An Empirical Framework for Discrete Games with Costly Information Acquisition

Youngjae Jeong*

Draft: October 16, 2025

Abstract

This paper develops a novel econometric framework for static discrete choice games with costly information acquisition. In traditional discrete games, players are assumed to perfectly know their own payoffs when making decisions, ignoring that information acquisition can be a strategic choice. In the proposed framework, I relax this assumption by allowing players to face uncertainty about their own payoffs and to optimally choose both the precision of information and their actions, balancing the expected payoffs from precise information against the information cost. The model provides a unified structure to analyze how information and strategic interactions jointly determine equilibrium outcomes. The model primitives are point identified, and the identification results are illustrated through Monte Carlo experiments. The empirical application of the U.S airline entry game shows that the low-cost carriers acquire less precise information about profits and incur lower information costs than other airlines, which is consistent with their business model that focuses on cost efficiency. The analysis highlights how differences in firms' information strategies can explain observed heterogeneity in market entry behavior and competition.

Keywords: Static discrete games, costly information acquisition, semiparametric identification.

JEL Classification: C57, D83, L10.

---

*Department of Economics, The Ohio State University, Email: `jeong.376@osu.edu`

1 Introduction

In many economic environments, decision makers may not perfectly know their own payoffs and acquire information about the payoffs when making choices. Such information acquisition is often costly and affects players' behavior. Recent empirical works have documented this phenomenon in various contexts such as insurance choice (Brown and Jeon, 2024), laundry detergent choice (Joo, 2023), and presidential election (Liao, 2024). Furthermore, experimental evidence shows that individuals tend to acquire costly information when facing payoff uncertainty (Dewan and Neligh, 2020; Dean and Neligh, 2023; Almog and Martin, 2024). These findings highlight the importance of incorporating costly information acquisition into economic models of decision making under payoff uncertainty.

This paper develops a novel econometric framework for static discrete choice games with costly information acquisition. Previous discrete choice games with incomplete information (e.g., Bajari et al. 2010; Aradillas-Lopez 2010) typically assume that each player fully observes their own payoff but not their rivals' payoffs. This paper relaxes this assumption and allows players to face uncertainty about their own payoffs and to acquire information about them. The players optimally choose how much information to acquire about the payoffs, taking into account the costs, and thus determine the precision of their information. This generalization provides a more realistic and flexible framework for analyzing strategic interactions under incomplete information.

Existing literature in industrial organization utilizes static discrete game models to analyze the strategic interactions between firms. These models typically impose strong assumptions about players' information, ignoring the possibility that information acquisition itself can be a strategic choice. For example, the models used to study firms' entry decisions, location decisions, and coordination (Bresnahan and Reiss, 1991; Magnolfi and Roncoroni, 2022; Berry, 1992; Ciliberto and Tamer, 2009; Koh, 2023; Seim, 2006; Sweeting, 2009) assume that players have either complete information or incomplete information about their own payoffs. These assumptions on information overlook the fact that players may not perfectly know their own payoffs and may choose to acquire information strategically.

Unlike the previous discrete games with incomplete information, this paper models the information structure as the outcome of a strategic choice involving costly information acquisition. Each player's information structure consists of signals about their own

payoffs and distributions of these signals. The precision of information is determined by the signal distributions. Acquiring more precise signals is more costly, and each player considers the trade-off between the benefits of acquiring more precise information and the associated costs. Thus, players strategically choose which signals to acquire and the distributions to maximize their expected payoffs net of the information costs.

Formally modeling costly information acquisition poses several challenges. The domain of information structures is infinite-dimensional, which makes it difficult to characterize players' optimal choices and to derive equilibrium conditions. To address these issues, I adopt the framework of rational inattention (Matějka and McKay, 2015). The theory of rational inattention provides a well-suited framework for analyzing the players' optimal strategies in games with costly information acquisition. In particular, within this framework, players' optimal strategies yield a tractable characterization of Nash equilibrium (Denti, 2023; Yang, 2015, 2020; Montes, 2022). This tractability simplifies equilibrium analysis and leads to a feasible model estimation.

In equilibrium, the model captures heterogeneity in players' ability to acquire information through variation in the information strategies across players. This contrasts with previous models that impose an exogenous information structure, where all players are assumed to have the same or arbitrarily specified information. By allowing endogenous information acquisition, the model provides a unified framework to analyze how differences in information strategies, information costs, and player characteristics jointly determine equilibrium outcomes.

Next, I show that the model primitives are identified in both parametric and semiparametric settings. The key assumption for the identification is the exclusion restriction. That is, a player-specific variable affects only one player's payoff with no effect on the rival players' payoffs. The exclusion restriction is not a strong assumption and is commonly used in the literature on discrete games (Tamer, 2003; Ciliberto and Tamer, 2009; Aradillas-Lopez, 2010; Grieco, 2014; Magnolfi and Roncoroni, 2022; Koh, 2023). I illustrate both parametric and semiparametric identification results via Monte Carlo experiments.

Finally, as an empirical application, I use my framework to analyze the entry decisions of airlines in the U.S. I specify a standard parametric linear payoff model and examine both the airlines' strategic behaviors and their costly information acquisition. The results indicate that the airlines' baseline profits are negative, which creates an incentive for them to acquire information about their profits. Furthermore, the results suggest that low-cost

carriers tend to acquire less information than other airlines, which is consistent with their business models that focus on cost efficiency and simplified operations. The analysis highlights how differences in firms' information strategies can explain observed heterogeneity in market entry behavior and competition.

This paper contributes to the literature on econometric discrete games (Bresnahan and Reiss, 1991; Berry, 1992; Tamer, 2003; Seim, 2006; Sweeting, 2009; Ciliberto and Tamer, 2009; Aradillas-Lopez, 2010; Grieco, 2014; Magnolfi and Roncoroni, 2022; Koh, 2023; Xie, 2024). Many of these studies assume that players have complete or incomplete information (or perfectly private information) among the players. A common feature across this literature is that players' information is exogenously specified by the econometrician. In contrast, this paper assumes that players have incomplete information about their own payoffs and does not impose any fixed information structure, instead allowing it to be optimally chosen by players.

The discrete game model in this paper closely resembles recent work by Grieco (2014); Magnolfi and Roncoroni (2022); Koh (2023), which relaxes standard information assumptions and leverages the information structure to model econometric discrete games under flexible and weak information assumptions. These studies adopt a novel approach to model player's information with the signals and the distribution of signals, allowing for a more flexible representation of players' information. However, in these papers, the information structure is still given exogenously by the econometrician and remains fixed. In contrast, this paper endogenizes the information structure by allowing players to optimally choose the information structure through the information strategy. By incorporating the cost of information acquisition, this paper fills a gap in the recent literature and contributes to a more realistic framework for modeling behavior and strategies in discrete games.

This paper also builds on a literature on costly information acquisition and rational inattention. Most of the recent papers theoretically analyze decision problems of a single agent and determine the agent's optimal behavior or decision (Matějka and McKay, 2015; Steiner et al., 2017; Fosgerau et al., 2020). Matějka and McKay (2015) show that a rationally inattentive decision maker's optimal decision rule follows a logit-like formula, linking the rational inattention framework to discrete choice models. Fosgerau et al. (2020) generalizes a decision problem in Matějka and McKay (2015), suggesting a generalized information cost function and proving that the rational inattention discrete choice problem

4

and the additive random utility discrete choice model are equivalent under this information cost function. Steiner et al. (2017) extends Matějka and McKay (2015)'s approach and constructs a rationally inattentive decision maker's problem in a dynamic environment. While I share major assumptions on the setup of the decision problem with these papers, my work extends the analysis to a multi-agent decision problem – discrete games – where strategic interaction plays a key role.

The related literature on game theory with costly information acquisition or rationally inattentive players includes Yang (2015, 2020); Denti (2023); Montes (2022). Yang (2015) investigates a two-player, two-action coordination investment game, and Montes (2022) focuses on attention-move games. In these papers, players have imperfect information about the payoff-relevant state and optimally acquire information. Denti (2023) extends this framework to a game where players acquire not only about the state but also each other's information. I adopt the structure of these theoretical models and expand them into an empirical model. Thus, my model enables the quantitative analysis and estimation of players' information acquisition strategies using data. By bridging the gap between theory and empirical application, the model enhances the ability to validate theoretical predictions and better understand strategic behavior in environments characterized by costly information acquisition.

The remainder of this paper is organized as follows. Section 2 sets up an econometric model for discrete games with costly information acquisition. Section 3 provides the parametric and semiparametric identification. In Section 4, I present identification via Monte Carlo experiments and discuss the estimation method. Section 5 applies my framework to the U.S. airline industry to analyze the entry decisions. Finally, Section 6 concludes. All proofs are in Appendix.

Notation. The boldface letters, e.g., **X** and **x**, represent vectors. The capital letters, e.g., Y_i, denote random variables, while the lowercase letters, e.g., y_i, represent their realized values.

2 Model

In this section, I describe a general model of 2×2 game with costly information acquisition and provide a two-player entry game as a running example.

Let $\mathcal{I} = \{1, 2\}$ be the set of players, and let the players be indexed by $i = 1, 2$. By

convention, I denote the rival of player i as $-i$ or j. [1] Each player i simultaneously chooses an action Y_i from their action set, denoted by $\mathcal{Y}_i$. The action set contains two elements, namely $\mathcal{Y}_i = \{0,1\}$. Let $\mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2$ represent the set of action profiles, and let $\mathbf{Y} = (Y_1, Y_2)^\top \in \mathcal{Y}$ denote the action profile.

2.1 Payoff Structure

The payoff structure of player i is additively separable and given by the following equation:

$$
U_i(Y_i, Y_j, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) = \begin{cases} u_i(Y_j; \mathbf{X}, \mathbf{Z}_i) + \varepsilon_i & \text{if } Y_i = 1, \\ 0, & \text{if } Y_i = 0, \end{cases} \tag{1}
$$

where $\mathbf{X} \in \mathcal{X} \subseteq \mathbb{R}^{d_x}$ is a vector of market-specific variables and $\mathbf{Z}_i \in \mathcal{Z}_i \subseteq \mathbb{R}^{d_z}$ is a vector of player-specific variables. The payoff function for action $Y_i = 1$, $u_i : \mathcal{Y}_j \times \mathcal{X} \times \mathcal{Z}_i \to \mathbb{R}$, can be specified parametrically or nonparametrically. The payoff shock $\varepsilon_i \in \mathcal{E}_i$ is unobservable to player i (and to the researcher) when making decisions, and I interpret it as the uncertainty in payoffs. The payoff for action $Y_i = 0$ is normalized to zero. In the next example below, I demonstrate parametric and nonparametric payoff functions.

The vector of market-specific variables $\mathbf{X}$ is common to all players and has influence on both players' payoffs. On the contrary, the vector of player-specific variables $\mathbf{Z}_i$ only affects player i's payoff, but not on rival's payoff. As shown in the payoff function (1), player i's payoff depends on $\mathbf{X}$ and $\mathbf{Z}_i$, but not on $\mathbf{Z}_j$. The variable $\mathbf{Z}_i$ is also referred to as the player-specific payoff shifter in the literature and satisfies the exclusion restriction assumption (Aradillas-Lopez, 2010; Bajari et al., 2010; Xie, 2024). The exclusion restriction assumption plays a key role in the identification, which will be shown in the next section. Examples of the market-specific variables $\mathbf{X}$ include market size, population, and consumer preference, whereas examples of player-specific variables $\mathbf{Z}_i$ include costs, technology, and product quality. I assume that the variables $(\mathbf{X}', \mathbf{Z}_i', \mathbf{Z}_j')'$ are observable to all players and the researcher. Furthermore, I do not restrict the support of each variables. That is, the support of each variables of $\mathbf{X}$ and $\mathbf{Z}_i$ can be either discrete or continuous.

Given the payoff structure (1), the payoff uncertainty arises from two sources; the rival's action Y_j and the unobservable payoff shock ε_i. When making a decision, each player

---

[1]I consider a two-player game, and denoting the rival by j does not violate the notation, as I can set $j = 3 - i$.

does not observe either the realized rival's action or the payoff shocks. In particular, the unobservable payoff shock ε_i directly affects the payoff, which creates an incentive for player i to acquire information about it. At the same time, because rival's decision is determined strategically, player i needs to form a belief about rival's behavior. Thus, the uncertainty in payoffs makes player's decision problem strategic since players weigh the value of acquiring information about their own payoff shocks while anticipating how rival will behave in equilibrium.

Since the main focus of this paper is on 2×2 games, the payoff function (1) can be expressed in a standard form commonly used in the literature. Define two functions $\pi_i(\cdot, \cdot)$ and $\delta_i(\cdot, \cdot)$ as follows:

$$\pi_i(\mathbf{X}, \mathbf{Z}_i) = u_i(Y_j = 0; \mathbf{X}, \mathbf{Z}_i),$$

$$\delta_i(\mathbf{X}, \mathbf{Z}_i) = u_i(Y_j = 1; \mathbf{X}, \mathbf{Z}_i) - u_i(Y_j = 0; \mathbf{X}, \mathbf{Z}_i).$$

I refer to the function $\pi_i(\cdot, \cdot)$ as the base payoff and the function $\delta_i(\cdot, \cdot)$ as the strategic effect. Models without a strategic effect can be considered as a single-agent model where a player chooses action in isolation without strategic interaction. Thus, I focus on models with non-zero strategic effect $\delta_i(\cdot, \cdot) \neq 0$ for all $\mathbf{X}$ and $\mathbf{Z}_i$ for each player i.

The sign of the strategic effect $\delta_i(\cdot, \cdot)$ depends on the strategic incentives of the game. The strategic effect is positive if the players want to coordinate, and the example is timing of radio commercials (Sweeting, 2009). On the other hand, the strategic effect is negative if the players want to differentiate, and the example is entry games. In a two-player entry game, the base payoff $\pi_i(\cdot, \cdot)$ corresponds to the monopoly profit, and the sum $\pi_i(\cdot, \cdot) + \delta_i(\cdot, \cdot)$ corresponds to the duopoly profit.

Using the above functions of the base payoff and the strategic effect, I can rewrite the payoff function as follows:

$$U_i(Y_i, Y_j, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) = \begin{cases} \pi_i(\mathbf{X}, \mathbf{Z}_i) + \delta_i(\mathbf{X}, \mathbf{Z}_i) \cdot \mathbb{1}(Y_j = 1) + \varepsilon_i, & \text{if } Y_i = 1, \\ 0, & \text{if } Y_i = 0, \end{cases} \tag{2}$$

where $\mathbb{1}(\cdot)$ is the indicator function. Given the additive separable payoff shock ε_i, the original payoff function (1) can be transformed into a sum of the base payoff and the competitive effect without imposing any restrictions. The analysis of the rest of this paper is mainly based on the payoff function (2).

Example 1 (A Two-player Entry Game). Consider a two-player entry game as in Bresna-han and Reiss (1991) and Tamer (2003). There are two players, $\mathcal{I} = \{1,2\}$. The players are firms, and each firm simultaneously chooses either to enter the market $Y_i = 1$ or not to enter the market $Y_i = 0$. The profit function for firm i is defined by

$$\Pi_i(Y_i, Y_j, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) = \begin{cases} \pi_i(\mathbf{X}, \mathbf{Z}_i) + \delta_i(\mathbf{X}, \mathbf{Z}_i) \cdot \mathbb{1}(Y_j = 1) + \varepsilon_i, & \text{if } Y_i = 1, \\ 0, & \text{if } Y_i = 0. \end{cases}$$

Firm i earns zero profit if it chooses not to enter $Y_i = 0$. If firm i chooses to enter, it gains monopoly profit if the rival stays out $Y_j = 0$ or duopoly profit if the rival enters $Y_j = 1$. In this model, $\delta_i(\cdot, \cdot)$ represents the competitive effect and is assumed to be negative. The matrix of the payoff structure is summarized as the table 1.

Payoffs $\pi_i(\cdot, \cdot)$ and $\delta_i(\cdot, \cdot)$ may be specified either as nonparametric functions or as parametric functions of covariates. The existing literature on entry games (Bresnahan and Reiss, 1991; Tamer, 2003; Ciliberto and Tamer, 2009) assumes linear parametric payoff functions, e.g.,

$$\pi_i(\mathbf{X}, \mathbf{Z}_i) = \mathbf{X}\alpha_i + \mathbf{Z}_i\beta_i,$$
$$\delta_i(\mathbf{X}, \mathbf{Z}_i) = \delta_i.$$

$\square$

	$Y_2 = 0$	$Y_2 = 1$
$Y_1 = 0$	0, 0	0, $\pi_2(\mathbf{X}, \mathbf{Z}_2) + \varepsilon_2$
$Y_1 = 1$	$\pi_1(\mathbf{X}, \mathbf{Z}_1) + \varepsilon_1$, 0	$\pi_1(\mathbf{X}, \mathbf{Z}_1) + \delta_1(\mathbf{X}, \mathbf{Z}_1) + \varepsilon_1$, $\pi_2(\mathbf{X}, \mathbf{Z}_2) + \delta_2(\mathbf{X}, \mathbf{Z}_2) + \varepsilon_2$

Table 1: The payoff matrix of 2×2 entry game in Example 1.

2.2 Information Structure

Every player i does not directly observe their own payoff shocks ε_i and rival's payoff shock ε_j when they make a decision. The payoff shock ε_i is realized and observable to each player i after the realization of action profile $\mathbf{Y}$. While the players are uncertain about the exact value of ε_i, they have a prior belief on the payoff shocks $\varepsilon = (\varepsilon_i, \varepsilon_j)$. The payoff shock ε is distributed according to the distribution F_i. I refer to the distribution F_i as player i's prior belief. The following assumption summarizes the prior beliefs.

Assumption 1.   1. The payoff shocks ε_i and ε_j are independent conditional on $(\mathbf{X}', \mathbf{Z}'_i, \mathbf{Z}'_j)'$. Moreover, for each player i, the payoff shock ε_i is independent of $(\mathbf{X}', \mathbf{Z}'_i, \mathbf{Z}'_j)'$.

2. The prior belief $F_i(\varepsilon)$ is absolutely continuous with respect to the Lebesgue measure. That is, the prior $F_i(\varepsilon)$ is continuously differentiable and strictly increasing almost everywhere over the support of ε.

I model players' information through the signals they receive and the corresponding distributions, following Magnolfi and Roncoroni (2022) and Koh (2023). Each player i receives a private signal $\tau_i^{\mathbf{X},\mathbf{Z}} \in \mathcal{T}_i^{\mathbf{X},\mathbf{Z}}$, where $\mathcal{T}_i^{\mathbf{X},\mathbf{Z}}$ denotes the set of all possible signals which player i may receive.[2] The private signal $\tau_i^{\mathbf{X},\mathbf{Z}}$ may contain information about the payoff shock ε, and its precision depends on the signal distribution.

Formally, the information structure $S_i^{\mathbf{X},\mathbf{Z}}$ of player i is defined as a pair consisting of the set of signals and the probability distributions over the signals:

$$S_i^{\mathbf{X},\mathbf{Z}} = \left(\mathcal{T}_i^{\mathbf{X},\mathbf{Z}}, \{P_i^{\mathbf{X},\mathbf{Z}}(\cdot|\varepsilon) : \varepsilon \in \mathcal{E}\} \right), \tag{3}$$

where $\{P_i^{\mathbf{X},\mathbf{Z}}(\cdot|\varepsilon)\}$ represents the family of conditional distributions of signals given the payoff shock ε. Each signal $\tau_i^{\mathbf{X},\mathbf{Z}} \in \mathcal{T}_i^{\mathbf{X},\mathbf{Z}}$ is realized according to its corresponding probability $P_i^{\mathbf{X},\mathbf{Z}}(\tau_i^{\mathbf{X},\mathbf{Z}}|\varepsilon)$.

One advantage of utilizing the information structure is that it allows us to express any type of information assumption from the existing literature. To illustrate the concept of the information structure, consider the following two-player entry game example.

Example 2 (Example 1 continued)**.** In Bresnahan and Reiss (1991), Tamer (2003), and Ciliberto and Tamer (2009), firm i not only observes his own payoff shock ε_i, but also rival's payoff shock ε_j. Thus, the firms have complete information.

The complete information can be interpreted as follows: each player i receives the exact value of ε_i through the signal τ_i. Formally, I can express the complete information structure as follows. The signal space for each firm is defined by $\mathcal{T}_1 = \mathcal{E}_1 \times \mathcal{E}_2$ and $\mathcal{T}_2 = \mathcal{E}_1 \times \mathcal{E}_2$. The distribution over signals for each firm is defined by

$$P_1 (\tau_1 = (\varepsilon_1, \varepsilon_2) | \varepsilon_1, \varepsilon_2) = 1,$$
$$P_2 (\tau_2 = (\varepsilon_1, \varepsilon_2) | \varepsilon_2, \varepsilon_1) = 1.$$

$\square$

---

[2]Here, $(\mathbf{X}, \mathbf{Z})$ is a shorthand for $(\mathbf{X}, \mathbf{Z}_i, \mathbf{Z}_j)$.

Example 3. I describe another information structure from Seim (2006), Aradillas-Lopez (2010), and Bajari et al. (2010). Firm i only observes his own payoff shock ε_i, but does not receive any signals about their rival's payoff shock ε_j. In this situation, the firms have perfectly private information or incomplete information. The signal space for each firm is expressed as, for example, $\mathcal{T}_1 = \mathcal{E}_1 \times \{-\infty\}$ and $\mathcal{T}_2 = \mathcal{E}_2 \times \{-\infty\}$. The distribution over signals for each firm is

$$P_1\left(\tau_1 = (\varepsilon_1, -\infty) \mid \varepsilon_1, \varepsilon_2\right) = 1,$$
$$P_2\left(\tau_2 = (\varepsilon_2, -\infty) \mid \varepsilon_2, \varepsilon_1\right) = 1.$$

Alternatively, the signal spaces can be defined as $\mathcal{T}_1 = \mathcal{E}_1$ and $\mathcal{T}_2 = \mathcal{E}_2$. If the signal τ_i and rival's payoff shock ε_j are conditionally independent given own payoff shock ε_i, [3] the distribution of the signals can be expressed as

$$P_1\left(\tau_1 = \varepsilon_1 \mid \varepsilon_1, \varepsilon_2\right) = P_1\left(\tau_1 = \varepsilon_1 \mid \varepsilon_1\right) = 1,$$
$$P_2\left(\tau_2 = \varepsilon_2 \mid \varepsilon_2, \varepsilon_1\right) = P_2\left(\tau_2 = \varepsilon_2 \mid \varepsilon_2\right) = 1.$$

This information structure is identical to the one in Seim (2006), Aradillas-Lopez (2010), and Bajari et al. (2010). □

Example 4. I can also consider a situation where each firm i has no information about his own payoff shock ε_i. In this case, each firm's signal space is singleton, e.g., $\mathcal{T}_i = \{-\infty\}$ for all realization of ε, and the distribution over the signal is, for $i = 1, 2$,

$$P_i\left(\tau_i = -\infty \mid \varepsilon_i, \varepsilon_j\right) = 1.$$

□

2.3 Timeline of the Game

The timing of the Bayesian game is as follows. In period 1, two players enter the game, and a vector of the covariates $(\mathbf{x}', \mathbf{z}_1', \mathbf{z}_2')'$ is realized. In addition, the payoff shock ε is drawn by nature. Each player has limited information about their payoffs due to the unobserved payoff shock. In period 2, each player determines how much information to acquire and pays information costs. I refer to this strategy as information strategy. In period 3, the

---

[3]Once player i's own payoff shock ε_i is given, the signal τ_i does not provide any information about rival's payoff shock ε_j.
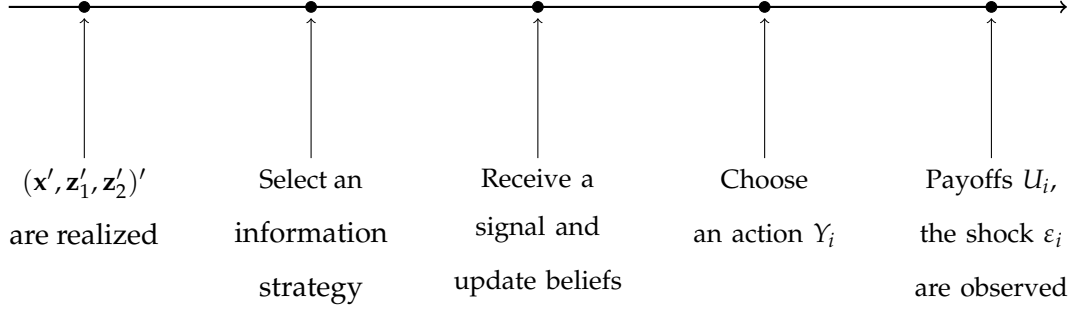
Figure 1: Timing of the Bayesian game

players receive private signals and update beliefs about the payoff shock. In period 4, the players simultaneously choose actions Y_i as a Bayesian expected utility maximizer. I refer to this strategy as action strategy. In period 5, the players' payoffs are realized. This is summarized in the Figure 1.

2.4 Costly Information Acquisition

Players have prior beliefs about the payoff shock ε and choose how much information to acquire to form posterior beliefs about the payoffs. However, acquiring information is costly. According to the rational inattention theory (Matějka and McKay, 2015; Caplin et al., 2019), players cannot gather and process all available information, as attention is scarce and limited. Moreover, acquiring information requires both time and effort, further adding to the cost of the information acquisition process.

Player i cannot influence the realization of ε, but can freely select which signals to observe and their distributions. The precision of these signals depends on the chosen distribution, and selecting more informative signals corresponds to acquiring more precise information. However, there is a cost of acquiring information with more informative signals being more costly. Given the prior belief F_i, observed covariates $(\mathbf{X}', \mathbf{Z}')'$, and the information cost function, player i optimally chooses the information structure

$$S_i^{\mathbf{X},\mathbf{Z}} = \left(\mathcal{T}_i^{\mathbf{X},\mathbf{Z}}, \{ P_i^{\mathbf{X},\mathbf{Z}}(\cdot|\varepsilon) \} \right),$$

which reflects a trade-off between the benefits of acquiring precise information and the costs of doing so. As a result, players obtain partially informative signals, which lie between fully informative and completely uninformative signals.

11

In specifying information cost function, I adopt the entropy-based mutual information function widely used in the rational inattention literature (Matějka and McKay, 2015). First, I introduce the entropy function, which measures uncertainty or unpredictability of an event. Next, I define the mutual information function using the entropy function. The mutual information function measures the expected uncertainty reduction. Finally, I define the information cost function, which is proportional to the mutual information.

Let X be a continuous random variable with the density function $f(x)$. Following (Cover and Thomas, 2006), the entropy of X is defined as follows [4]:

$$H(f) = H(X) = -\int_{x \in \mathcal{X}} f(x) \ln f(x) \, dx.$$

To measure the acquired information, I use the mutual information. Let X and Y be two random variables with the joint density function $f(x, y)$ and marginal density functions $f(x)$ and $f(y)$. The mutual information $I(X, Y)$ between two random variables X and Y is defined as follows:

$$I(X, Y) = \int_{\mathcal{Y}} \int_{\mathcal{X}} f(x, y) \ln \left(\frac{f(x, y)}{f(x) f(y)} \right) \, dx \, dy$$
$$= H(X) - \mathrm{E}_Y \left[H(X \mid Y) \right].$$

The first term $H(X)$ denotes the entropy of X, which measures uncertainty of X without any information about Y. The second term $H(X \mid Y)$ denotes the entropy of X given Y, which represents the uncertainty of X once Y is known. Thus, the mutual information $I(X, Y)$ quantifies the expected uncertainty reduction of X due to the knowledge of Y, which corresponds to the amount of information acquired. In other words, the mutual information measures the decrease in prior uncertainty after information acquisition.

Finally, the information cost function $C(X, Y)$ is defined to be proportional to the mutual information, which measures the amount of information acquired. Formally, the information cost function is expressed as:

$$C(X, Y) = \lambda \cdot I(X, Y),$$

where $\lambda > 0$ is unit cost of information and $I(X, Y)$ is the mutual information between X and Y.

---

[4]The random variable X corresponds to the payoff shock ε in the model. In the econometrics literature, the payoff shock follows the normal distribution or the type 1 extreme value distribution. On the other hand, theory literature assumes a discrete payoff shock in many cases. The entropy function with a discrete random variable X is defined by $H(X) = H(p) = \sum_{x \in \mathcal{X}} p(x) \ln p(x)$ where $p(x)$ is a probability mass function.

Based on the above definitions, the information cost function in my setting is defined as follows. First, I define the mutual information between the prior belief about the payoff shock and the posterior belief after observing the signal

$$I(F_i, P_i^{\mathbf{X,Z}}) = H(F_i(\cdot)) - \mathrm{E}_{\tau_i}\left[H(P_i^{\mathbf{X,Z}}(\cdot|\tau_i^{\mathbf{X,Z}}))\right],\tag{4}$$

where $H(F_i(\cdot))$ is the entropy of the prior belief $F_i(\varepsilon)$, and $H(P_i^{\mathbf{X,Z}}(\cdot|\tau_i^{\mathbf{X,Z}}))$ is the entropy of the posterior belief $P_i^{\mathbf{X,Z}}(\varepsilon|\tau_i^{\mathbf{X,Z}})$ after observing the signal $\tau_i^{\mathbf{X,Z}}$. This measures the information acquired about the payoff shock ε by observing the signal $\tau_i^{\mathbf{X,Z}}$.

Next, the information cost function is defined as:

$$C_i(F_i, P_i^{\mathbf{X,Z}}) = \lambda_i \cdot I(F_i, P_i^{\mathbf{X,Z}}),\tag{5}$$

where the unit cost of information $\lambda_i > 0$. The information cost function is proportional to the mutual information $I(F_i, P_i^{\mathbf{X,Z}})$. If a signal $\tau_i^{\mathbf{X,Z}}$ does not contain any information about the payoff shock ε or is independent of the payoff shock, the mutual information is zero, $I(F_i, P_i^{\mathbf{X,Z}}) = 0$, and so is the information cost $C_i(F_i, P_i^{\mathbf{X,Z}}) = 0$. The following example illustrates the mutual information and the information cost function.

Example 5. Consider a simple entry decision with unobservable payoff shocks given as $\varepsilon \in \{G, B\}$. Good market conditions are represented by $\varepsilon = G$, while bad market conditions are represented by $\varepsilon = B$. Suppose that the firm's prior belief about the payoff shock is given as $P(G) = 0.5$ and $P(B) = 0.5$. The entropy of the prior belief is

$$H(P) = -\sum_{x \in \{G,B\}} P(x) \ln P(x) \approx 0.693.$$

Now, consider two specific information structures $S_1 = (\{0,1\}, P_1)$ and $S_2 = (\{0,1\}, P_2)$. Each information structure generates signals $\tau \in \{0,1\}$, where signal $\tau = 1$ can be interpreted as a recommendation to enter the market, and $\tau = 0$ as a recommendation not to enter. Let the signal distributions of the S_1 be

$$P_1(\tau = 1|G) = P_1(\tau = 0|B) = 0.8, \quad P_1(\tau = 1|B) = P_1(\tau = 0|G) = 0.2,$$

and the signal distributions of the S_2 be

$$P_2(\tau = 1|G) = P_2(\tau = 0|B) = 0.6, \quad P_2(\tau = 1|B) = P_2(\tau = 0|G) = 0.4.$$

Applying Bayes' rule, the posterior beliefs under S_1 are

$$P_1(G|1) = 0.8, \quad P_1(B|1) = 0.2,$$
$$P_1(G|0) = 0.2, \quad P_1(B|0) = 0.8,$$

and the mutual information is:

$$I(P, P_1(\cdot|\tau)) = H(P) - \mathrm{E}_\tau[H(P_1(\cdot|\tau))] \approx 0.193.$$

Assuming the unit cost of information to be $\lambda = 1$, the information cost for S_1 is:

$$C(P, P_1) \approx 0.193.$$

Similarly, the posterior beliefs under S_2 are:

$$P_2(G|1) = 0.6, \quad P_2(B|1) = 0.4,$$
$$P_2(G|0) = 0.4, \quad P_2(B|0) = 0.6,$$

with the mutual information:

$$I(P, P_2(\cdot|\tau)) = H(P) - \mathrm{E}_\tau[H(P_2(\cdot|\tau))] \approx 0.020,$$

and the information cost:

$$C(P, P_2) \approx 0.020.$$

Therefore, we conclude that the information structure S_1 is more informative than S_2. This result arises because the signals in S_1 are more correlated with the true market conditions. Under S_1, the signal correctly matches the underlying market condition with higher probability than under S_2. As a result, the posterior beliefs under S_1 are more informative, leading to a larger expected reduction in uncertainty. □

The intuition of the entropy-based information cost function can be explained as follows. Consider a situation where a player asks a series of binary questions (yes or no questions) at a cost per question to determine the outcome of the payoff shocks. The more questions he asks, the more information he can acquire, but also the higher the costs he pays. This cost is proportional to the number of questions he asked and reduction in uncertainty about the payoff shocks. Initially, the player faces a certain level of ex-ante uncertainty, measured by the entropy of the prior belief. As he gathers information, this uncertainty is reduced. The mutual information function precisely quantifies this average reduction in uncertainty, reflecting how much the player learns about the true payoff shocks by observing a signal. Therefore, a higher mutual information value implies a greater reduction in uncertainty, which, in this model, corresponds to a higher information cost.

2.5 Strategy and Equilibrium

Each player's decision problem involves two strategies. First, in period 2 of Figure 1, the player chooses the information strategy. The information strategy is the selection of the optimal information structure $S_i^{\mathbf{X,Z}} = (\mathcal{T}_i^{\mathbf{X,Z}}, P_i^{\mathbf{X,Z}}(\cdot \,|\, \varepsilon_i))$ given the information cost function (5). [5] In other words, it is the choice of how precise the signals should be, balancing the benefits of better information against the costs of acquiring it.

Second, once the player has chosen the optimal information structure, received the signals $\tau_i^{\mathbf{X,Z}}$, the player updates his prior beliefs to form the posterior beliefs using Bayes' rule:

$$P_i^{\mathbf{X,Z}}(\varepsilon_i \,|\, \tau_i^{\mathbf{X,Z}}) = \frac{P_i^{\mathbf{X,Z}}(\tau_i^{\mathbf{X,Z}} \,|\, \varepsilon_i) f_i(\varepsilon_i)}{P_i^{\mathbf{X,Z}}(\tau_i^{\mathbf{X,Z}})},$$

where $f_i(\cdot)$ is the density of the prior belief $F_i(\cdot)$. After forming the posterior beliefs $P_i^{\mathbf{X,Z}}(\cdot \,|\, \tau_i^{\mathbf{X,Z}})$, the player chooses an action Y_i. This stage corresponds to periods 3 and 4 in Figure 1. The action strategy is a mapping from the signal to the action space $\mathcal{Y}_i$, denoted by $\sigma_i(\tau_i^{\mathbf{X,Z}})$.

Formally, the action strategy is defined as the best response to the rival's action strategy given the optimal information structure and the beliefs about the rival's action $b_j(Y_j|\mathbf{X,Z})$:

$$\sigma_i^{\mathbf{X,Z}}(\tau_i^{\mathbf{X,Z}}) = Y_i \iff \arg\max_{Y_i \in \mathcal{Y}_i} \mathrm{E}_{\varepsilon_i \sim P_i^{\mathbf{X,Z}}(\cdot|\tau_i^{\mathbf{X,Z}})} \left[\sum_{Y_j} U_i(Y_i, Y_j, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \cdot b_j(Y_j|\mathbf{X}, \mathbf{Z}_j) \right],$$

where the expectation is taken with respect to the posterior belief $P_i^{\mathbf{X,Z}}(\varepsilon \,|\, \tau_i^{\mathbf{X,Z}})$. [6]

One important result of decision-making with the costly information acquisition is that the information strategy and action strategy can be combined into a single strategy under optimal behavior (Matějka and McKay, 2015; Yang, 2015, 2020). Since information acquisition is costly, players will avoid collecting useless or redundant information, which determines the optimal signal space. In particular, the signal space of each player's optimal

---

[5] Player i's payoff shock ε_i is independent from ε_j conditional on $(\mathbf{X}', \mathbf{Z}')'$ by Assumption 1. I further assume that the signal $\tau_i^{\mathbf{X,Z}}$ is independent from ε_j given ε_i. That is, ε_j is the private information of rival, and the signal $\tau_i^{\mathbf{X,Z}}$ does not provide any information about ε_j once ε_i is given. Thus, the signal distribution $P_i^{\mathbf{X,Z}}(\cdot|\varepsilon)$ can be simplified to $P_i^{\mathbf{X,Z}}(\cdot|\varepsilon_i)$.

[6] The expectation taken with respect to the posterior belief $P_i^{\mathbf{X,Z}}(\varepsilon_i \,|\, \tau_i^{\mathbf{X,Z}})$ is:

$$\int_{\varepsilon_i} \left[\sum_{Y_j} U_i(Y_i, Y_j, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \cdot b_j(Y_j|\mathbf{X}, \mathbf{Z}_j) \right] P_i^{\mathbf{X,Z}}(\varepsilon_i \,|\, \tau_i^{\mathbf{X,Z}}) \mathrm{d}\varepsilon_i.$$

information structure contains at most as many signals as the number of elements of the action set $\mathcal{Y}_i$. Moreover, under optimal behavior, each signal leads to exactly one action. Distinguishing distinct signals that induce the same action is inefficient, as it requires more costs but gives the same payoff. It is more efficient for the players to collapse such distinct signals into one single signal. The following lemma from Matějka and McKay (2015); Yang (2015, 2020) formalizes this property.

Lemma 1. Let $S_i^{\mathbf{X,Z}} = (\mathcal{T}_i^{\mathbf{X,Z}}, P_i^{\mathbf{X,Z}})$ be the optimal information structure for player $i \in \mathcal{I}$. Then, player i always plays pure strategies, and the signal space $\mathcal{T}_i^{\mathbf{X,Z}}$ contains no more elements than those in $\mathcal{Y}_i$.

Proof. In the Appendix $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Lemma 1 implies that the optimal information structure can be represented as

$$S_i^{\mathbf{X,Z}} = \left(\mathcal{Y}_i, \{P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i)\} \right).$$

Here, the signal space contains two signals, i.e., $\mathcal{T}_i^{\mathbf{X,Z}} \equiv \mathcal{Y}_i = \{0,1\}$, with the corresponding signal distribution $P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i)$. Intuitively, receiving the signal $\tau_i^{\mathbf{X,Z}} = 1$ can be interpreted as "accept/yes", whereas $\tau_i^{\mathbf{X,Z}} = 0$ corresponds to "reject/no". The player's optimal action strategy is simply to follow the received signal. Specifically, if player i receives a signal of "accept/yes", $\tau_i^{\mathbf{X,Z}} = 1$, with probability $P_i^{\mathbf{X,Z}}(1|\varepsilon_i)$, then the player chooses action $Y_i = 1$ as recommended. Conversely, if player i receives a signal of "reject/no", $\tau_i^{\mathbf{X,Z}} = 0$, with probability $P_i^{\mathbf{X,Z}}(0|\varepsilon_i)$, then the player chooses action $Y_i = 0$ as recommended. Thus, the action strategy is

$$\sigma_i^{\mathbf{X,Z}}(\tau_i^{\mathbf{X,Z}} = Y_i) = Y_i.$$

Hence, once the optimal information structure is chosen, the action strategy is automatically determined by following the signal. In this sense, it is without loss of generality to regard player i's strategy as focusing solely on the information strategy.

Each player i maximizes the ex-ante expected payoff less the information cost. Let $b_j(Y_j|\mathbf{X}, \mathbf{Z})$ be player i's belief about the rival's action Y_j. Using the result of Lemma 1, player i's ex-ante expected payoff maximization problem can be expressed as:

$$\max_{P_i^{\mathbf{X,Z}}:\mathcal{E}_i \to (0,1)} \int_{\varepsilon_i} \sum_{Y_i} \sum_{Y_j} U_i(Y_i, Y_j, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \cdot b_j(Y_j|\mathbf{X}, \mathbf{Z}) \cdot P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i) \, dF_i(\varepsilon_i) - C_i(F_i, P_i^{\mathbf{X,Z}}) \quad (6)$$

subject to

$$C_i(F_i, P_i^{\mathbf{X},\mathbf{Z}}) = \lambda_i I(F_i, P_i^{\mathbf{X},\mathbf{Z}}), \quad \sum_{Y_i} P_i^{\mathbf{X},\mathbf{Z}} = 1, \quad P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon) \geq 0, \quad \text{for all } i \in \mathcal{I}, \varepsilon_i \in \mathcal{E}_i.$$

The solution to the optimization problem has been well established in Matějka and McKay (2015) and Caplin et al. (2019). [7] Proposition 1 below characterizes the player's optimal strategy.

Proposition 1. Let $P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i)$ be a strategy of player i in the decision problem (6). Then the strategy $P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i)$ is optimal if and only if it satisfies the following:

$$P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) = \frac{P_i^{\mathbf{X},\mathbf{Z}}(Y_i) \exp\left(\sum_{Y_j} U_i(Y_i, Y_j, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) b_j(Y_j|\mathbf{X}, \mathbf{Z})\right)^{1/\lambda_i}}{\sum_{Y_{i'}} P_i^{\mathbf{X},\mathbf{Z}}(Y_i') \exp\left(\sum_{Y_j} U_i(Y_i', Y_j, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) b_j(Y_j|\mathbf{X}, \mathbf{Z})\right)^{1/\lambda_i}}, \quad \text{a.s.} \quad (7)$$

with $P_i^{\mathbf{X},\mathbf{Z}}(Y_i) = \int_{\varepsilon_i} P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) \, dF_i(\varepsilon_i) > 0$.

Proof. In the Appendix $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Next, I formally define the equilibrium of the game in the context of costly information acquisition. The solution concept is the Bayesian-Nash equilibrium (BNE), which is the standard in the literature on incomplete information games.

Definition 1. A Bayesian-Nash equilibrium (BNE) is a profile of strategies $(P_i^{\mathbf{X},\mathbf{Z}}(Y_i))_{i\in\mathcal{I}}$ such that for each player $i \in \mathcal{I}$ with $j \neq i$,

1. Given beliefs about the rival's action $b_j(Y_j|\mathbf{X}, \mathbf{Z})$, the strategy $P_i^{\mathbf{X},\mathbf{Z}}$ solves the maximization problem (6).

2. Beliefs are consistent. That is, $b_j(Y_j|\mathbf{X}, \mathbf{Z}) = P_j^{\mathbf{X},\mathbf{Z}}(Y_j)$.

The equilibrium concept in Definition 1 captures the strategic interactions between players in a game with costly information acquisition. First, each player's strategy in a Bayesian-Nash equilibrium optimally balances the expected payoff from each action with the costs of acquiring information, conditional on the beliefs about rival's actions. Second, consistency of beliefs ensures that each player's expectation about the rival's strategies aligns with the actual equilibrium strategies. Finally, the equilibrium defined here

---

[7]Matějka and McKay (2015) established the necessary conditions for the rational inattention optimization problem, but did not prove sufficiency. Subsequently, Caplin et al. (2019) derived the necessary and sufficient conditions using the concept of consideration sets.

generalizes the standard BNE concept to account for endogenous information acquisition, providing a foundation for empirical estimation of discrete games with costly information.

In the next subsection, I focus on the equilibrium in the choice probability space, which facilitates the characterization of players' strategies and the computation of equilibrium outcomes.

2.6 Equilibrium Existence and Uniqueness

Given the equilibrium strategies defined above, it is often convenient to represent the equilibrium in terms of conditional choice probabilities (CCPs), which aggregate over the unobserved payoff shocks. Associated with each player's equilibrium strategy, the CCP $P_i(Y_i|\mathbf{X}, \mathbf{Z}_i, \mathbf{Z}_j)$ is defined as

$$P_i(Y_i|\mathbf{X}, \mathbf{Z}_i, \mathbf{Z}_j) \equiv P_i^{\mathbf{X}, \mathbf{Z}}(Y_i) = \int_{\varepsilon_i} P_i^{\mathbf{X}, \mathbf{Z}}(Y_i|\varepsilon_i) \, dF_i(\varepsilon_i). \tag{8}$$

Since the action is binary, the CCPs can be simplified to a single probability of choosing action $Y_i = 1$:

$$P_i(\mathbf{X}, \mathbf{Z}_i, \mathbf{Z}_j) \equiv P_i(Y_i = 1|\mathbf{X}, \mathbf{Z}_i, \mathbf{Z}_j).$$

Moreover, the equilibrium condition can be expressed as a fixed-point equation. Let $\mathbf{P}(\mathbf{X}, \mathbf{Z}) = (P_1(\mathbf{X}, \mathbf{Z}), P_2(\mathbf{X}, \mathbf{Z}))'$ denote the vector of conditional choice probabilities. For $i, j \in \mathcal{I}$ with $j \neq i$, define

$$\psi_i(p_i, p_j, \mathbf{X}, \mathbf{Z}) = \int_{\varepsilon_i} \frac{p_i \exp\left(\pi_i(\mathbf{X}, \mathbf{Z}_i) + \delta_i(\mathbf{X}, \mathbf{Z}_i) \cdot p_j + \varepsilon_i\right)^{1/\lambda_i}}{p_i \exp\left(\pi_i(\mathbf{X}, \mathbf{Z}_i) + \delta_i(\mathbf{X}, \mathbf{Z}_i) \cdot p_j + \varepsilon_i\right)^{1/\lambda_i} + 1 - p_i} \, dF_i(\varepsilon_i). \tag{9}$$

Stacking the two functions, define the mapping $\Psi : [0, 1]^2 \to [0, 1]^2$ as:

$$\Psi(p_1, p_2, \mathbf{X}, \mathbf{Z}) = \begin{pmatrix} \psi_1(p_1, p_2, \mathbf{X}, \mathbf{Z}) \\ \psi_2(p_2, p_1, \mathbf{X}, \mathbf{Z}) \end{pmatrix}. \tag{10}$$

From equations (7) and (8), the equilibrium choice probabilities satisfy

$$P_i(\mathbf{X}, \mathbf{Z}) = \psi_i(P_i(\mathbf{X}, \mathbf{Z}), P_j(\mathbf{X}, \mathbf{Z}), \mathbf{X}, \mathbf{Z}).$$

Thus, the equilibrium conditional choice probabilities are characterized as the fixed-point solution of

$$\mathbf{P}(\mathbf{X}, \mathbf{Z}) = \Psi(\mathbf{P}(\mathbf{X}, \mathbf{Z}), \mathbf{X}, \mathbf{Z}). \tag{11}$$

The following proposition establishes existence of a Bayesian Nash equilibrium. The proof follows from Brouwer's fixed-point theorem.

Proposition 2. Under the Assumption 1, there exists a Bayesian Nash equilibrium of the games with costly information acquisition.

Proof. In the Appendix ◻

As shown in Figure 2, the multiple solutions to equation (11) may arise, a common feature in discrete choice games. Aradillas-Lopez (2010) investigates the conditions under which the equilibrium is unique. Building on Gale-Nikaido conditions (Gale and Nikaido, 1965), Aradillas-Lopez (2010) derives the sufficient conditions for equilibrium uniqueness. The following proposition formalizes uniqueness conditions in the context of the game with costly information acquisition.

Proposition 3. Let $\nabla_{\mathbf{P}}\Psi(\mathbf{P}(\mathbf{X},\mathbf{Z}),\mathbf{X},\mathbf{Z})$ denote the Jacobian of (9) evaluated at the equilibrium choice probabilities $\mathbf{P}(\mathbf{X},\mathbf{Z})$. The solution to the equation

$$\mathbf{P}(\mathbf{X},\mathbf{Z}) - \Psi(\mathbf{P}(\mathbf{X},\mathbf{Z}),\mathbf{X},\mathbf{Z}) = 0$$

is unique if no principal minors of the matrix

$$I_{2\times2} - \nabla_{\mathbf{P}}\Psi(\mathbf{P}(\mathbf{X},\mathbf{Z}),\mathbf{X},\mathbf{Z}) \tag{12}$$

are zero, where $I_{2\times2}$ is the 2-dimensional identity matrix. That is,

1. the $(1,1)$ element of the matrix (12) is positive, i.e., $1 - \left.\dfrac{\partial\psi_1}{\partial P_1}\right|_{\mathbf{P}(\mathbf{X},\mathbf{Z})} > 0$,

2. the $(2,2)$ element of the matrix (12) is positive, i.e., $1 - \left.\dfrac{\partial\psi_2}{\partial P_2}\right|_{\mathbf{P}(\mathbf{X},\mathbf{Z})} > 0$, and

3. the determinant of the matrix (12) is positive, i.e., $\det\left(I_{2\times2} - \nabla_{\mathbf{P}}\Psi(\mathbf{P}(\mathbf{X},\mathbf{Z}),\mathbf{X},\mathbf{Z})\right) > 0$.

The third condition in Proposition 3 is automatically satisfied if $\delta_1(\mathbf{X},\mathbf{Z}_1) \times \delta_2(\mathbf{X},\mathbf{Z}_2) < 0$, or if one player's action is strategic substitute and the other's is strategic complement. The panel (c) in Figure 2 illustrates this case. The sign of $\delta_i(\mathbf{X},\mathbf{Z}_i)$ determines the slope of player i's best response function. A positive $\delta_i(\mathbf{X},\mathbf{Z}_i)$ indicates that the best response function is increasing, while a negative $\delta_i(\mathbf{X},\mathbf{Z}_i)$ indicates that it is decreasing. Thus, when the two slopes have opposite signs, the best response functions are monotone in opposite directions and intersect at one point.
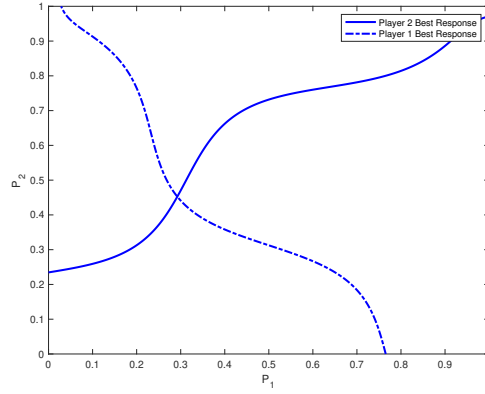
19

Figure 2: Unique Equilibrium and Multiple Equilibria
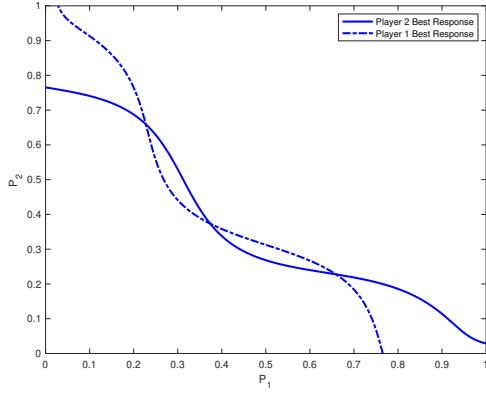
(a) Multiple Equilibria, $\delta_1 \cdot \delta_2 > 0$
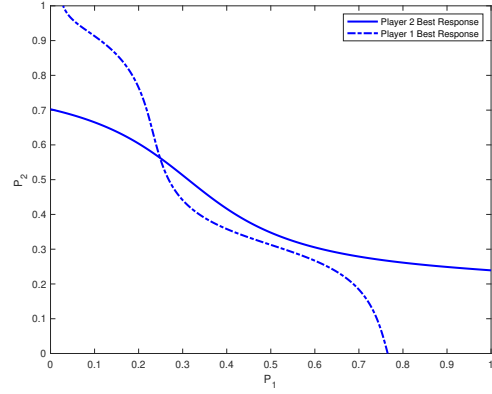
(b) Unique Equilibrium, $\delta_1 \cdot \delta_2 > 0$

(c) Unique Equilibrium, $\delta_1 \cdot \delta_2 < 0$

Notes: The figures illustrate the best response functions. The horizontal axis represents player 1's choice probability $P_1(\mathbf{X}, \mathbf{Z})$, and the vertical axis represents player 2's choice probability $P_2(\mathbf{X}, \mathbf{Z})$. The dotted line represents player 1's best response function, and the solid line represents player 2's best response function. The intersection points of the two lines indicate the equilibria of the game.

3 Identification

In this section, I describe the identification result of model primitives. First, I present the data generating process and the identification conditions. Next, I demonstrate that the model primitives, the payoff function, can be identified.

3.1 Data Generating Process and Identification Conditions

Suppose that an econometrician observes an independent dataset of M different markets, indexed by $m = 1, 2, \ldots, M$. The dataset $\{(\mathbf{y}'_m, \mathbf{x}'_m, \mathbf{z}'_{1m}, \mathbf{z}'_{2m})'\}_{m=1}^M$ consists of the realized action profiles $\mathbf{y}_m$, vectors of market-specific variables $\mathbf{x}_m$, and a vector of player-specific variables $\mathbf{z}_{im}$ for each market $m = 1, 2, \ldots, M$. I assume that both the market-specific and player-specific variables are observable to the econometrician.

Assumption 2. The realizations of the vector $\{(\mathbf{y}'_m, \mathbf{x}'_m, \mathbf{z}'_{1m}, \mathbf{z}'_{2m})'\}_{m=1}^M$ are independent and identically distributed across markets and observable to the econometrician.

Let Z_{ik} denote the k-th variable in $\mathbf{Z}_i$, and let $\mathbf{Z}_{i,-k} = (Z_{i1}, \ldots, Z_{i,k-1}, Z_{i,k+1}, \ldots, Z_{id_z})'$ denote the vector of $\mathbf{Z}_i$ excluding Z_{ik}. I assume that the variable Z_{ik} varies over its support $\mathcal{Z}_{ik}$, conditional on player i's other variables $\mathbf{Z}_{i,-k}$, the rival's variables $\mathbf{Z}_j$, and the market-specific variables $\mathbf{X}$.

The supports of each $\mathbf{Z}_i \in \mathcal{Z}_i \subseteq \mathbb{R}^{d_z}$ can be either discrete or continuous. If $\mathbf{Z}_i$ has a continuous support, then at least one of player-specific variables, Z_{ik} for some k, is continuously distributed for $d_z \geq 1$. In other words, the support of $\mathcal{Z}_{ik}$ contains a non-degenerate interval.

I further assume that the vector of player-specific variables $\mathbf{Z}_i$ enters only player i's payoff function and is excluded from the rival's payoff function. Changes in $\mathbf{Z}_i$ affect only player i's payoff and do not directly affect the rival's payoff. This assumption plays a key role in identification and is referred to as exclusion restriction. Assumption 3 summarizes the conditions on the player-specific variable $\mathbf{Z}_i$.

Assumption 3.   1. The player-specific variable $\mathbf{Z}_i$ can be discrete or continuous, and $\mathbf{Z}_i$ enters player i's payoff but is excluded from the payoffs of other player.

2. The variable Z_{ik} varies over its support $\mathcal{Z}_{ik}$, conditional on $(\mathbf{X}', \mathbf{Z}'_{i,-k}, \mathbf{Z}'_j)'$.

Next, I assume that the equilibrium conditional choice probabilities (CCPs) can be consistently estimated or identified from the data. Specifically, each player's conditional choice probabilities $P_i(\mathbf{X}, \mathbf{Z}_i, \mathbf{Z}_j)$ are assumed to be known to the econometrician. As defined in equation (8), the CCPs vary with both player-specific variables $(\mathbf{Z}'_i, \mathbf{Z}'_j)'$.

Assumption 4. The conditional choice probabilities $P_i(\mathbf{X}, \mathbf{Z}_i, \mathbf{Z}_j)$ vary with $(\mathbf{X}', \mathbf{Z}'_i, \mathbf{Z}'_j)'$. Furthermore, the CCPs are identified or consistently estimated from the data.

Finally, I derive the relationship between equilibrium CCPs and the model primitives. From the equations (9) through (11), the equilibrium CCPs satisfy

$$P_i(\mathbf{X}, \mathbf{Z}) = \int_{\varepsilon_i} \frac{P_i(\mathbf{X}, \mathbf{Z}) \exp\left(\pi_i(\mathbf{X}, \mathbf{Z}_i) + \delta_i(\mathbf{X}, \mathbf{Z}_i) P_j(\mathbf{X}, \mathbf{Z}) + \varepsilon_i\right)^{1/\lambda_i}}{P_i(\mathbf{X}, \mathbf{Z}) \exp\left(\pi_i(\mathbf{X}, \mathbf{Z}_i) + \delta_i(\mathbf{X}, \mathbf{Z}_i) P_j(\mathbf{X}, \mathbf{Z}) + \varepsilon_i\right)^{1/\lambda_i} + (1 - P_i(\mathbf{X}, \mathbf{Z}))} \, dF_i(\varepsilon_i),$$

(13)

for $i, j \in \mathcal{I}$ with $j \neq i$. The terms in the exponent represent player i's expected payoff, and the deterministic component of the expected payoff is

$$EU_i(\mathbf{X}, \mathbf{Z}) = \pi_i(\mathbf{X}, \mathbf{Z}_i) + \delta_i(\mathbf{X}, \mathbf{Z}_i) P_j(\mathbf{X}, \mathbf{Z}). \tag{14}$$

Since the integrand of (13) is strictly increasing in $EU_i(\mathbf{X}, \mathbf{Z})$ and the prior belief $F_i(\varepsilon_i)$ is bijective, there exists a strictly increasing function $G_i(\cdot)$ such that

$$P_i(\mathbf{X}, \mathbf{Z}) = G_i\left(\frac{1}{\lambda_i} EU_i(\mathbf{X}, \mathbf{Z})\right) = G_i\left(\frac{1}{\lambda_i}\left(\pi_i(\mathbf{X}, \mathbf{Z}_i) + \delta_i(\mathbf{X}, \mathbf{Z}_i) P_j(\mathbf{X}, \mathbf{Z})\right)\right). \tag{15}$$

This equation shows that the equilibrium CCPs can be expressed as a function of the players' deterministic components of expected payoffs, providing a key link between the CCPs and the underlying structure of the model.

3.2 Identification Result: Semiparametric Model

Given these assumptions, I now turn to the identification of model primitives. In this subsection, the model primitives — the base payoffs $\pi_i(\cdot)$ and the competitive effects $\delta_i(\cdot)$ — are nonparametric functions, and I demonstrate that they are identified.

Fix an arbitrary realization of market-specific variables and player i's variables, $\mathbf{X} = \mathbf{x}$ and $\mathbf{Z}_i = \mathbf{z}_i$. Given the equilibrium choice probabilities $P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{Z}_j)$, I can apply the mapping in equation (15) to recover $EU_i(\mathbf{x}, \mathbf{z}_i, \mathbf{Z}_j)$. Recall that $G_i(\cdot)$ is a bijective function and hence invertible. Therefore,

$$G_i^{-1}(P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{Z}_j)) = \frac{1}{\lambda_i}\left(\pi_i(\mathbf{x}, \mathbf{z}_i) + \delta_i(\mathbf{x}, \mathbf{z}_i) P_j(\mathbf{x}, \mathbf{Z}_j, \mathbf{z}_i)\right). \tag{16}$$

Given that $\lambda_i > 0$, model primitives $\{\pi_i(\cdot), \delta_i(\cdot), \lambda_i\}$ are only identified up to scale. Multiplying $\pi_i(\cdot)$, $\delta_i(\cdot)$ and λ_i by a positive constant $c > 0$ does not change the right-hand side of equation (16). I therefore normalize the unit cost of information λ_i to one. Under this normalization, the model primitives, $\pi_i(\cdot)$, and $\delta_i(\cdot)$ can be identified.

A necessary condition for identification of $\pi_i(\cdot)$ and $\delta_i(\cdot)$ is that, for each $\mathbf{x}$ and $\mathbf{z}_i$, there exist at least two distinct points, $\mathbf{z}_j^1$ and $\mathbf{z}_j^2$ in the support of $\mathbf{Z}_j$ conditional on $\mathbf{x}$ and $\mathbf{z}_i$. A

sufficient condition for identification is that the two distinct points satisfy the following:

$$P_j(\mathbf{x}, \mathbf{z}_j^1, \mathbf{z}_i) \neq P_j(\mathbf{x}, \mathbf{z}_j^2, \mathbf{z}_i).$$

By Assumption 4, the conditional choice probabilities $P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{Z}_j)$ and $P_j(\mathbf{x}, \mathbf{Z}_j, \mathbf{z}_i)$, vary sufficiently with $\mathbf{Z}_j$, so the inequality condition holds. Therefore, variation in $P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{Z}_j)$ and $P_j(\mathbf{x}, \mathbf{Z}_j, \mathbf{z}_i)$ across $\mathbf{Z}_j$ allows the identification of $\pi_i(\cdot)$ and $\delta_i(\cdot)$.

Proposition 4. Suppose that Assumptions 1 through 4 hold and the unit cost of information λ_i is normalized to 1. Suppose further that, for each $\mathbf{x} \in \mathcal{X}$ and $\mathbf{z}_i \in \mathcal{Z}_i$, there exist at least two distinct points, $\mathbf{z}_j^1$ and $\mathbf{z}_j^2$ in the support of $\mathbf{Z}_j$ conditional on $\mathbf{x}$ and $\mathbf{z}_i$, such that

$$P_j(\mathbf{x}, \mathbf{z}_j^1, \mathbf{z}_i) \neq P_j(\mathbf{x}, \mathbf{z}_j^2, \mathbf{z}_i).$$

Then, for each $i \in \mathcal{I}$, the base payoff $\pi_i(\mathbf{X}, \mathbf{Z}_i)$ and the competitive effects $\delta_i(\mathbf{X}, \mathbf{Z}_i)$ are nonparametrically identified.

Proof. In the Appendix ☐

The proposition shows that, under the maintained assumptions and sufficient variation in $\mathbf{Z}_j$, which induces variation in $P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{Z}_j)$ and $P_j(\mathbf{x}, \mathbf{Z}_j, \mathbf{z}_i)$, both the structural base payoffs and the strategic effects can be nonparametrically identified from the equilibrium choice probabilities. Intuitively, variation in $\mathbf{Z}_j$ changes both players' equilibrium choice probabilities without directly affecting the payoff functions of player i. By observing how player i's choice probabilities respond to changes in $\mathbf{Z}_j$, the payoff primitives can be identified even in the presence of costly information acquisition.

3.3   Identification Result: Parametric Model

In this subsection, the model primitives are specified as parametric functions, and I show that the parameter vector is point identified. Let the payoff function $\pi_i(\cdot)$ and the competitive effect $\delta_i(\cdot)$ be defined by parametric functions as follows:

$$\pi_i(\mathbf{X}, \mathbf{Z}_i) = \mathbf{X}\beta_i + \mathbf{Z}_i\gamma_i,$$
$$\delta_i(\mathbf{X}, \mathbf{Z}_i) = \delta_i,$$

where $\beta_i \in \mathbb{R}^{d_x}$, $\gamma_i \in \mathbb{R}^{d_z}$, and $\delta_i \in \mathbb{R}$ are constant parameters. The competitive effect δ_i may also take a linear functional form, such as $\delta_i(\mathbf{X}, \mathbf{Z}_i) = \mathbf{X}\alpha_i + \mathbf{Z}_i\kappa_i$, where $\alpha_i \in \mathbb{R}^{d_x}$ and $\kappa_i \in \mathbb{R}^{d_z}$ are constant parameters.

The deterministic expected payoff (14) can be expressed as

$$EU_i(\mathbf{X}, \mathbf{Z}, \theta_i) = \mathbf{X}\beta_i + \mathbf{Z}_i\gamma_i + \delta_i \cdot P_j(\mathbf{X}, \mathbf{Z}, \theta_j),$$

where $\theta_i = (\beta_i', \gamma_i', \delta_i)'$ is the parameter vector.

To identify the parameters, I adopt the strategy known as identification at infinity following Tamer (2003). Identification requires an assumption that there exists one variable Z_{ik} such that $\gamma_{ik} \neq 0$ and has an everywhere positive Lebesgue density conditional on $\mathbf{Z}_{i,-k}$. This assumption ensures that the variation in Z_{ik} can be used to identify the effect of Z_{ik} on the payoff function, which is crucial for point identification of the parameter vector. Moreover, this assumption is standard in the literature on discrete choice games (Tamer, 2003; Ciliberto and Tamer, 2009). Under the assumption, the parameter vector $\theta = (\theta_i', \theta_j')'$ is point identified.

Proposition 5. Suppose that Assumptions 1 through 4 hold. I further assume that there exists one variable Z_{ik} such that $\gamma_{ik} \neq 0$ and has an everywhere positive Lebesgue density conditional on $\mathbf{Z}_{i,-k}$. Then the parameter vector θ is point identified if the matrix $(\mathbf{X}, \mathbf{Z}_i)$ has full column rank for all $i \in \mathcal{I}$.

Proof. In the Appendix □

This result formalizes the point identification of the parametric version of the model. While the nonparametric setting established identification of the underlying payoff primitives, this proposition demonstrates that, under rank conditions, the finite-dimensional parameter vector θ can be uniquely identified from the equilibrium choice probabilities. Hence, the model provides a fully identified structure that can be directly estimated using standard econometric estimation methods.

4 Monte Carlo Experiments

In this section, I illustrate the identification results in the Monte Carlo experiments.

4.1 Parametric Model

I conduct a Monte Carlo experiment to illustrate the parametric identification results of games with costly information acquisition. I construct a sample of M independent markets with two players and two actions. The payoff function for each player is linear in

parameters, and the payoff for each player from $Y_i = 1$ is given by

$$U_1(Y_2, \varepsilon_1; Z_1, \theta) = 1.8 + 0.5Z_1 - 1.3 \cdot \mathbb{1}(Y_2 = 1) + \varepsilon_1,$$

$$U_2(Y_1, \varepsilon_2; Z_2, \theta) = 1.6 + 0.8Z_2 - 1.3 \cdot \mathbb{1}(Y_1 = 1) + \varepsilon_2,$$

where θ is the vector of parameters. The payoff from $Y_i = 0$ is zero for both players. The true parameters in the data generating process $\theta^0 = (\beta_1^0, \gamma_1^0, \beta_2^0, \gamma_2^0, \delta_1^0, \delta_2^0)$ are given by $(1.8, 0.5, 1.6, 0.8, -1.3, -1.3)$. I assume no market-specific variables in this exercise, and the player-specific variable Z_i is randomly drawn from the uniform distribution with support $[0, 1]$. Each player's prior belief $F_i(\varepsilon_i)$ follows the normal distribution with mean zero and variance 4.

The sample data are generated from the unique equilibrium in each market conditional on (Z_1, Z_2), and I estimate the model using the maximum likelihood (ML) with the constrained optimization approach (Su, 2014). One challenge in estimating discrete games with incomplete information is the potential existence of multiple equilibria, which makes computing all the equilibria often infeasible. As a result, the ML estimation is often demanding and computationally burdensome. In this exercise, however, the equilibrium is unique by construction, and thus the ML estimation is feasible. Among the alternative estimation methods, such as two-step estimators and pseudo-likelihood estimators, the constrained ML approach is known to be efficient (Su, 2014).

Let $\mathbf{P}_m = (P_{1m}, P_{2m})'$ be a vector of choice probabilities in market m. Given the parameter vector θ and the data $\mathbf{z}_m = (z_{1m}, z_{2m})'$ and $\mathbf{y}_m = (y_{1m}, y_{2m})'$, the log likelihood function in market m is

$$\ell_m(\mathbf{y}_m, \mathbf{z}_m, \mathbf{P}_m, \theta) = \sum_{i=1,2} y_{im} \log(P_{im}) + (1 - y_{im}) \log(1 - P_{im}).$$

The log likelihood for all markets $m = 1, 2, \ldots, M$ is

$$\mathcal{L}(\mathbf{P}, \theta) = \sum_{m=1}^{M} \ell_m(\mathbf{y}_m, \mathbf{z}_m, \mathbf{P}_m, \theta), \tag{17}$$

where $\mathbf{P} = (\mathbf{P}_m)_{m=1}^{M}$ is the vector of choice probabilities for each market.

The probability $\mathbf{P}$ in the equation (17) is not necessarily an equilibrium choice probability. I impose the Nash equilibrium restriction (11) as a constraint. Hence, the ML estimation is to solve the maximization problem:

$$\max_{\mathbf{P}, \theta} \mathcal{L}(\mathbf{P}, \theta) \quad \text{s.t.} \quad \mathbf{P} - \Psi(\mathbf{P}, \mathbf{z}, \theta) = 0. \tag{18}$$

I report the estimation results in Table 2. The table shows the result for samples of $M = 1,000$ markets and $R = 300$ repetitions. The result displays strong evidence of convergence of each estimates to the true parameters.

Parameter	True Value	Mean	Median	S.D.	RMSE	Mean Bias	Median Bias
β_1	1.8	1.7803	1.7432	0.4633	0.4638	-0.0197	-0.0568
γ_1	0.5	0.5050	0.5050	0.1558	0.1559	0.0050	0.0050
β_2	1.6	1.6172	1.6127	0.3679	0.3683	0.0172	0.0127
γ_2	0.8	0.8060	0.8092	0.1535	0.1537	0.0060	0.0092
δ_1	-1.3	-1.2754	-1.2194	0.5847	0.5853	0.0246	0.0806
δ_2	-1.3	-1.3241	-1.3207	0.4469	0.4476	-0.0241	-0.0207

Table 2: Monte Carlo Results: Parametric Model, $M = 1000$

4.2 Semiparametric Model

Next, I illustrate the semiparametric identification results of games with costly information acquisition. In this Monte Carlo experiment, the maximum likelihood estimation by the method of sieve (Chen, 2007) in implemented to estimate nonparametric functions of the base return and the competitive effect.

The data generating process mainly follows Xie (2024). I construct a sample of M independent markets. I consider two identical players in each market. The payoff function for each player from $Y_i = 1$ is given by

$$\pi_i(Z_i) = 3 - \log(1 + 2Z_i),$$
$$\delta_i(Z_i) = -\frac{4 \exp(Z_i)}{1 + \exp(Z_i)},$$

where Z_i is randomly drawn from the uniform distribution with the support $[0, 3]$. The payoff from $Y_i = 0$ is zero for both players. Each player's prior belief $F_i(\varepsilon_i)$ follows from the uniform distribution with the support $[-5, 5]$.

Let $\theta_i = (\pi_i, \delta_i)$ be model primitives. θ_i contains all unknown functions which is to be estimated by the sieve MLE. Moreover, each element of the model primitives θ_i satisfies $\pi_i \in \Pi$ and $\delta_i \in \Delta$, where Π and Δ are the spaces of functions π_i and δ_i. The approximating spaces, i.e. sieves, are represented by $\Pi_M \subseteq \Pi$ and $\Delta_M \subseteq \Delta$, and the sieve space for the model primitive θ_i is $\Theta_M = \Pi_M \times \Delta_M$.

I use the space of polynomials as the sieve space for both π_i and δ_i. The sieve spaces are

$$\Pi_M = \left\{ \sum_{l=0}^{d_{\pi,M}} \beta_l \cdot Z_i^l \,\middle|\, \beta_l \in \mathbb{R} \right\}, \quad \Delta_M = \left\{ \sum_{l=0}^{d_{\delta,M}} \beta_l \cdot Z_i^l \,\middle|\, \beta_l \in \mathbb{R} \right\}.$$

The equilibrium choice probabilities follow from (15)

$$P_i(Z_i, Z_j; \theta_i) = G_i \left(\pi_i(Z_i) + \delta_i(Z_i) P_j(Z_j, Z_i; \theta_j) \right), \tag{19}$$

where $G_i(\cdot)$ is the inverse of the uniform distribution with the support $[-5, 5]$.

Under the regularity conditions (Chen, 2007), the sieve MLE is consistent, and the model primitives θ_i can be estimated by maximizing the following log-likelihood function

$$\max_{\theta_i \in \Theta_M} \sum_{i=1}^{2} \sum_{m=1}^{M} y_{im} \log \left(P_{im}(z_i, z_{3-i,m}, \theta_i) \right) + (1 - y_{im}) \log \left(1 - P_{im}(z_{im}, z_{3-i,m}, \theta_i) \right) \tag{20}$$

Figure 3 shows the estimation results. I use the sample of $M = 400$ independent markets, and repeated $R = 100$ times. The degree of polynomial was set to 2. The blue solid lines in each panel indicate the true value of π_i and δ_i. The red solid lines in each panel represent the averages of the estimates from 100 repetition. The red dashed lines show 90% confidence intervals. The sieve MLE give estimates that are consistent and close to the true model primitives.

5 Empirical Application

In this section, I apply my model to the market-level entry decisions of airline firms in the U.S. airline industry.

5.1 Motivation

I investigate the U.S. airline industry as an empirical application. In this industry, a market is defined as a city pair. Each airline's entry decision involves deciding whether or not to serve a route between two cities.

In contrast to much of the existing literature, my model assumes that firms do not perfectly know their profits at the time they make their entry decisions due to payoff shocks. In previous models, firms are assumed to observe their payoff shocks, and thus their profits when making decisions. However, in my model, the uncertainty arising from unobserved payoff shocks induces firms to acquire information.
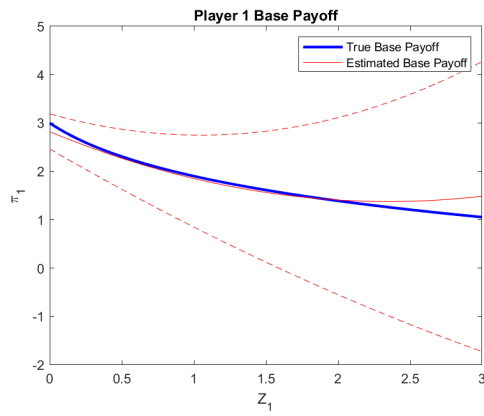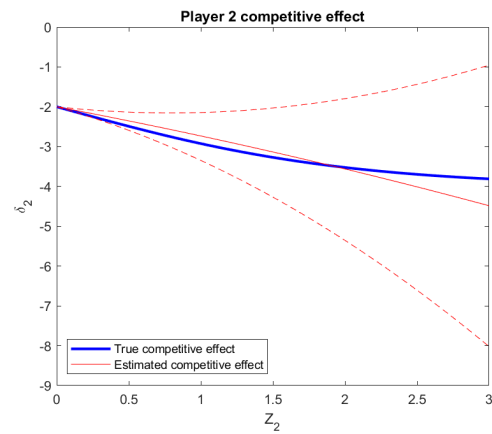
(a) Player 1's base return
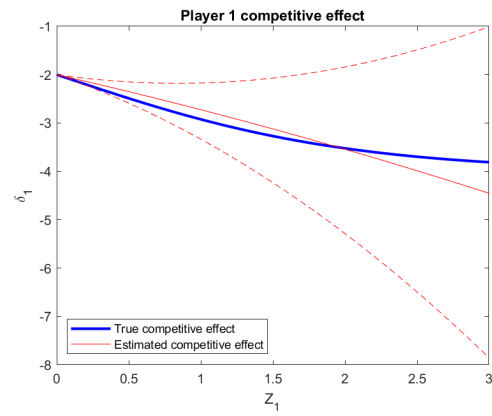
(b) Player 1's competitive effect
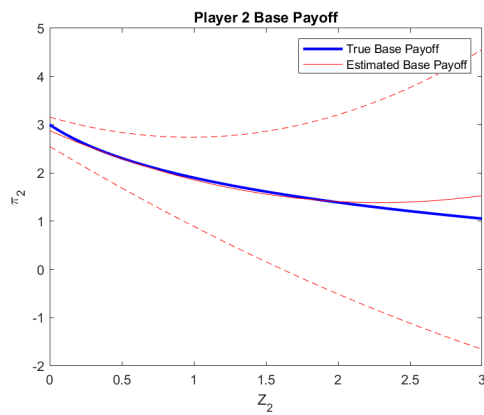
(c) Player 2's base return

(d) Player 2's competitive effect

Figure 3: Monte Carlo Results: Semiparametric Model, $M = 400$.

The uncertainty in profits, and thus the need for information acquisition, stems from three sources. First, airlines do not know the exact demand for a route before selling flight tickets and therefore need to gather information to forecast it. If a firm serves a route that attracts few passengers, costs can exceed revenue, leading to negative profits. In such a case, the firm would be better off not entering the market. Thus, gathering information about future demand plays a key role in their decision making.

Second, the operational feasibility at airports is important for airlines' entry decisions. This includes factors such as airport infrastructures, slot availability, and ground handling capacities at a certain airport. Assessing the operational feasibility might require airlines to acquire information from airport authorities.

Lastly, regulatory environments are a crucial factor in airlines' decisions. To provide services between two cities, an airline needs to comply with the local aviation regulations and acquire information about them. Moreover, taxes and fees can differ across airports, and the firms should be aware of these prior to the operation to make an optimal decision.

5.2 Data

My primary dataset on the airline industry comes from Kline and Tamer (2016) [8]. According to them, the data are collected from the second quarter of the 2010 Airline Origin and Destination Survey (DB1B). The unit of observation is a market, and each observation contains information on airlines' market entry decisions and the market-specific and firm-specific explanatory variables.

The data contain 7,882 markets indexed by $m = 1, 2, \ldots, 7882$, and each market is defined as routes between two airports irrespective of stopping. As in Kline and Tamer (2016), I consider the entry behavior of two players: (1) low-cost carriers (LCC) [9], and (2) other airlines (OA) [10]. In the following analysis, the players LCC and OA are treated as a single player, and $Y_{LCC} = 1$ (respectively, $Y_{OA} = 1$) means that at least one airline firm in LCC (respectively, OA) serves the market. On the contrary, $Y_{LCC} = 0$ (respectively, $Y_{OA} = 0$) means that none of firms in LCC (respectively, OA) serves the market. The uncondi-

---

[8]Chen et al. (2018) and Gu et al. (2025) use the same dataset and apply their models to airline entry game. In addition, Kline and Tamer (2024) uses a similar dataset, except that the data are collected from the year 2023.

[9]According to Kline and Tamer (2016), AirTran, Allegiant Air, Frontier, JetBlue, Midwest Air, Southwest, Spirit, Sun Country, USA3000, and Virgin America are aggregated into one single player, LCC.

[10]All airlines other than low-cost carriers are aggregated into one single player, OA.

tional choice probabilities are $(\widehat{\Pr}(0,0), \widehat{\Pr}(1,0), \widehat{\Pr}(0,1), \widehat{\Pr}(1,1)) = (0.15, 0.07, 0.61, 0.16)$, where $\widehat{\Pr}(Y_{LCC}, Y_{OA})$ is the sample frequency of the market outcome (Y_{LCC}, Y_{OA}).

Two explanatory variables are used in my empirical analysis. The first explanatory variable is a market- and firm-specific variable which has been used in many papers: market presence (Berry, 1992; Ciliberto and Tamer, 2009). The market presence, denoted by Z_{im}, is a proxy for a scale of operation of airline i in market m (Berry, 1992). Moreover, the market presence is a measure of the airline's market power and profitability since a large scale of operation may deter entry of other airlines (Gu et al., 2025). The market presence Z_{im} varies across airlines and only enters airline i's profits. This variable satisfies the exclusion restriction and is essential for the payoff identification. The market presence of airline i at a given market m is computed as follows. For each airline and for each airport, the number of routes served by the airline from the airport is counted. Next, the proportion of routes served from the airport is calculated by dividing the number of routes served by the airline by the total number of routes from that airport by all airlines. The market presence for each airline at a given market m is the average of these proportions at two endpoints of the market. Since there are two players in my analysis, the market presence variable is the maximum of market presence for each airline in the group of LCC and OA.

The second explanatory variable is market-specific variable: market size. The market size, denoted by X_m, is defined by the population of the endpoints [11]. Both LCC and OA in the same market have the same market size, and the market size varies across the markets. In the analysis, the unit of market size is billions of people.

Table 3 reports the summary statistics of my data. The decision variables are binary, and both the market presence and market size are continuous variables. Apparently, the entry probability of OA is about 3 times higher than LCC. Moreover, the airport presence of OA, a measure of operation scale, is about 7.9 times higher than LCC. On average, markets that LCC serves tend to be larger than the markets that OA serves.

---

[11]The original data source, Kline and Tamer (2016), does not explicitly specify which of the two endpoints' population was used, nor does it clarify the source of the population data. However, Gu et al. (2025), which uses the same dataset, mentions that the population data is collected from the U.S. Census Bureau in 2010 and the market size is defined by the maximum of the two endpoints' population.

Table 3: Summary Statistics

	Mean	Std. Dev.	Median	Max	Min	Obs.
LCC Entry	0.2322	0.4222	0	1	0	7,882
OA Entry	0.7742	0.4182	1	1	0	7,882
Market Presence	0.1986	0.1596	0.1000	0.5007	0	15,764
Market Presence (*LCC*)	0.0447	0.0204	0.0429	0.1128	0	7,882
Market Presence ($Y_{LLC} = 1$)	0.0676	0.0154	0.0677	0.1128	0.0149	1,830
Market Presence ($Y_{LLC} = 0$)	0.0377	0.0161	0.0386	0.1053	0	6,052
Market Presence (*OA*)	0.3525	0.0562	0.3581	0.5007	0.0349	7,882
Market Presence ($Y_{OA} = 1$)	0.3646	0.0437	0.3652	0.5007	0.1318	6,102
Market Presence ($Y_{OA} = 0$)	0.3111	0.0723	0.3279	0.4655	0.0349	1,780
Market Size	1,229,550	648,935	1,067,422	4,177,793	171,466	7,882
Market Size ($Y_{LLC} = 1$)	1,310,375	662,521	1,159,922	4,044,754	231,001	1,830
Market Size ($Y_{LLC} = 0$)	1,205,111	642,828	1,046,574	4,177,793	171,466	6,052
Market Size ($Y_{OA} = 1$)	1,286,312	646,809	1,131,214	4,177,793	173,002	6,102
Market Size ($Y_{OA} = 0$)	1,034,966	617,953	874,540	4,177,793	171,466	1,780

5.3 Preliminary Analysis

Before estimating the structural model, I conduct a preliminary analysis using simple Probit regressions to examine the data patterns. The Probit regression equation is given by the following equation:

$$Y_{im} = \mathbb{1}\left\{\beta^{cons} + \beta^{pres} Z_{im} + \beta^{size} X_m + \delta Y_{jm} + \varepsilon_{im} \geq 0\right\}, \tag{21}$$

where the error term ε_{im} follows the standard normal distribution. In simple Probit regressions, each observation represents a specific player within a particular market. The dependent variable is entry decisions, $Y_{im} \in \{0, 1\}$. I regress the binary decisions on firm-specific and market-specific variables.

Table 4 reports the estimation results from the Probit regressions. The coefficients indicate the impact of each variable on the underlying latent profit for the entry decision. The estimation results are reasonable in two ways. First, both market presence and market

Table 4: Estimation results from Probit regressions

	(1)	(2)	(3)
Market Presence	5.6688	5.7304	5.3937
	(0.0757)	(0.0768)	(0.0899)
Market Size		0.2396	0.2582
		(0.0182)	(0.0185)
Competitive Effect			-0.1936
			(0.0277)
Constant	-1.1020	-1.4030	-1.2632
	(0.0186)	(0.0300)	(0.0358)
Log Likelihood	-7648.11	-7560.80	-7536.69
Observations	15,764	15,764	15,764

Notes: This table reports the estimation results of the Probit regressions. Each observation represents a specific player within a particular market. Standard errors are in the parentheses.

size are positively correlated with the entry decisions. This result is consistent with the intuition that an airline with a large operational scale in a large demand market is more likely to serve the route. Second, the competitive effect is negatively correlated with the entry decisions, indicating that a rival's entry reduces the player's profits.

Although the preliminary analysis result is consistent with the intuition on airlines' entry decisions, the estimation results may be biased. In the above analysis, the airlines' profits are assumed to be perfectly observable to airlines. As I have discussed in the previous section, the profits may not be perfectly known to airlines due to uncertainty, operational feasibility, and regulatory environments. I model the structural entry game with costly information acquisition to properly incorporate strategic interactions in the next subsection.

5.4   Structural Entry Game Analysis

I develop a standard parametric entry game that builds on the existing literature by adding costly information acquisition. The two firms are *LCC* and *OA*, and they choose a binary

action $Y_{im} \in \{0,1\}$ where $Y_{im} = 1$ if firm i enters a market m and $Y_{im} = 0$ if i stays out. The profit function for firm i is specified as follows:

$$U_{im}(Y_{im}, Y_{jm}, \varepsilon_{im}; X_m, Z_{im}, \theta) = Y_{im} \left(\beta_i^{cons} + \beta_i^{pres} Z_{im} + \beta_i^{size} X_m + \delta_i Y_{jm} + \varepsilon_{im} \right), \quad (22)$$

where θ is the vector of parameters of interest.

Firm i's profit from entering the market is $\beta_i^{cons} + \beta_i^{pres} Z_{im} + \beta_{im}^{size} X_m + \delta_i Y_{jm} + \varepsilon_{im}$ where Z_{im} represents the market presence, X_m represents the market size, Y_{jm} represents rival's entry, and ε_{im} is the payoff shock. The payoff shock ε_{im} is unobserved by firm i (and also by rival j and econometricians) when the firm i makes entry decision. The profit from staying out is normalized to zero. I assume that each player's prior belief about ε_{im} follows an i.i.d. standard normal distribution. This is because the normal distribution describes the maximum entropy among all distributions with the same variance. Given this prior belief, firms make entry decisions under the condition of maximum uncertainty.

I estimate the parameters by using the two-step estimation method due to multiple equilibria. Under the multiple equilibria, a vector of parameters θ corresponds to multiple choice probabilities. Since I do not know which choice probability is selected by the firms in the data, I estimate the choice probabilities in the first step. In the second step, given the choice probabilities, I estimate the parameters by maximizing the pseudo log-likelihood function.

In the first step, I nonparametrically estimate the conditional choice probabilities. I denote the conditional choice probability of entering the market as $P_{im}(x_m, \mathbf{z}_m) = P_{im}(Y_{im} = 1 | x_m, \mathbf{z}_m)$. Since the entry decision is binary, the conditional choice probability of staying out can be expressed as $1 - P_{im}(x_m, \mathbf{z}_m)$. The conditional choice probabilities are estimated using sieve method (Chen, 2007). Let $q^\kappa(x_m, \mathbf{z}_m) = (q_1(x_m, \mathbf{z}_m), \ldots, q_\kappa(x_m, \mathbf{z}_m))^\top$ be the $\kappa \times 1$ vector of basis functions with degree κ. The sieve estimator of the conditional choice probabilities are:

$$\widehat{P}_{im}(x_m, \mathbf{z}_m) = \frac{\exp(q^\kappa(x_m, \mathbf{z}_m)'\hat{\gamma})}{1 + \exp(q^\kappa(x_m, \mathbf{z}_m)'\hat{\gamma})}, \quad (23)$$

where $\hat{\gamma}$ is the maximizer of the log likelihood function

$$\hat{\gamma} = \arg\max_{\gamma} \sum_{m=1}^{M} \left[\mathbb{1}\{Y_{im} = 1\} q^\kappa(x_m, \mathbf{z}_m)'\gamma - \ln\left(1 + \exp(q^\kappa(x_m, \mathbf{z}_m)'\gamma)\right) \right]. \quad (24)$$

In the first step estimation, I set the degree $\kappa = 3$. The sieve estimator $\widehat{P}_{im}(x_m, \mathbf{z}_m)$ will converge to the true conditional choice probability $P_{im}(x_m, \mathbf{z}_m)$.

In the second step estimation, I take as given the first step estimates $\widehat{P}_{im}(x_m, \mathbf{z}_m)$ of the conditional choice probabilities. I then construct the pseudo log-likelihood function and find the maximizer of the pseudo log-likelihood function. For a given value of $(\mathbf{x}, \mathbf{z})$, for a given $y = (y_{LCC}, y_{OA})$, I estimate the parameters θ by maximizing the following pseudo log-likelihood function:

$$\ell(\theta, \widehat{\mathbf{P}}) = \sum_{m=1}^{M} \sum_{i} y_{im} \ln \Psi_i \left(\widehat{\mathbf{P}}_m(x_m, \mathbf{z}_m), x_m, \mathbf{z}_m, \theta \right) + (1 - y_{im}) \ln \left(1 - \Psi_i \left(\widehat{\mathbf{P}}_m(x_m, \mathbf{z}_m), x_m, \mathbf{z}_m, \theta \right) \right),$$

(25)

where $\Psi_i(\cdot)$ is the equilibrium condition defined by

$$\Psi_i \left(\widehat{\mathbf{P}}_m(x_m, \mathbf{z}_m), x_m, \mathbf{z}_m, \theta \right)$$

$$= \int_{\varepsilon_i} \frac{\widehat{P}_{im}(x_m, \mathbf{z}_m) \exp \left(\beta_i^{cons} + \beta_i^{pres} z_{im} + \beta_i^{size} x_m + \delta_i \widehat{P}_j(x_m, \mathbf{z}_m) + \varepsilon_{im} \right)}{\widehat{P}_{im}(x_m, \mathbf{z}_m) \exp \left(\beta_i^{cons} + \beta_i^{pres} z_{im} + \beta_i^{size} x_m + \delta_i \widehat{P}_{jm}(x_m, \mathbf{z}_m) + \varepsilon_{im} \right) + 1 - \widehat{P}_{im}(x_m, \mathbf{z}_m)} \mathrm{d}\Phi(\varepsilon_{im}).$$

5.5 Estimation Result

Table 5 presents the estimation results under the assumption that players acquire costly information. For both LCC and OA, the coefficients on market presence and market size are positive and statistically significant. This means that markets with existing airline presence and larger market sizes generally offer higher expected profits for entry. The constant terms for both LCC and OA, which capture the average profit from entry independent of the observed firm and market characteristics, are negative and statistically significant. The negative constant term can be interpreted as the entry barrier faced by firms, and the results indicate that this barrier is more substantial for LCC than OA.

The competitive effects are negative for both LCC and OA, consistent with the expectation that increased competition reduces a firm's entry profit. However, the coefficient for LCC is not statistically significant. This suggests that, conditional on other variables, the entry of OA does not influence LCC's entry decision. By contrast, LCC's entry decision has a statistically significant negative effect on OA's entry decision. In the context of costly information acquisition, this implies that LCC relies more heavily on its own payoff uncertainty, making the acquisition of precise information more critical for its decision.

Table 5: Estimation results of the entry game with costly information acquisition

	LCC	OA
Market Presence	16.4100	2.5118
	(0.5552)	(0.1185)
Market Size	0.0407	0.0853
	(0.0124)	(0.0104)
Competitive Effect	-0.0301	-0.1963
	(0.0651)	(0.0517)
Constant	-1.0847	-0.7013
	(0.0653)	(0.0469)
Log Likelihood	-5839.21	
Observations	7882	

Notes: This table reports the estimation results of the model with costly information acquisition. The standard errors are computed using the bootstrap method with 1,000 replications. The standard errors are in parentheses.

5.6 Information Acquisition

I can compute the amount of information acquired by each firm in each market using the estimated parameters. The amount of information acquired by firm i in market m is given by

$$I_{im}(F_{im}, P_{im}(\mathbf{x}_m, \mathbf{z}_m)) = H(F_{im}(\cdot)) - \mathrm{E}\left[H(P_{im}(\cdot|y_{im}; \mathbf{x}_m, \mathbf{z}_m))\right], \tag{26}$$

where $P_{im}(\mathbf{x}_m, \mathbf{z}_m)$ is the equilibrium choice probability and $P_{im}(\cdot|, y_{im}; \mathbf{x}_m, \mathbf{z}_m)$ is the posterior beliefs.

Table 6 reports the mean, median, and maximum of the amount of information acquired by LCC and OA. The mean of the amount of information acquired by LCC is from 0.0363, while that by OA is 0.0543. The LCC tends to acquire less information than OA. The same pattern is observed in the median and maximum of the amount of information acquired. The median and the maximum of the amount of information acquired by LCC are 0.0250 and 0.1198, respectively, which are lower than those of OA. These results indicate that LCC tends to acquire less information than OA, which may be due to their

business model that focuses on cost efficiency.

Table 6: The Amount of Information Acquired

	LCC	OA
Mean	0.0363	0.0543
Median	0.0250	0.0546
Max	0.1198	0.1330

Figure 4 presents the distribution of the amount of information acquired by *LCC* and *OA*, along with fitted normal density functions. The distribution of the amount of information acquired by *OA* appears to have a higher mean and a smaller variance than that of *LCC*. This indicates that *OA* tends to acquire more information than *LCC* and suggests systematic differences in the role of information across the two firms.
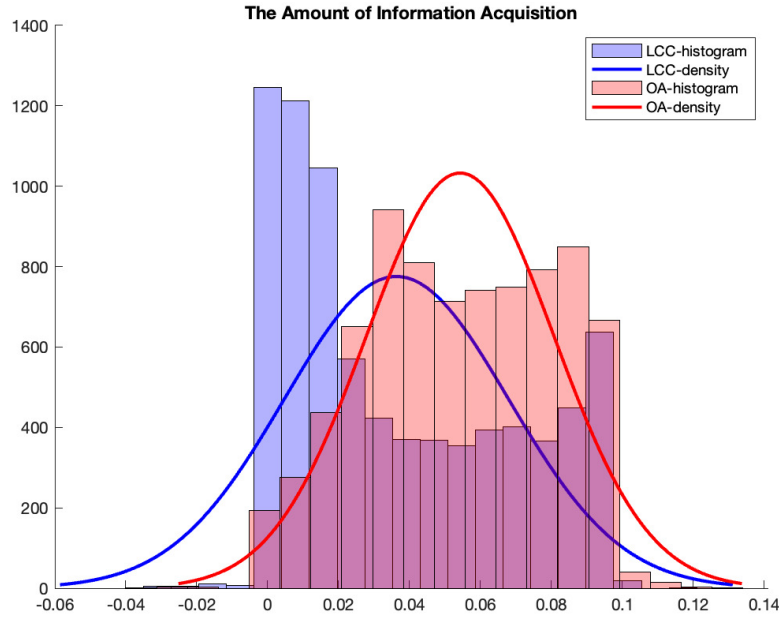

Figure 4: The amount of information acquired

Notes: This figure shows the histogram of the estimated amount of information acquired by *LCC* and *OA*. The solid lines represent the fitted normal density functions. The negative values of the amount of information acquired are due to the numerical approximation errors.

5.6.1 Comparison to Perfectly Private Information Structure

I compare the model with costly information acquisition to the model without such costs. When the cost of information is zero, each player acquires full information about the exact value of their own payoff shock ε_i. This corresponds to the perfectly private information structure studied in Seim (2006); Bajari et al. (2010); Aradillas-Lopez (2010). The key distinction between the two models lies in whether firms observe their own payoff shock when making entry decisions. In my model with costly information acquisition, firms do not observe payoff shocks prior to entry decisions. In contrast, under the perfectly private information structure, firms perfectly observe their own payoff shock. The estimation results for the perfectly private information structure model are reported in Table 7.

In the model without information acquisition, the estimated coefficients on market presence and market size are substantially larger for both LCC and OA compared to the model with costly information acquisition. This suggests that when firms observe their private payoff shocks, the observed market characteristics become more influential in explaining their entry decisions. The constant terms also become significantly more negative in the perfectly private information model, implying a larger barrier of entry.

The estimated competitive effects differ across the two models. In the model with costly information acquisition, the absolute value of estimates of competitive effect for LCC is smaller than that for LCC in perfectly private information model. However, the competitive effect remains insignificant for LCC. The estimates for OA show similar results as in the perfectly private information model.

All the coefficients, except for the LCC's competitive effect, increase in the magnitude of coefficients from the perfectly private information model, as compared to the costly information acquisition model. This suggests that when firms are modeled with perfectly private information structure, the estimation may attribute a greater influence to observed firm and market characteristics. When firms are assumed to perfectly observe their own payoff shocks, conditional choice probabilities are more affected by observable variables, such as market presence and market size. As a result, the model may overestimate the influence of these variables.

Finally, the lower log likelihood for the perfectly private information model indicates that the model with costly information acquisition provides a better statistical fit to the observed data. This finding supports the idea that firms' actual decision making likely

37

incorporates a more sophisticated, endogenous information strategy.

Table 7: Estimation results of the entry game without information acquisition

	LCC	OA
Market Presence	62.9220	9.6254
	(1.4240)	(0.3466)
Market Size	0.1141	0.3315
	(0.0331)	(0.0302)
Competitive Effect	-0.0091	-0.3254
	(0.1057)	(0.0610)
Constant	-4.1228	-2.8645
	(0.1245)	(0.1242)
Log Likelihood	-5986.84	
Observations	7882	

Notes: This table reports the estimation results of the perfectly private information model. In this information structure, each firm only observe their own payoff shock ε_i, but not rival's payoff shock ε_j. The standard errors are in the parentheses.

6 Conclusion

This paper has proposed a novel econometric framework for static discrete games with costly information acquisition. In contrast to the existing literature, in the model, players acquire information about their own payoffs, but acquiring the information incurs costs, which is proportional to the mutual information. I characterized the Nash equilibrium conditions and established equilibrium existence and uniqueness condition. I showed that the model primitives are identified for for both parametric and semiparametric settings, and I presented the identification results via Monte Carlo experiments. Finally, I applied the proposed framework to U.S airline market entry decisions and found that the low-cost carriers acquire less information about profits compared to other airlines, which may be due to their business model that focuses on cost efficiency.

References

Almog, D. and D. Martin (2024). Rational inattention in games: experimental evidence. *Experimental Economics 27*(4), 715–742.

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

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

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica: Journal of the Econometric Society*, 889–917.

Bresnahan, T. F. and P. C. Reiss (1991). Empirical models of discrete games. *Journal of Econometrics 48*(1-2), 57–81.

Brown, Z. Y. and J. Jeon (2024). Endogenous information and simplifying insurance choice. *Econometrica 92*(3), 881–911.

Caplin, A., M. Dean, and J. Leahy (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies 86*(3), 1061–1094.

Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. *Handbook of econometrics 6*, 5549–5632.

Chen, X., T. M. Christensen, and E. Tamer (2018). Monte carlo confidence sets for identified sets. *Econometrica 86*(6), 1965–2018.

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

Cover, T. and J. Thomas (2006). Elements of information theory, 2nd ed. *John Wiley & Sons: Hoboken, NJ, USA*.

Dean, M. and N. Neligh (2023). Experimental tests of rational inattention. *Journal of Political Economy 131*(12), 3415–3461.

Denti, T. (2023). Unrestricted information acquisition. *Theoretical Economics 18*(3), 1101–1140.

Dewan, A. and N. Neligh (2020). Estimating information cost functions in models of rational inattention. *Journal of Economic Theory 187*, 105011.

Fosgerau, M., E. Melo, A. De Palma, and M. Shum (2020). Discrete choice and rational inattention: A general equivalence result. *International economic review 61*(4), 1569–1589.

Gale, D. and H. Nikaido (1965). The jacobian matrix and global univalence of mappings. *Mathematische Annalen 159*(2), 81–93.

Grieco, P. L. (2014). Discrete games with flexible information structures: An application to local grocery markets. *The RAND Journal of Economics 45*(2), 303–340.

Gu, J., T. M. Russell, and T. Stringham (2025). Counterfactual identification and latent space enumeration in discrete outcome models. Working paper.

Joo, J. (2023). Rational inattention as an empirical framework for discrete choice and consumer-welfare evaluation. *Journal of Marketing Research 60*(2), 278–298.

Kline, B. and E. Tamer (2016). Bayesian inference in a class of partially identified models. *Quantitative Economics 7*(2), 329–366.

Kline, B. and E. Tamer (2024). Counterfactual analysis in empirical games. Working paper.

Koh, P. S. (2023). Stable outcomes and information in games: An empirical framework. *Journal of Econometrics 237*(1), 105499.

Liao, M. (2024). Identification of a rational inattention discrete choice model. *Journal of Econometrics 240*(1), 105670.

Magnolfi, L. and C. Roncoroni (2022). Estimation of discrete games with weak assumptions on information. *forthcoming The Review of Economic Studies*.

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review 105*(1), 272–98.

Montes, A. (2022). Canonical equilibria in games with flexible information acquisition. *Available at SSRN 4029000*.

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

Steiner, J., C. Stewart, and F. Matějka (2017). Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica 85*(2), 521–553.

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

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

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies 70*(1), 147–165.

Xie, E. (2024). Nonparametric identification of incomplete information discrete games with non-equilibrium behaviors. *Bank of Canada Staff Working Paper*.

Yang, M. (2015). Coordination with flexible information acquisition. *Journal of Economic Theory 158*, 721–738.

Yang, M. (2020). Optimality of debt under flexible information acquisition. *The Review of Economic Studies 87*(1), 487–536.

A Proofs

A.1 Proof of Lemma 1

Proof. Let $\sigma_i(\tau_i^{\mathbf{X},\mathbf{Z}})$ be player i's optimal action decision rule after receiving a signal $\tau_i^{\mathbf{X},\mathbf{Z}} \in \mathcal{T}_i^{\mathbf{X},\mathbf{Z}}$. Suppose to the contrary that player i plays not only pure strategies but also mixed strategies. Define the following sets

$$\mathcal{T}_{i0} := \{\tau_i^{\mathbf{X},\mathbf{Z}} \in \mathcal{T}_i^{\mathbf{X},\mathbf{Z}} \mid \sigma_i(\tau_i^{\mathbf{X},\mathbf{Z}}) = 0\},$$
$$\mathcal{T}_{i1} := \{\tau_i^{\mathbf{X},\mathbf{Z}} \in \mathcal{T}_i^{\mathbf{X},\mathbf{Z}} \mid \sigma_i(\tau_i^{\mathbf{X},\mathbf{Z}}) = 1\},$$
$$\mathcal{T}_{i,mixed} := \{\tau_i^{\mathbf{X},\mathbf{Z}} \in \mathcal{T}_i^{\mathbf{X},\mathbf{Z}} \mid \sigma_i(\tau_i^{\mathbf{X},\mathbf{Z}}) \in (0,1)\}.$$

The three sets form a partition of player i's signal space $\mathcal{T}_i^{\mathbf{X},\mathbf{Z}}$. Due to the information cost, the player does not have an incentive to investigate signal realizations that induce the same action. For instance, the player does not distinguish $\tau_i' \neq \tau_i$ such that $\tau_i, \tau_i' \in \mathcal{T}_{i1}$ and $\sigma_i(\tau_i') = \sigma_i(\tau_i) = 1$. The effort to process the details of signals incurs information costs but gives no extra benefits. Thus, player i's signal space is equivalent to a set with at most three elements.

Given the optimal information structure $S_i^{\mathbf{X},\mathbf{Z}}$ and the optimal action decision rule $\sigma_i(\cdot)$, the expected payoff under $S_i^{\mathbf{X},\mathbf{Z}}$ excluding the information cost is

$$\int_{\varepsilon_i} \int_{\tau_i^{\mathbf{X},\mathbf{Z}}} U_i(Y_i, Y_{-i}, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \sigma_i(\tau_i^{\mathbf{X},\mathbf{Z}}) \, P_i^{\mathbf{X},\mathbf{Z}}(\tau_i^{\mathbf{X},\mathbf{Z}} \mid \varepsilon_i) \mathrm{d}\tau_i^{\mathbf{X},\mathbf{Z}} \mathrm{d}F_i(\varepsilon_i).$$

Next, consider a new information structure $\tilde{S}_i^{\mathbf{X},\mathbf{Z}} = (\tilde{\mathcal{T}}_i^{\mathbf{X},\mathbf{Z}}, \tilde{P}_i^{\mathbf{X},\mathbf{Z}})$ such that $\tilde{\mathcal{T}}_i^{\mathbf{X},\mathbf{Z}} = \{t_0, t_1\}$. Let the optimal action decision rule $\tilde{\sigma}_i(\cdot)$ associated with the new information structure $\tilde{S}_i^{\mathbf{X},\mathbf{Z}}$ be

$$\tilde{\sigma}_i(t_0) = 0$$

$$\tilde{\sigma}_i(t_1) = 1.$$

The distribution of the signals are defined by

$$\tilde{P}_i^{\mathbf{X},\mathbf{Z}}(t_0 \mid \varepsilon_i) = \int_{\mathcal{T}_{i0} \cup \mathcal{T}_{i,mixed}} P_i^{\mathbf{X},\mathbf{Z}}(\tau_i^{\mathbf{X},\mathbf{Z}} \mid \varepsilon_i) \mathrm{d}\tau_i^{\mathbf{X},\mathbf{Z}}$$
$$\tilde{P}_i^{\mathbf{X},\mathbf{Z}}(t_1 \mid \varepsilon_i) = \int_{\mathcal{T}_{i1}} P_i^{\mathbf{X},\mathbf{Z}}(\tau_i^{\mathbf{X},\mathbf{Z}} \mid \varepsilon_i) \mathrm{d}\tau_i^{\mathbf{X},\mathbf{Z}}$$

for any payoff shock $\varepsilon_i \in \mathcal{E}_i$. The expected payoff under $\tilde{S}_i^{\mathbf{X,Z}}$ excluding the information cost is identical to that under S_i since

$$\int_{\varepsilon_i} U_i(Y_i, Y_{-i}, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \tilde{\sigma}_i(t_0) \, \tilde{P}_i^{\mathbf{X,Z}}(t_0 \mid \varepsilon_i) \mathrm{d}F_i(\varepsilon_i) + \int_{\varepsilon_i} U_i(Y_i, Y_{-i}, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \tilde{\sigma}_i(t_1) \, \tilde{P}_i^{\mathbf{X,Z}}(t_1 \mid \varepsilon_i) \mathrm{d}F_i(\varepsilon_i)$$

$$= \int_{\varepsilon_i} \int_{\mathcal{T}_{i0} \cup \mathcal{T}_{i,mixed}} U_i(Y_i, Y_{-i}, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \sigma_i(\tau_i^{\mathbf{X,Z}}) \, P_i^{\mathbf{X,Z}}(\tau_i^{\mathbf{X,Z}} \mid \varepsilon_i) \mathrm{d}\tau_i^{\mathbf{X,Z}} \mathrm{d}F_i(\varepsilon_i)$$

$$+ \int_{\varepsilon_i} \int_{\mathcal{T}_{i1}} U_i(Y_i, Y_{-i}, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \sigma_i(\tau_i^{\mathbf{X,Z}}) \, P_i^{\mathbf{X,Z}}(\tau_i^{\mathbf{X,Z}} \mid \varepsilon_i) \mathrm{d}\tau_i^{\mathbf{X,Z}} \mathrm{d}F_i(\varepsilon_i)$$

$$= \int_{\varepsilon_i} \int_{\tau_i^{\mathbf{X,Z}}} U_i(Y_i, Y_{-i}, \varepsilon_i; \mathbf{X}, \mathbf{Z}_i) \sigma_i(\tau_i^{\mathbf{X,Z}}) \, P_i^{\mathbf{X,Z}}(\tau_i^{\mathbf{X,Z}} \mid \varepsilon_i) \mathrm{d}\tau_i^{\mathbf{X,Z}} \mathrm{d}F_i(\varepsilon_i)$$

Notice that the information structure $\tilde{S}_i^{\mathbf{X,Z}}$ is less informative than $S_i^{\mathbf{X,Z}}$ if $\mathcal{T}_{i,mixed}$ is nonempty or $|\mathcal{T}_i| > 2$. If $\mathcal{T}_{i,mixed}$ is nonempty or $|\mathcal{T}_i| > 2$, then $\tilde{S}_i^{\mathbf{X,Z}}$ does not require player i to process signal realizations within $\mathcal{T}_{i0}$ and $\mathcal{T}_{i,mixed}$. The mutual information cost function has a property that the more informative it is, the more it costs. By this property, $\tilde{S}_i^{\mathbf{X,Z}}$ is cheaper than $S_i^{\mathbf{X,Z}}$, and the player's net expected payoff is greater under the new information structure $\tilde{S}_i^{\mathbf{X,Z}}$. This contradicts the optimal information structure $S_i^{\mathbf{X,Z}}$. Therefore, the player does not choose mixed strategy and the cardinality of the signal space is at most two. $\qquad\square$

A.2   Proof of Proposition 1.

Let the rival's information strategy be $P_j^{\mathbf{X,Z}}(Y_j|\varepsilon_j)$. Define the deterministic expected payoff (14) by

$$EU_i(\mathbf{X, Z}) = \pi_i(\mathbf{X, Z}_i) + \delta_i(\mathbf{X, Z}_i) P_j^{\mathbf{X,Z}}(Y_j|\varepsilon_j).$$

I construct the Lagrangian for the expected payoff maximization problem:

$$\mathcal{L}(P_i) = \int_{\varepsilon_i} \sum_{Y_i \in \mathcal{Y}_i} (EU_i(\mathbf{X, Z}) + \varepsilon_i) \, P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i) \, \mathrm{d}F(\varepsilon_i)$$

$$- \lambda_i \left[\int_{\varepsilon_i} \left(\sum_{Y_i \in \mathcal{Y}_i} P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i) \log P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i) \right) \mathrm{d}F_i(\varepsilon_i) - \sum_{Y_i \in \mathcal{Y}_i} P_i^{\mathbf{X,Z}}(Y_i) \log P_i^{\mathbf{X,Z}}(Y_i) \right]$$

$$+ \sum_{Y_i \in \mathcal{Y}_i} \int_{\varepsilon_i} \xi(\varepsilon_i) P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i) \, \mathrm{d}F_i(\varepsilon_i) - \int_{\varepsilon_i} \mu(\varepsilon_i) \left(\sum_{Y_i \in \mathcal{Y}_i} P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i) - 1 \right) \mathrm{d}F_i(\varepsilon_i)$$

where

$$P_i^{\mathbf{X,Z}}(Y_i) = \int_{\varepsilon_i} P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i) \, \mathrm{d}F_i(\varepsilon_i),$$

$\xi(\varepsilon_i) \geq 0$ is the Lagrangian multiplier for the constraint:

$$P_i^{\mathbf{X,Z}}(Y_i|\varepsilon_i) \geq 0, \quad \forall \varepsilon_i \in \mathcal{E}_i$$

43

and $\mu(\varepsilon_i)$ is the Lagrangian multiplier for the constraint:

$$\sum_{Y_i} P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) = 1, \quad \forall \varepsilon_i \in \mathcal{E}_i$$

If $\int_{\varepsilon_i} P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) \, dF_i(\varepsilon_i) > 0$, the first-order condition with respect to $P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i)$ is:

$$(EU_i(\mathbf{X},\mathbf{Z}) + \varepsilon_i) - \lambda_i \left[\log P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) + 1 - \log \left(P_i^{\mathbf{X},\mathbf{Z}}(Y_i) \right) - 1 \right] + \xi(\varepsilon_i) - \mu(\varepsilon_i) = 0$$

Since $\int_{\varepsilon_i} P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) \, dF_i(\varepsilon_i) > 0$ implies $P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) > 0$, the complementary-slackness condition gives $\xi(\varepsilon_i) = 0$ and $\mu(\varepsilon_i) > 0$. Thus I have

$$P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) = P_i^{\mathbf{X},\mathbf{Z}}(Y_i) \exp \left(\frac{(EU_i(\mathbf{X},\mathbf{Z}) + \varepsilon_i) - \mu(\varepsilon_i)}{\lambda_i} \right)$$

Plugging the above expression for $P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i)$ into the constraint $\sum_{Y_i} P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) = 1$ gives

$$P_i^{\mathbf{X},\mathbf{Z}}(Y_i|\varepsilon_i) = \frac{P_i^{\mathbf{X},\mathbf{Z}}(Y_i) \exp \left(EU_i(\mathbf{X},\mathbf{Z}) + \varepsilon_i \right)^{1/\lambda_i}}{\sum_{Y_i'} P_i^{\mathbf{X},\mathbf{Z}}(Y_i') \exp \left(EU_i(\mathbf{X},\mathbf{Z}) + \varepsilon_i \right)^{1/\lambda_i}} \tag{27}$$

A.3 Proof of Proposition 2.

Proof. Let $\mathbf{P}(\mathbf{X},\mathbf{Z}) \equiv (P_1(\mathbf{X},\mathbf{Z}_1,\mathbf{Z}_2), P_2(\mathbf{X},\mathbf{Z}_2,\mathbf{Z}_1)) \in [0,1]^2$. I define the fixed point mapping $\Phi : [0,1]^2 \rightarrow [0,1]^2$

$$\Phi(\mathbf{P}(\mathbf{X},\mathbf{Z}),\mathbf{X},\mathbf{Z}) = \begin{pmatrix} \int_{\varepsilon_1} \frac{P_1(\mathbf{X},\mathbf{Z}) \exp(EU_1(\mathbf{X},\mathbf{Z}) + \varepsilon_1)^{1/\lambda_1}}{P_1(\mathbf{X},\mathbf{Z}) \exp(EU_1(\mathbf{X},\mathbf{Z}) + \varepsilon_1)^{1/\lambda_1} + 1 - P_1(\mathbf{X},\mathbf{Z})} dF_1(\varepsilon_1) \\ \int_{\varepsilon_2} \frac{P_2(\mathbf{X},\mathbf{Z}) \exp(EU_2(\mathbf{X},\mathbf{Z}) + \varepsilon_2)^{1/\lambda_2}}{P_2(\mathbf{X},\mathbf{Z}) \exp(EU_2(\mathbf{X},\mathbf{Z}) + \varepsilon_2)^{1/\lambda_2} + 1 - P_2(\mathbf{X},\mathbf{Z})} dF(\varepsilon_2) \end{pmatrix}.$$

Given Assumption 1, Φ satisfies the conditions for Brouwer's Fixed Point Theorem. Therefore, there exists a fixed point $(P_1(\mathbf{X},\mathbf{Z}_1,\mathbf{Z}_2), P_2(\mathbf{X},\mathbf{Z}_2,\mathbf{Z}_1)) \equiv \mathbf{P}(\mathbf{X},\mathbf{Z}) \in (0,1)^2$ such that solves

$$\Phi(\mathbf{P}(\mathbf{X},\mathbf{Z}),\mathbf{X},\mathbf{Z}) = \mathbf{P}(\mathbf{X},\mathbf{Z}).$$

$\square$

A.4 Proof of Proposition 4.

Proof. Fix an arbitrary realization of player i's player-specific variable $\mathbf{Z}_i = \mathbf{z}_i$ and market-specific variable $\mathbf{X} = \mathbf{x}$. By Assumption 3, the rival's player-specific variable $\mathbf{Z}_j$ has continuous variation conditional on $\mathbf{Z}_i = \mathbf{z}_i$. I can find two realizations of $\mathbf{Z}_j$, and say $\mathbf{z}_j^1$ and

$\mathbf{z}_j^2$. The equation (16) for these realizations leads to the following system of equations

$$G_i^{-1}(P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{z}_j^1)) = \pi_i(\mathbf{x}, \mathbf{z}_i) + \delta_i(\mathbf{x}, \mathbf{z}_i) P_j(\mathbf{x}, \mathbf{z}_j^1, \mathbf{z}_i)$$

$$G_i^{-1}(P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{z}_j^2)) = \pi_i(\mathbf{x}, \mathbf{z}_i) + \delta_i(\mathbf{x}, \mathbf{z}_i) P_j(\mathbf{x}, \mathbf{z}_j^2, \mathbf{z}_i)$$

By the Assumption 4, both values of $P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{z}_j^1)$ and $P_i(\mathbf{x}, \mathbf{z}_i, \mathbf{z}_j^2)$ are known to the econometrician. Moreover, by the Assumption 1, the function $G_i^{-1}(\cdot)$ can be derived by the econometrician. The above system of equations contains two equations with two unknowns, $\pi_i(\mathbf{x}, \mathbf{z}_i)$ and $\delta_i(\mathbf{x}, \mathbf{z}_i)$.

Assumptions 3 and 4 imply that $P_j(\mathbf{x}, \mathbf{z}_j^1, \mathbf{z}_i) \neq P_j(\mathbf{x}, \mathbf{z}_j^2, \mathbf{z}_i)$, and the rank condition for the above system of equations is satisfied. Therefore, the unknown functions $\pi_i(\mathbf{x}, \mathbf{z}_i)$ and $\delta_i(\mathbf{x}, \mathbf{z}_i)$ are identified. Since the choice of $\mathbf{z}_i$ and $\mathbf{x}$ was arbitrary, the functions $\pi_i(\mathbf{X}, \mathbf{Z}_i)$ and $\delta_i(\mathbf{X}, \mathbf{Z}_i)$ are identified for all values of $\mathbf{Z}_i$ and $\mathbf{X}$ in its support.

$\square$

A.5 Proof of Proposition 5.

Proof. For simplicity, suppose that $\pi_i(\mathbf{Z}_i) = \mathbf{Z}_i \beta_i$ and $\delta_i(\mathbf{Z}_i) = \delta_i$. The parameter vector of interest is $\theta = (\beta_1, \beta_2, \delta_1, \delta_2)$. Assume that Z_{ik} has an everywhere positive Lebesgue density conditional on $\mathbf{Z}_{i,-k}$ and the matrices $\mathbf{Z}_i$ and $\mathbf{Z}_j$ have full column rank.

Without loss of generality, assume $\beta_{2k} > 0$, and let Z_{2k} be small enough conditional on $\mathbf{Z}_{2,-k}$. Consequently, $\beta_2 Z_{2k} \approx -\infty$, and player 2's expected deterministic payoff of choosing $Y_2 = 1$ is

$$EU_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) = \mathbf{Z}_2 \beta_2 + \delta_2 \cdot P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) \approx -\infty.$$

The conditional choice probability for player 2 is

$$P_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) = \int_{\varepsilon_2} \frac{P_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) \exp\left(EU_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) + \varepsilon_2\right)^{1/\lambda_2}}{P_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) \exp\left((EU_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) + \varepsilon_2)\right)^{1/\lambda_2} + (1 - P_2(\mathbf{Z}_1, \mathbf{Z}_2; \theta))} \, dF_2(\varepsilon_2)$$

$$\approx \int_{\varepsilon_2} \frac{P_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) \cdot 0}{P_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) \cdot 0 + (1 - P_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta))} \, dF_2(\varepsilon_2) = 0$$

as $EU_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) \approx -\infty$. Thus, player 2 will choose the action $Y_2 = 0$ for all values of Y_1 and $\mathbf{Z}_1$.

Given that $P_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) \approx 0$ as $EU_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) \approx -\infty$, player 1's expected deterministic payoff of choosing $Y_1 = 1$ is

$$EU_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) = \mathbf{Z}_1 \beta_1 + \delta_1 \cdot P_2(\mathbf{Z}_2, \mathbf{Z}_1; \theta) = \mathbf{Z}_1 \beta_1,$$

and the conditional choice probability is

$$P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) = \int_{\varepsilon_1} \frac{P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) \exp\left(\mathbf{Z}_1\beta_1 + \varepsilon_1\right)^{1/\lambda_1}}{P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) \exp\left(\mathbf{Z}_1\beta_1 + \varepsilon_1\right)^{1/\lambda_1} + (1 - P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta))} \, dF_1(\varepsilon_1)$$

Next, I claim that $P_1(1 \mid \mathbf{Z}_1, \mathbf{Z}_2; \theta) \neq P_1(1 \mid \mathbf{Z}_1, \mathbf{Z}_2; b)$ for $b \neq \theta$.

Suppose to the contrary $P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) = P_1(\mathbf{Z}_1, \mathbf{Z}_2; b) \in (0, 1)$. The full rank condition on $\mathbf{Z}_1$ guarantees $\mathbf{Z}_1\beta_1 \neq \mathbf{Z}_1 b_1$ for $b_1 \neq \beta_1$ and the exponential function is strictly monotonic. It follows that

$$\exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1\beta_1)\right) \neq \exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1 b_1)\right).$$

Without loss of generality, assume $\mathbf{Z}_1\beta_1 > \mathbf{Z}_1 b_1$ for $b_1 \neq \beta_1$ and $\exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1\beta_1)\right) > \exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1 b_1)\right)$.

Let the function $I(x)$ be defined by

$$I(x) = \frac{p \exp\left(\frac{1}{\lambda}(x + \varepsilon)\right)}{p \exp\left(\frac{1}{\lambda}(x + \varepsilon)\right) + (1 - p)}$$

This function is strictly increasing in x since the derivative with respect to x is always positive:

$$\frac{dI}{dx} = \frac{\frac{1}{\lambda} p(1 - p) \exp\left(\frac{1}{\lambda}(x + \varepsilon)\right)}{[p \exp\left(\frac{1}{\lambda}(x + \varepsilon)\right) + (1 - p)]^2} > 0.$$

Thus I have

$$P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) = \int_{\varepsilon_1} \frac{P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) \exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1\beta_1 + \varepsilon_1)\right)}{P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta) \exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1\beta_1 + \varepsilon_1)\right) + (1 - P_1(\mathbf{Z}_1, \mathbf{Z}_2; \theta))} \, dF_1(\varepsilon_1)$$

$$= \int_{\varepsilon_1} \frac{P_1(\mathbf{Z}_1, \mathbf{Z}_2; b) \exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1\beta_1 + \varepsilon_1)\right)}{P_1(\mathbf{Z}_1, \mathbf{Z}_2; b) \exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1\beta_1 + \varepsilon_1)\right) + (1 - P_1(\mathbf{Z}_1, \mathbf{Z}_2; b))} \, dF_1(\varepsilon_1)$$

$$> \int_{\varepsilon_1} \frac{P_1(\mathbf{Z}_1, \mathbf{Z}_2; b) \exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1 b_1 + \varepsilon_1)\right)}{P_1(\mathbf{Z}_1, \mathbf{Z}_2; b) \exp\left(\frac{1}{\lambda_1}(\mathbf{Z}_1 b_1 + \varepsilon_1)\right) + (1 - P_1(\mathbf{Z}_1, \mathbf{Z}_2; b))} \, dF_1(\varepsilon_1)$$

$$= P_1(\mathbf{Z}_1, \mathbf{Z}_2; b)$$

Since I have assumed $\mathbf{Z}_1\beta_1 > \mathbf{Z}_1 b_1$ for $b_1 \neq \beta_1$, it follows that $P_1(\mathbf{Z}_1, \mathbf{Z}_2; \beta) > P_1(\mathbf{Z}_1, \mathbf{Z}_2; b)$. Therefore, the assumption that $P_1(\mathbf{Z}_1, \mathbf{Z}_2; \beta) = P_1(\mathbf{Z}_1, \mathbf{Z}_2; b)$ leads to a contradiction. Hence, $P_1(\mathbf{Z}_1, \mathbf{Z}_2; \beta) \neq P_1(\mathbf{Z}_1, \mathbf{Z}_2; b)$ for $b \neq \beta$, and this result identifies β_1.

As for the identification of β_2, assume $\beta_{1k} > 0$, and choose Z_{1k} small enough such that $Z_{1k}\beta_{1k} \approx -\infty$. β_2 is identified using the similar argument.

As for the identification of δ_i for $i = 1, 2$, let $Z_{ik} \approx \infty$. δ_1 and δ_2 are identified using the similar argument. $\square$