^j\left[\mathbf{p}_i(z_i, z_{-i}^2)\right]\ \ \forall\, 0 \leq j \leq J_1.$$

Suppose $\mathbf{p}_i(z_i, z_{-i}^1) \neq \mathbf{p}_i(z_i, z_{-i}^2)$: it implies that $g_i(z_i, z_{-i}^1) - g_i(z_i, z_{-i}^2) \neq 0$.[39] Note that the difference of subjective expectations does not depend on which action is taken by player i. Therefore, for any two actions j and k, we can cancel the term $g_i(z_i, z_{-i}^1) - g_i(z_i, z_{-i}^2)$ by division and get

$$\frac{\delta_i(z_i, a_i = j, a_{-i} = 1)}{\delta_i(z_i, a_i = k, a_{-i} = 1)} = \frac{F_i^j\left[\mathbf{p}_i(z_i, z_{-i}^1)\right] - F_i^j\left[\mathbf{p}_i(z_i, z_{-i}^2)\right]}{F_i^k\left[\mathbf{p}_i(z_i, z_{-i}^1)\right] - F_i^k\left[\mathbf{p}_i(z_i, z_{-i}^2)\right]}.$$

Since the terms on the right-hand side are known, $\frac{\delta_i(z_i, a_i=j, a_{-i}=1)}{\delta_i(z_i, a_i=k, a_{-i}=1)}$ is identified. Furthermore, given multiplicative separable assumption 5, $\frac{\delta_i(z_i, a_i=j, a_{-i})}{\delta_i(z_i, a_i=k, a_{-i})} = \frac{\delta_i(z_i, a_i=j, a_{-i}=1)}{\delta_i(z_i, a_i=k, a_{-i}=1)}$ for any a_{-i}. This proves Proposition 3 (a).

Given the results in Proposition 1, we have the following equation for any two actions j and k:

$$\frac{F_i^j\left[\mathbf{p}_i(z_i, z_{-i}, s)\right] - \pi_i(z_i, a_i = j)}{F_i^k\left[\mathbf{p}_i(z_i, z_{-i}, s)\right] - \pi_i(z_i, a_i = k)} = \frac{\delta_i(z_i, s, a_i = j, a_{-i} = 1)}{\delta_i(z_i, s, a_i = k, a_{-i} = 1)}.$$

Under Assumption 6 (b), there must exist $s^1$and s^2, such that $\frac{\delta_i(z_i, s^1, a_i=j, a_{-i}=1)}{\delta_i(z_i, s^1, a_i=k, a_{-i}=1)} \neq \frac{\delta_i(z_i, s^2, a_i=j, a_{-i}=1)}{\delta_i(z_i, s^2, a_i=k, a_{-i}=1)}$; therefore, the previous equation turns into

$$\pi_i(z_i, a_i = j) - \tfrac{\delta_i(z_i, s^1, a_i=j, a_{-i}=1)}{\delta_i(z_i, s^1, a_i=k, a_{-i}=1)}\pi_i(z_i, a_i = k) = F_i^j\left[\mathbf{p}_i(z_i, z_{-i}, s^1)\right] - \tfrac{\delta_i(z_i, s^1, a_i=j, a_{-i}=1)}{\delta_i(z_i, s^1, a_i=k, a_{-i}=1)}F_i^k\left[\mathbf{p}_i(z_i, z_{-i}, s^1)\right],$$
$$\pi_i(z_i, a_i = j) - \tfrac{\delta_i(z_i, s^2, a_i=j, a_{-i}=1)}{\delta_i(z_i, s^2, a_i=k, a_{-i}=1)}\pi_i(z_i, a_i = k) = F_i^j\left[\mathbf{p}_i(z_i, z_{-i}, s^2)\right] - \tfrac{\delta_i(z_i, s^2, a_i=j, a_{-i}=1)}{\delta_i(z_i, s^2, a_i=k, a_{-i}=1)}F_i^k\left[\mathbf{p}_i(z_i, z_{-i}, s^2)\right].$$

Since $\frac{\delta_i(z_i, s, a_i=j, a_{-i}=1)}{\delta_i(z_i, s, a_i=k, a_{-i}=1)}$ is identified, this is a linear equation system containing two unknowns (i.e. $\pi_i(z_i, a_i = j)$ and $\pi_i(z_i, a_i = k)$) and two equations. Given $\frac{\delta_i(z_i, s^1, a_i=j, a_{-i}=1)}{\delta_i(z_i, s^1, a_i=k, a_{-i}=1)} \neq \frac{\delta_i(z_i, s^2, a_i=j, a_{-i}=1)}{\delta_i(z_i, s^2, a_i=k, a_{-i}=1)}$, $\pi_i(z_i, a_i = j)$ and $\pi_i(z_i, a_i = k)$ are uniquely determined through this system. In addition, according to equation (14), $\delta_i(z_i, s, a_i = k, a_{-i} = 1) \cdot g_i(z_i, z_{-i}, s)$ is identified thereafter. Next, for an action $l \neq j,\ k$, it is clear that $\delta_i(z_i, s, a_i = l, a_{-i} = 1) \cdot g_i(z_i, z_{-i}, s)$ is identified as $\delta_i(z_i, s, a_i = k, a_{-i} = 1) \cdot g_i(z_i, z_{-i}, s) \cdot \frac{\delta_i(z_i, s, a_i=l, a_{-i}=1)}{\delta_i(z_i, s, a_i=k, a_{-i}=1)}$. Finally, the researcher can uniquely determine the value of $\pi_i(z_i, s, a_i = l)$ according to equation (14), given the identification of $\delta_i(z_i, s, a_i = l, a_{-i} = 1) \cdot g_i(z_i, z_{-i}, s)$. This completes the proof. □

[39] Conditional on z_i, player i's payoff functions $\pi_i(z_i, a_i)$ and $\delta_i(z_i, a_i, a_{-i} = 1)$ are fixed; therefore, the only reason that player i's choice probability varies as z_{-i} varies is because player i's subjective expectation $g_i(z_i, z_{-i})$ varies as z_{-i} varies.

Proof of Proposition 4: Suppress $\mathbf{x}$ for notational simplicity as the identification results hold true for each $\mathbf{x}$. Define $\tilde{h}(\omega|\tilde{x},J_1,J_2)$ as the distribution of ω conditional on market structure (J_1,J_2). By definition

$$\tilde{h}(\omega|\tilde{x},J_1,J_2) = \frac{h(\omega|\tilde{x})p_{1,t_1}^{J_1}(\tilde{x},\omega)p_{2,t_1}^{J_2}(\tilde{x},\omega)}{\sum_{l=1}^{2} h(\omega^l|\tilde{x})p_{1,t_1}^{J_1}(\tilde{x},\omega^l)p_{2,t_1}^{J_2}(\tilde{x},\omega^l)}$$
$$= \frac{h(\omega|\tilde{x})p_{1,t_1}^{J_1}(\tilde{x},\omega)p_{2,t_1}^{J_2}(\tilde{x},\omega)}{Pr(J_1,J_2|\tilde{x})}. \quad (15)$$

First, consider any pair of $J_1,\ J_2 \geq 1$, and define the following matrices:

$$\boldsymbol{Pr}(\tilde{x},J_1,J_2) = \begin{bmatrix} Pr(a_1=0,a_2=0|\tilde{x},J_1,J_2) & Pr(a_1=0,a_2=1|\tilde{x},J_1,J_2) \\ Pr(a_1=1,a_2=0|\tilde{x},J_1,J_2) & Pr(a_1=1,a_2=1|\tilde{x},J_1,J_2) \end{bmatrix},$$
$$\mathbf{P}_{i,t_2}(J_1,J_2) = \begin{bmatrix} p_{i,t_2}^0(\omega^1,J_1,J_2) & p_{i,t_2}^0(\omega^2,J_1,J_2) \\ p_{i,t_2}^1(\omega^1,J_1,J_2) & p_{i,t_2}^1(\omega^2,J_1,J_2) \end{bmatrix},$$
$$Diag\big(\tilde{\mathbf{h}}(\tilde{x},J_1,J_2)\big) = \begin{bmatrix} \tilde{h}(\omega^1|\tilde{x},J_1,J_2) & 0 \\ 0 & \tilde{h}(\omega^2|\tilde{x},J_1,J_2) \end{bmatrix}.$$

As shown by Kasahara and Shimotsu (2014), the matrix $\boldsymbol{Pr}(\tilde{x},J_1,J_2)$ has full rank since ω can take on two possible values. Moreover, such a matrix is observed by econometricians. For any $\tilde{x}$, we have

$$\boldsymbol{Pr}(\tilde{x},J_1,J_2) = \mathbf{P}_{1,t_2}(J_1,J_2)\cdot Diag\big(\tilde{\mathbf{h}}(\tilde{x},J_1,J_2)\big)\cdot \mathbf{P}_{2,t_2}(J_1,J_2)'. \quad (16)$$

Consequently, for any two values $\tilde{x}^1$ and $\tilde{x}^2$, we have the following:

$$\boldsymbol{Pr}(\tilde{x}^1,J_1,J_2)\cdot\boldsymbol{Pr}(\tilde{x}^2,J_1,J_2)^{-1} = \mathbf{P}_{1,t_2}(J_1,J_2)\cdot Diag\big(\tilde{\mathbf{h}}(\tilde{x}^1,J_1,J_2)\big)\cdot Diag\big(\tilde{\mathbf{h}}(\tilde{x}^2,J_1,J_2)\big)^{-1}\cdot\mathbf{P}_{1,t_2}(J_1,J_2)^{-1}.$$

Therefore, $Diag\big(\tilde{\mathbf{h}}(\tilde{x}^1,J_1,J_2)\big)\cdot Diag\big(\tilde{\mathbf{h}}(\tilde{x}^1,J_1,J_2)\big)^{-1}$ represents a diagonal matrix with each element corresponding to an eigenvalue of matrix $\boldsymbol{Pr}(\tilde{x}^1,J_1,J_2)\cdot\boldsymbol{Pr}(\tilde{x}^2,J_1,J_2)^{-1}$. Moreover, $\mathbf{P}_{1,t_2}(J_1,J_2)$ is a matrix of eigenvector. Under condition that $\boldsymbol{Pr}(\tilde{x}^1,J_1,J_2)\cdot\boldsymbol{Pr}(\tilde{x}^2,J_1,J_2)^{-1}$ has full rank, it then has two distinct eigenvalues and each element in $\mathbf{P}_{1,t_2}(J_1,J_2)$ is identified up to a scaler.[40] In addition, matrix $\boldsymbol{Pr}(\tilde{x},J_1,J_2)$

[40]The full rank condition for $\boldsymbol{Pr}(\tilde{x}^1,J_1,J_2)\cdot\boldsymbol{Pr}(\tilde{x}^2,J_1,J_2)^{-1}$ is a weak condition. It requires diagonal elements of

and $\mathbf{P}_{i,t_2}(J_1,J_2)$ can be defined on any arbitrary pair of player i's action. It implies $p^j_{1,t_2}(\omega,J_1,J_2)$ is identified up to a scaler for each $j \leq J_1$. With the restriction that probability sums to unit, $p^j_{1,t_2}(\omega,J_1,J_2)$ is identified for any $J_1,\ J_2 \geq 1$. By the same argument, we can establish the identification of $p^j_{2,t_2}(\omega,J_1,J_2)$.

Given the identification of $p^j_{i,t_2}(\omega,J_1,J_2)$, it is easy to see that $\tilde{h}(\omega|\tilde{x},J_1,J_2)$ is also identified by equation (16) for any $J_1,\ J_2 \geq 1$. By the definition of $\tilde{h}(\omega|\tilde{x},J_1,J_2)$ in equation (15), $h(\omega|\tilde{x})p^{J_1}_{1,t_1}(\tilde{x},\omega)p^{J_2}_{2,t_1}(\tilde{x},\omega)$ is identified since $Pr(J_1,J_2|\tilde{x})$ is known by researchers. Consequently, for any $J_i',\ J_i'' \geq 1$, $p^{J_i'}_{i,t_1}(\tilde{x},\omega)/p^{J_i''}_{i,t_1}(\tilde{x},\omega)$ is identified. It implies that player i's first-stage conditional choice probability is identified up to two values, $p^0_{i,t_1}(\tilde{x},\omega)$ and $p^1_{i,t_1}(\tilde{x},\omega)$.

To identify $p^0_{i,t_1}(\tilde{x},\omega)$ and $p^1_{i,t_1}(\tilde{x},\omega)$, it suffices to consider four different market structures (J_1,J_2) where $J_i = 0,\ 1$. It implies the following restrictions:

$$\begin{aligned}
h(\omega^1|\tilde{x})p^1_{1,t_1}(\omega^1,\tilde{x})p^1_{2,t_1}(\omega^1,\tilde{x}) &= \tilde{h}(\omega^1|\tilde{x},J_1=1,J_2=1)Pr(J_1=1,J_2=1|\tilde{x}),\\
h(\omega^2|\tilde{x})p^1_{1,t_1}(\omega^2,\tilde{x})p^1_{2,t_1}(\omega^2,\tilde{x}) &= \tilde{h}(\omega^2|\tilde{x},J_1=1,J_2=1)Pr(J_1=1,J_2=1|\tilde{x}),\\
\sum_{l=1}^{2} h(\omega^l|\tilde{x})p^1_{1,t_1}(\omega^l,\tilde{x})p^0_{2,t_1}(\omega^l,\tilde{x}) &= Pr(J_1=1,J_2=0|\tilde{x}),\\
\sum_{l=1}^{2} h(\omega^l|\tilde{x})p^1_{1,t_1}(\omega^l,\tilde{x})p^0_{2,t_1}(\omega^l,\tilde{x})p^1_{1,t_2}(\omega^l,J_1=1,J_2=0) &= Pr(J_1=1,J_2=0,a_1=1|\tilde{x}),\\
\sum_{l=1}^{2} h(\omega^l|\tilde{x})p^0_{1,t_1}(\omega^l,\tilde{x})p^1_{2,t_1}(\omega^l,\tilde{x}) &= Pr(J_1=0,J_2=1|\tilde{x}),\\
\sum_{l=1}^{2} h(\omega^l|\tilde{x})p^0_{1,t_1}(\omega^l,\tilde{x})p^1_{2,t_1}(\omega^l,\tilde{x})p^1_{2,t_2}(\omega^l,J_1=0,J_2=1) &= Pr(J_1=0,J_2=1,a_2=1|\tilde{x}).
\end{aligned} \tag{17}$$

All terms on the right-hand side of equation system (17) are either observed or identified based on previous arguments. Denote $|\tilde{x}|$ as the number of support for $\tilde{x}$. Therefore, this equation system contains $6|\tilde{x}|$ restrictions. As shown on Table 7, there are $5|\tilde{x}|+4$ unknowns. Therefore, when $6|\tilde{x}| \geq 5|\tilde{x}|+4 \Leftrightarrow |\tilde{x}| \geq 4$, the order condition is satisfied. Moreover, no equation can be written as a linear combination of other equations. Therefore, all unknowns are identified.

Intuitively, consider the fourth equation in equation system (17). Since $\tilde{x}$ does not enter into $p^1_{1,t_2}(\cdot)$, it mimics the model studied by Henry et al. (2014) in which an exclusion restriction only affects the

$Diag\big(\tilde{\mathbf{h}}(\tilde{x}^1,J_1,J_2)\big)\cdot Diag\big(\tilde{\mathbf{h}}(\tilde{x}^2,J_1,J_2)\big)^{-1}$ to be different. Equivalently, it only requires $\tilde{x}$ to affect the distribution of ω conditional on market structure (J_1,J_2).

Table 7: Number of Unknowns

Unknowns	# of Unknowns
$h(\omega^1 \vert \tilde{x})$	$\vert\tilde{x}\vert$
$p^1_{i,t_1}(\omega,\tilde{x})$	$4\vert\tilde{x}\vert$
$p^1_{i,t_2}(\omega \vert J_i = 1, J_{-i} = 0)$	4

Note: $h(\omega^2|\tilde{x})$, $p^0_{i,t_1}(\omega,\tilde{x})$ and $p^0_{i,t_2}(\omega|J_i = 1, J_{-i} = 0)$ do not count as unknowns since they are perfectly determined by $h(\omega^1|\tilde{x})$, $p^1_{i,t_1}(\omega,\tilde{x})$ and $p^1_{i,t_2}(\omega|J_i = 1, J_{-i} = 0)$.

mixture weights. They have shown that such a model is point identified up to the mixture weights at two realizations of the exclusion restriction. Applying their results, suppose $h(\omega^1|\tilde{x}) \cdot p^1_{1,t_1}(\omega^1,\tilde{x}) \cdot p^0_{2,t_1}(\omega^1,\tilde{x})$ is known at two points of $\tilde{x}$, say $\tilde{x}^1$ and $\tilde{x}^2$. Then $h(\omega|\tilde{x})$ and $p^j_{i,t_1}(\omega,\tilde{x})$ are identified for each $\tilde{x}$ just using the first five equations in equation system (17). In addition, instead of considering the fourth equation, we can start with the last one in equation system (17) and replicate the same procedure. It also yields the point identification of $h(\cdot)$ and $p^j_{i,t_1}(\cdot)$. Finally, the estimated functions starting from two different equations must be the same for every $\tilde{x}$; these are additional restrictions that identify the weight mixture and CCPs at points $\tilde{x}^1$ and $\tilde{x}^2$.

Finally, by a similar argument, $p^j_{i,t_2}(\omega, J_i, J_{-i} = 0)$ is identified for any $J_i \geq 1$. This completes the proof. $\square$

A.2 Relaxation of Known Distribution of $G_i(\cdot)$

In the main text, player i's private information is assumed to be independent across players and independent of public information $\mathbf{x}$. Moreover, the distribution of this private information is assumed to be known by researchers. The commonly used distributional assumptions in practice include i.i.d. type 1 extreme value distribution (i.e. Logit Model) and i.i.d. standard normal distribution (i.e. Multinomial Probit Model). These assumptions are restrictive in the sense that they restrict the private information among actions to be independent and homoscedastic. However, the distributional assumption of ε_i can be relaxed to capture heteroskedasticity and potential correlation among actions in a fairly flexible way. This subsection formally establishes this point with the help of exclusion restriction z_i and variation in players' action sets. First, consider an assumption 1' that is a weaker version of Assumption 1.

Assumption 1′. *(a) For each* $i = 1,2$, $\varepsilon_i = \big(\varepsilon_i(0), \varepsilon_i(1), \cdots, \varepsilon_i(J_i)\big)'$ *follows a CDF* $G_i(\cdot;\boldsymbol{\beta}_{i,x})$ *that is* *absolutely continuous with respect to Lebesgue measure in* $\mathbb{R}^{J_i+1}$. $\boldsymbol{\beta}_{i,x} = (\beta_{1,i,x}, \cdots, \beta_{L_i,i,x})'$ *is a vector of parameters with* $L_i < \infty$ *dimensions. Moreover, researchers know the functional form of* $G_i(\cdot)$ *but not the parameters* $\boldsymbol{\beta}_{i,x}$ *for* $i = 1,2$.

(b) ε_i *is independent of* ε_{-i} *conditional on* $(\mathbf{x}, z_1, z_2)$. *Moreover, conditional on* $\mathbf{x}$, ε_i *is independent of* (z_1, z_2) *for* $i = 1,2$.

Assumption $1'$ parametrizes the distribution of ε_i by a vector $\boldsymbol{\beta}_{i,x}$ which is unknown by researchers. Such a parametrization relaxes Assumption 1 in several directions. First, $\boldsymbol{\beta}_{i,x}$ can contain the standard deviation of $\varepsilon_i(a_i)$ for different a_i and potential correlation between $\varepsilon_i(j)$ and $\varepsilon_i(k)$ for $j \neq k$. As a consequence, it captures the possible heteroskedasticity and correlation of private information among different actions in a fairly flexible way. Second, Assumption $1'$ allows the distribution of ε_i to be correlated with $\mathbf{x}$, and such a correlation is captured by the dependence of $\boldsymbol{\beta}_{i,x}$ on $\mathbf{x}$.[41] Since I investigate identification conditional on $\mathbf{x}$, I write $\boldsymbol{\beta}_{i,x}$ as $\boldsymbol{\beta}_i$ for notation simplicity. Note that the identification results do not require exclusion restriction s.

An implication of Assumption $1'$ is that the inverse of best response probability function $F_i(\cdot)$ will depend on $\boldsymbol{\beta}_i$. We then have the following equations:

$$\pi_i(z_i, a_i = j) + \sum_{k=1}^{J_{-i}} \delta_i(z_i, a_i = j, a_2 = k) \cdot b_i^k(z_i, z_{-i}) = F_i^j\left[\mathbf{p}_i(z_i, z_{-i}); \boldsymbol{\beta}_i\right] \;\; \forall\, 0 \leq j \leq J_1.$$

Given Assumption $1'$, $F_i(\cdot, \boldsymbol{\beta}_i)$ is known by researchers up to the unknown parameters $\boldsymbol{\beta}_i$. Addition-

[41] Lewbel and Tang (2015) generalize the *special regressor* approach considered in Matzkin (1992) and Lewbel (2000) to a two-player binary choice game with incomplete information. They show that if researchers can observe a variable that affects a player's payoff linearly, then the distribution of the error term is non-parametrically identified under the Bayesian Nash Equilibrium condition. Despite its power, this special regressor approach does not work in my framework because I non-parametrically specify the payoff function; as a result, I parametrize the function $G_i(\cdot)$.

ally, Assumption 6 (a) implies that there exist h realizations of z_{-i}, say z_{-i}^1, z_{-i}^2 up to z_{-i}^h, such that

$$\begin{aligned}
F_i^1\left[\mathbf{p}_{i,J_{-i}=0}(z_i,z_{-i});\boldsymbol{\beta}_i\right] &= \pi_i(z_i,a_i=1),\\
&\vdots\\
F_i^{J_i}\left[\mathbf{p}_{i,J_{-i}=0}(z_i,z_{-i});\boldsymbol{\beta}_i\right] &= \pi_i(z_i,a_i=J_i),\\
F_i^1\left[\mathbf{p}_{i,J_{-i}=1}(z_i,z_{-i}^1);\boldsymbol{\beta}_i\right] &= \pi_i(z_i,a_i=1)+\delta_i(z_i,a_i=1,a_{-i}=1)\cdot b_i^1(z_i,z_{-i}^1),\\
F_i^2\left[\mathbf{p}_{i,J_{-i}=1}(z_i,z_{-i}^1);\boldsymbol{\beta}_i\right] &= \pi_i(z_i,a_i=2)+\delta_i(z_i,a_i=2,a_{-i}=1)\cdot b_i^1(z_i,z_{-i}^1),\\
&\vdots\\
F_i^{J_i}\left[\mathbf{p}_{i,J_{-i}=1}(z_i,z_{-i}^h);\boldsymbol{\beta}_i\right] &= \pi_i(z_i,a_i=J_i)+\delta_i(z_i,a_i=J_i,a_{-i}=1)\cdot b_i^1(z_i,z_{-i}^h).
\end{aligned} \tag{18}$$

Equation system (18) consists of $J_i(h+1)$ equations with $2J_i+h-1+L_i$ unknowns.[42] A necessary order condition for identification is $J_i(h+1)\geq 2J_i+h-1+L_i$, which yields $(J_i-1)(h-1)\geq L_i$. Moreover, denote $\mathbf{F}(z_i,\mathbf{z}_{-i}^{1:h},J_i;\boldsymbol{\beta}_i)=(F_i^1\left[\mathbf{p}_{i,J_{-i}=0}(z_i,z_{-i});\boldsymbol{\beta}_i\right],\cdots,F_i^{J_i}\left[\mathbf{p}_{i,J_{-i}=1}(z_i,z_{-i}^h);\boldsymbol{\beta}_i\right])'$ as a $J_i(h+1)\times 1$ vector of inversion of choice probability, and the following Assumption 8 establishes a sufficient condition for the identification of $\boldsymbol{\beta}_i$.

Assumption 8. *Conditional on $(\mathbf{x},J_i>1,z_i)$, there are $h\geq 2$ realizations of z_2 such that $(J_i-1)(h-1)\geq L_i$. Moreover, let $\frac{\partial\mathbf{F}(z_i,\mathbf{z}_{-i}^{1:h},J_i;\boldsymbol{\beta}_i)}{\partial\boldsymbol{\beta}_i}$ be a Jacobian matrix with dimension $J_i(h+1)\times L_i$; such a matrix has column rank L_i.*

Proposition 5. *Under Assumption 6 (a), 8 and conditions met in Proposition 1 with Assumption 1 replaced by Assumption 1′, $\boldsymbol{\beta}_i$ is identified in a neighborhood of its true value.*

Proof. $\boldsymbol{\beta}_i$ does not enter into the right-hand side of equation system (18); as shown in Subsection 3.3, all unknowns on the right-hand side are identified if researchers know the value of $\boldsymbol{\beta}_i$. Since there are $2J_i+h-1$ unknowns on the right-hand side, there still remain $(J_i-1)(h-1)\geq L_i$ restrictions that can

[42] There are J_i unknowns for $\pi_i(z_i,a_i=j)\ \forall\ 1\leq j\leq J_1$, h unknowns for $\delta_i(z_i,a_i=1,a_{-i}=1)\cdot b_i^1(z_i,z_{-i})$, (J_i-1) unknowns for $\frac{\delta_i(z_i,a_i=j,a_{-i}=1)}{\delta_i(z_i,a_i=1,a_{-i}=1)}\ \forall\ 1<j\leq J_i$ and L_i unknowns for $\boldsymbol{\beta}_i$. Note that as shown in Subsection 3.3, only $\delta_i(z_i,a_i=1,a_{-i}=1)\cdot b_i^1(z_i,z_{-i})$ is identified while $\delta_i(z_i,a_i=1,a_{-i}=1)$ and $b_i^1(z_i,z_{-i})$ are not distinguishable from each other, so I treat the perceived interactive effects as unknowns. In addition, $\delta_i(z_i,a_i=1,a_{-i}=1)\cdot b_i^1(z_i,z_{-i})$ and $\frac{\delta_i(z_i,a_i=j,a_{-i}=1)}{\delta_i(z_i,a_i=1,a_{-i}=1)}$ perfectly determine the value of $\delta_i(z_i,a_i=j,a_{-i}=1)\cdot b_i^1(z_i,z_{-i})$ and therefore $\delta_i(z_i,a_i=j,a_{-i}=1)\cdot b_i^1(z_i,z_{-i})\ \forall\ j\neq 1$ does not count as an unknown.

be exploited to identify β_i. Under Assumption 8, both the rank and order condition are satisfied and therefore β_i is locally identified. $\square$

Assumption 8 is a generic assumption. Suppose β_i is identified when researchers perfectly know player i's payoff and belief, then the Jacobian matrix $\frac{\partial \mathbf{F}(z_i,\mathbf{z}_{-i}^{1:h},J_i;\beta_i)}{\partial \beta_i}$ will have full column rank with probability one. This implies that researchers can capture heteroskedasticity and correlation as flexible as in the standard discrete choice model.

A.3 Identification Results when $J_{-i} > 1$

In this section, I present the identification result of player i's payoff function when player $-i$ has more than two actions. As in the main text, the distribution of private information is assumed to be known by the researcher in this section.

For some $k \le J_i - J_{-i}$, define a $J_{-i} \times J_{-i}$ matrix of interaction effect $\boldsymbol{\Delta}_i^{k:J_{-i}+k-1}$ as

$$\begin{bmatrix} \delta_i(a_i=k,a_{-i}=1), & \delta_i(a_{-i}=k,a_{-i}=2), & \cdots, & \delta_i(a_i=k,a_{-i}=J_{-i}) \\ \delta_i(a_i=k+1,a_{-i}=1), & \delta_i(a_{-i}=k+1,a_{-i}=2), & \cdots, & \delta_i(a_i=k+1,a_{-i}=J_{-i}) \\ \vdots & \vdots & \ddots & \vdots \\ \delta_i(a_i=k+J_{-i}-1,a_{-i}=1), & \delta_i(a_i=k+J_{-i}-1,a_{-i}=2) & \cdots, & \delta_i(a_i=k+J_{-i}-1,a_{-i}=J_{-i}) \end{bmatrix}.$$

Moreover, let $\boldsymbol{\pi}_i^{k:J_{-i}+k-1} = \big(\pi_i(a_i=k),\cdots,\pi_i(a_i=k+J_{-i}-1)\big)'$ and $\mathbf{F}_i^{k:J_{-i}+k-1} = \big(F_i^k(\mathbf{p}_i),\cdots,F_i^{k+J_{-i}-1}(\mathbf{p}_i)\big)'$, and we then have following proposition.

Proposition 6. *(a) Under Assumptions 1 to 4 and suppose $\boldsymbol{\Delta}^{k:J_{-i}+k-1}$ is invertible for any k and data contain observations with $J_i > J_{-i}$; then for any $k',k \le J_i - J_{-i}$, the identified set of player i's payoff is given by the set of values that satisfies the following restrictions:*

$$\big[\boldsymbol{\Delta}_i^{k:J_{-i}+k-1}\big]^{-1}\big[\mathbf{F}_i^{k:J_{-i}+k-1} - \boldsymbol{\pi}_i^{k:J_{-i}+k-1}\big] = \big[\boldsymbol{\Delta}_i^{k':J_{-i}+k'-1}\big]^{-1}\big[\mathbf{F}_i^{k':J_{-i}+k'-1} - \boldsymbol{\pi}_i^{k':J_{-i}+k'-1}\big].$$

(b) Suppose further that data also contain observations with $J_{-i}=0$, then $\pi_i(a_i=k)$ is identified by $F_i^k[\mathbf{p}_i(J_{-i}=0)]$ $\forall$ k. Furthermore, the identified set of player i's interactive effect and belief is given by

the set of values that satisfies the following restriction:

$$\sum_{j=1}^{J_{-i}} \delta_i(a_i = k, a_{-i} = j) b_i^j = F_i^k[\mathbf{p}_i(J_{-i} = 1)] - F_i^k[\mathbf{p}_i(J_{-i} = 0)].$$

Proof. By construction, we have

$$\boldsymbol{\pi}_i^{k:J_{-i}+k-1} + \boldsymbol{\Delta}_i^{k:J_{-i}+k-1} \mathbf{b}_i = \mathbf{F}_i^{k:J_{-i}+k-1}.$$

Under the assumption such that $\boldsymbol{\Delta}_i^{k:J_{-i}+k-1}$ is invertible, we then have the vector of belief equals to

$$\mathbf{b}_i = \left[\boldsymbol{\Delta}_i^{k:J_{-i}+k-1}\right]^{-1} \left[\mathbf{F}_i^{k:J_{-i}+k-1} - \boldsymbol{\pi}_i^{k:J_{-i}+k-1}\right].$$

Similarly, for another value of k', we have $\mathbf{b}_i = \left[\boldsymbol{\Delta}_i^{k':J_{-i}+k'-1}\right]^{-1} \left[\mathbf{F}_i^{k':J_{-i}+k'-1} - \boldsymbol{\pi}_i^{k':J_{-i}+k'-1}\right]$. Consequently, it yields

$$\left[\boldsymbol{\Delta}_i^{k:J_{-i}+k-1}\right]^{-1} \left[\mathbf{F}_i^{k:J_{-i}+k-1} - \boldsymbol{\pi}_i^{k:J_{-i}+k-1}\right] = \left[\boldsymbol{\Delta}_i^{k':J_{-i}+k'-1}\right]^{-1} \left[\mathbf{F}_i^{k':J_{-i}+k'-1} - \boldsymbol{\pi}_i^{k':J_{-i}+k'-1}\right].$$

Furthermore, if there exist observations with $J_{-i} = 0$, then Proposition 6 (b) follows a similar proof as Proposition 1 (b) and is omitted. □

A.4 Identification in Games with Multiple Actions, Multiple Players

Consider a game with N players where $N > 2$. Player is indexed by $i, n \in \{1, 2, \cdots, N\}$. Each player i has an action set $A_i = \{0, 1, \cdots, J_i\}$. Consequently, Cartesian product $A = A_1 \times A_2 \cdots \times A_N$ represents the space of action profiles in this game. Each player i simultaneously chooses an action a_i from her action set A_i . Let $\mathbf{a} = (a_1, a_2, \cdots, a_N) \in A$ be a realized outcome or action profile in this game. Player i's payoff under $\mathbf{a}$ is

$$\Pi_i(\mathbf{x}, \varepsilon_i, \mathbf{a}) = \pi_i(\mathbf{x}, a_i) + \sum_{n=1, n \neq i}^{N} \delta_{i,n}(\mathbf{x}, a_i, a_n) \cdot \mathbb{1}(a_n \neq 0) + \varepsilon_i(a_i). \tag{19}$$

The term $\delta_{i,n}(\mathbf{x}, a_i, a_n)$ in equation (19) represents player n's impact on player i. It allows different players to have heterogeneous interactive effects on player i but requires they are additive separable. When there is variation in players' choice sets, we can identify π_i and $\delta_{i,n}$ by studying the situation that $J_{i'} = 0 \; \forall \; i' \neq i, \; n$. This is essentially the game with two players considered in the main text and all identification results directly follow.

When players' action sets are constant across observations, I consider a restriction on the payoff function that allows for dimension reduction similar to Assumption 5:

$$\delta_{i,n}(\mathbf{x}, a_i, a_n) = \delta_{i,n}(\mathbf{x}, a_i, a_n = 1) \cdot \eta_{i,n}(\mathbf{x}, a_n) \text{ where } \eta_{i,n}(\mathbf{x}, a_n = 1) = 1.$$

Consequently, player i's expected payoff for action a_i is

$$\begin{aligned} E[\Pi_i(\mathbf{x}, \varepsilon_i, a_i)] &= \pi_i(\mathbf{x}, a_i) + \sum_{n=1, n\neq i}^{N} \delta_{i,n}(\mathbf{x}, a_i, a_n = 1) \big[\sum_{j=1}^{J_n} \eta_{i,n}(\mathbf{x}, a_n = j) b_{i,n}^{j}(\mathbf{x}) \big] + \varepsilon_i(a_i) \\ \Rightarrow E[\Pi_i(\mathbf{x}, \varepsilon_i, a_i)] &= \pi_i(\mathbf{x}, a_i) + \sum_{n=1, n\neq i}^{N} \delta_{i,n}(\mathbf{x}, a_i, a_n = 1) \cdot g_{i,n}(\mathbf{x}) + \varepsilon_i(a_i), \end{aligned}$$

where $b_{i,n}^{j}(\mathbf{x})$ represents player i's belief about the probability that player n will choose action $a_n = j$. Term $g_{i,n}(\mathbf{x}) = \sum_{j=1}^{J_n} \eta_{i,n}(\mathbf{x}, a_n = j) b_{i,n}^{j}(\mathbf{x})$ represents player i's subjective expected value of $\eta_{i,n}$. Given the distribution of ε_i, we can invert player i's conditional choice probability:

$$\pi_i(\mathbf{x}, a_i = j) + \sum_{n=1, n\neq i}^{N} \delta_{i,n}(\mathbf{x}, a_i, a_n = 1) \cdot g_{i,n}(\mathbf{x}) = F_i^{j} \big[\mathbf{p}_i(\mathbf{x}) \big].$$

Suppose that $N \leq \min\{J_1, \cdots, J_N\}$ (i.e. the number of players is smaller than the number of actions): the above equation shares the same structure as the asymmetric game described in Appendix A.3: for instance, a game with two players where player i has $J_i + 1$ actions while player $-i$ has N actions. Therefore, the result of Proposition 6 (a) applies. Furthermore, when there exist exclusion restrictions z_i and s, the base return $\pi_i(\cdot)$ and the perceived interactive payoff $\sum_{n=1, n\neq i}^{N} \delta_{i,n}(\cdot) g_{i,n}(\cdot)$ are point identified.

A.5 Data Construction and Supplementary Tables

Construction of Cinema Clusters

I obtain a list of cinemas, including address, for each county. I then construct clusters of cinemas by the following algorithm:

1. Start with an arbitrary cinema denoted by $cinema_i$ which is not assigned into any cluster, and assign this cinema to a new cluster.

2. Draw a radius of 2 km around such a $cinema_i$. If no other cinema exists in this area, the algorithm terminates and we start a new unassigned cinema by step 1. If some cinemas exist in this area, assign those cinemas in the same cluster as $cinema_i$.

3. For each newly assigned cinema, repeat step 2.

Supplementary Tables and Graphs

Figure 7: KFC's Interactive Payoff of Two 24-H Stores

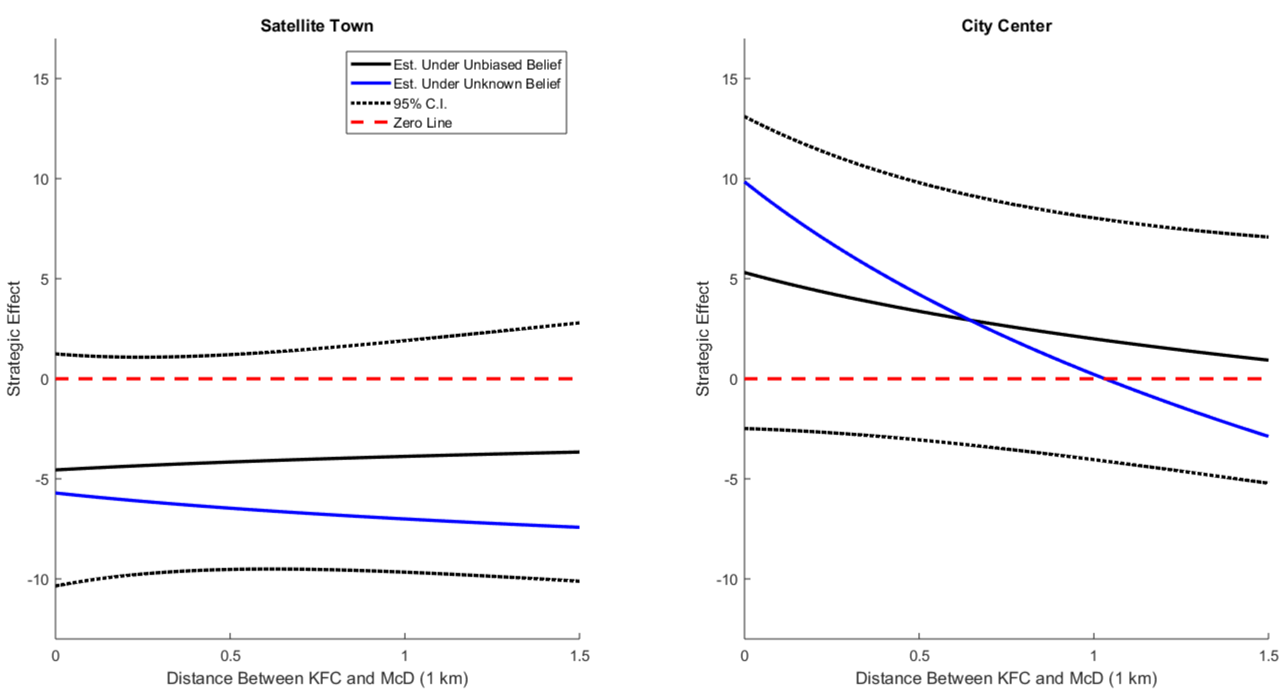


Table 8: Reduced Form Logit Regression: Number of McD's 24-Hour Stores

	One 24 H
$Income$	-0.1872
	(0.1946)
Pop	0.6420^{***}
	(0.1949)
$z_{KFC} \times \mathbb{1}(KFCStores > 0)$	-0.2096
	(0.4719)
$z_{McD} \times \mathbb{1}(KFCStores = 0)$	0.1919
	(0.4196)
$z_{McD} \times \mathbb{1}(KFCStores > 0)$	-1.0361^{**}
	(0.4454)
$z_{McD} \times (Income - \overline{Income})$	0.1752^{**}
	(0.0785)
$\log(1 + KFCDist)$	-2.2674
	(1.7501)
$KFCStores$	0.8460
	(0.7563)
$\log(1 + s)$	0.5764
	(1.1502)
$Center$	5.2309^{***}
	(1.8663)
$Light$	0.2656^{**}
	(0.1134)
$Cinema$	0.2962
	(0.2797)
Regional Dummies (3 Regions)	Yes
# of Unobserved Types	2
Observations	1207

Table 9: Reduced Form Logit Regression: Entry Decision

	KFC's Existence	McD's Existence
Income	0.4861***	1.0444***
	(0.1056)	(0.3331)
$Income^2$	−0.0124***	−0.0460***
	(0.0030)	(0.0152)
Pop	0.2279**	0.1699
	(0.0974)	(0.2097)
Pop^2	-0.0053	0.0053
	0.0043	0.0091
z_{KFC}	0.6705***	0.8150*
	(0.2098)	(0.4686)
z_{KFC}^2	0.0214	0.1322
	(0.0402)	(0.0948)
z_{McD}	−0.6048***	−1.3479***
	(0.1580)	(0.4310)
z_{McD}^2	0.0869***	0.1618**
	(0.0270)	(0.0694)
Center	5.5938***	2.8541***
	(1.2496)	(0.9293)
$Income \times Pop$	0.0317**	0.0063
	(0.0139)	(0.0223)
$Pop \times Center$	−0.2945**	-0.0009
	(0.1011)	(0.0874)
$z_{KFC} \times z_{McD}$	−0.1283**	−0.2530**
	(0.0574)	(0.1268)
Light	0.0241**	0.0121
	(0.0102)	(0.0260)
GDPGrowth	−0.0932***	−0.0983**
	(0.0225)	(0.0475)
Regional Dummies (3 Regions)	Yes	
# of Unobserved Types	2	
Observations	1207	