CIRJE-F-797

Duopoly in the Japanese Airline Market: Bayesian Estimation for the Entry Game

Shinya Sugawara
Graduate School of Economics, University of Tokyo

Yasuhiro Omori
University of Tokyo

April 2011

Duopoly in the Japanese Airline Market: Bayesian Estimation for the Entry Game

SHINYA SUGAWARA * YASUHIRO OMORI †

April 8, 2011

Abstract

This paper provides an econometric analysis on a duopoly game in the Japanese domestic airline market. We establish a novel Bayesian estimation approach for the entry game, which allows the incorporation of flexible inference techniques. We find asymmetric strategic interactions between Japanese firms, which implies that competition is still influenced by the former regulation regime. Furthermore, our prediction analysis indicates that the new Shizuoka airport will suffer from a lack of demand in the future.

Key words: Japanese airline market, Bayesian analysis, Entry game, Markov chain Monte Carlo, Multiple equilibria, Mixed strategy.

1 Introduction

The Japanese domestic airline market today is characterized by a serious rivalry between two giant firms, Japan Airlines (JAL) and All Nippon Airways (ANA). This paper presents an econometric analysis of their duopoly competition in the domestic market. In comparison with the well-studied US airline market, it is difficult to capture the strategic interactions between firms in Japanese market. In the United States, consecutive deregulation policies have introduced a severe substitutive competition. However, the Japanese government had once imposed a selective regulation on the firms: JAL was mainly assigned to operate international flights, while ANA was assigned to operate domestic flights. Although it has been abandoned since the 1980s, this regulation makes it difficult to arrive at a clear consensus on the current pattern of the firms' strategic interaction.

This paper is concerned with the econometric analysis for the duopoly competition. We describe it as a static, complete information entry game: for each air route,

*Graduate School of Economics, University of Tokyo, Tokyo 113-0033, Japan E-mail: ee077007@ecc.u-tokyo.ac.jp

†Faculty of Economics, University of Tokyo, Tokyo 113-0033, Japan. E-mail:omori@e.u-tokyo.ac.jp.

the firms decide whether to operate a flight, taking account of the action of the counterpart. For an empirical study, there is an econometric difficulty in the treatment of multiple Nash equilibria. It induces the multiplicity of underlying data generating processes, which causes an identification problem for the model parameters.

Following the seminal paper of Bresnahan and Reiss (1991), several estimation methods have been proposed to deal with this problem. A major approach is to limit our consideration to a specific equilibrium. For example, Jia (2008), in an analysis of the competition between Wal-Mart and Kmart, focused on an equilibrium in which Kmart plays the leader's role, as their history has shown. Aguirregabiria and Mira (2002), in the repeated game setting, added an assumption that players stick to their first choice of an equilibrium. On the other hand, Tamer (2003) and Ciliberto and Tamer (2009) explicitly modeled the players' choice among the multiple options using latent variables. This part of the model structure is called a selection rule.

For the model with the selection rule approach, this paper proposes a novel Bayesian estimation procedure. The advantage of our methodology is that it allows to employ flexible inference techniques. Specifically, we study two applied issues that are not captured in the previous papers. First, we adopt mixed-strategy Nash equilibria in addition to conventional pure Nash equilibria as a data generating process. Second, we carry out a statistical prediction for a policy experiment on the future of the new Shizuoka Airport.

Our estimation result provides a clear perspective on the complicate situation of the Japanese airline market. It is shown that ANA, which had a monopolistic power during the former regulation regime, receives a negative impact from the presence of JAL, while JAL enjoys a positive benefit from the presence of ANA. This conclusion implies that the former regulation is still influential in this market. Furthermore, the prediction analysis yields a pessimistic forecast that the most routes from the Shizuoka airport will be difficulty surviving.

The organization of this paper is as follows. In Section 2, we provide a brief review of the Japanese airline market. Section 3 describes the economic model and and our Bayesian estimation method for the entry game model. The proposed method is applied to the Japanese airline data in Section 4. Section 5 concludes the paper.

2 The Japanese airline market

This section provides a brief review of the Japanese domestic airline market from the perspectives of an international comparison and a history. There are various statistics to show the considerable scale of the Japanese civil aviation industry[1]: in the year

[1]The country and airport-level statistics come from International Civil Aviation Organization (2006), and the company-level statistics come from International Air Transport Association (2007)

2006, Japan was the world's fifth largest country in terms of passengers and Tokyo (Haneda) was the world's fourth-largest airport in terms of passengers. There are two Japanese airline firms in the world ranking in terms of the passengers times distance: in 2007, Japan Airlines (JAL) was ranked the 12th largest, and All Nippon Airways (ANA) was the 21st largest. Focusing on the domestic market, we have more striking figures for the industry. Although it has only the world's 61st largest surface area[2], Japan has the world's third largest number of passengers on domestic flights, just after the United States and China. In this large industry, only two firms, JAL and ANA, occupy more than 90% of the market share.

Despite the large scale and the unique duopoly property, to the best of our knowledge, our study is the first game estimation made for this market. Comparatively, for the US domestic airline market, there have been several empirical studies using the entry game framework such as Berry (1992) and Ciliberto and Tamer (2009). Based on these studies, it has been a stylized fact that the US airline companies have been experiencing severe substitutive competition since airline deregulation, which began in 1978.

In the same manner as in the United States, competition in the Japanese airline market was once restricted by the government. Private air transportation in Japan had been prohibited for six years after the World War II by the General Headquarters/Supreme Commander for the Allied Powers. Although several private companies entered the market after the prohibition was lifted, the Ministry of Transport issued a notice, so-called "the aviation constitution" in 1972 to stabilize the immature market. Under this notice, there were only three companies that were permitted to run flight operations: JAL, ANA and Japan Air System (JAS, formerly called Toa Domestic Airlines). JAL was a flag carrier assigned to operate international flights and the main domestic routes, ANA was assigned the main domestic and local routes and JAS was assigned the domestic local routes. The airfares and routes were controlled by the government. This system is also called the 45/47 system, where 45 and 47 represent the years 1970 and 1972 of the Showa era in Japan.

Under the 45/47 system, the Japanese airline industry expanded rapidly and met a growing demand for airline deregulation like the US open-sky policy. Japanese deregulation began in 1985 when the 45/47 system was abandoned. In the 1990s, some new companies entered the market. However, they failed to expand their market shares because of a long-lasting recession in Japan. Instead, JAL merged with JAS in 2001, and the Japanese airline market has been dominated by only two firms, ANA and (the new) JAL.

Although their histories have much in common, there are several differences between the Japanese and US airline markets. In the United States, the hub-and-spoke

[2]United Nations Statistical Office (2009)

system provides flights connecting through larger airports to reach their destinations. In Japan, direct flights are mainly used because of the slender shape of the country. On the routes between the largest cities, each airline company has many flights, and there has been fierce competition. However, on the local routes, the former restrictions have still affected the behavior of companies. Therefore, there is no clear consensus over whether their strategic interaction is substitutive, compensative or asymmetric.

3 Econometrics

3.1 Econometric models with pure Nash equilibria

We describe the duopoly competition between ANA and JAL as a static, complete information entry game, in a manner similar to that of Ciliberto and Tamer (2009) who analyzed the US market. In this paper, our model takes the form of a two-player and two-strategy (2×2) game.

Let $y_{im} \in \{0,1\}$ denote the strategy of of i-th player in the m-th market. The indices for players $i = 1$ and 2 represent JAL and ANA, and those for the markets $m = 1, 2, \ldots, M$ are routes between two airports, respectively. The strategy $y_{im} = 1$ implies the entrance or operation of a flight in the m-th route by the i-th player while the strategy $y_{im} = 0$ implies no entrance. An important assumption is that competitions across the markets are independent, which implies that the firms do not make a network-level decision.

Our main target of estimation is the payoff function of the players. We set the payoff when a firm does not enter the market to be zero as a reference[3]. When a firm enters a market, its payoff is assumed to be a linear function of observed regressors $\boldsymbol{x}_{im} = (x_{1im}, \ldots, x_{Kim})'$ with $K \times 1$ coefficient parameter $\boldsymbol{\beta}_i$ and an unobserved payoff component u_{im}. Furthermore, we introduce a strategic interaction term denoted by Δ_i, which appears only when the counterpart enters the market. The resulting payoff matrix is given in Table 1.

		Player 2: 1	Player 2: 0
Player 1	1	$\boldsymbol{x}'_{1m}\boldsymbol{\beta}_1 + \Delta_1 + u_{1m}, \boldsymbol{x}'_{2m}\boldsymbol{\beta}_2 + \Delta_2 + u_{2m}$	$\boldsymbol{x}'_{1m}\boldsymbol{\beta}_1 + u_{1m}$, 0
	0	0, $\boldsymbol{x}'_{2m}\boldsymbol{\beta}_2 + u_{2m}$	$0, 0$

Table 1: Payoff matrix for an entry game

For the information structure of players, we assume complete information, which implies that $(\boldsymbol{x}_{im}, \boldsymbol{\beta}_i, \Delta_i, u_{im})$, $i = 1$ and $i = 2$ are all known to both players in the m-th market. Among the common knowledge of players, only $\boldsymbol{x}$ is observed by econometricians. Throughout this paper, $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \Delta_1, \Delta_2)$ is treated as the

[3]See Kooreman (1994) for the estimation without such a standardization.

model parameters and u_{im} as a random error term. For simplicity, we assume that $\boldsymbol{u}_m = (u_{1m}, u_{2m})$ is independent and identically distributed across the market and the distribution is known by econometricians.

A conventional assumption for the data generating process is that the players take pure Nash equilibria. In static games, different values of the payoff yield a variety of pure Nash equilibria, each of which represents a data generating system in the estimation. Under our functional assumption, we have distinct patterns for the equilibria according to the signs of (Δ_1, Δ_2). We call $(\Delta_1 < 0,\ \Delta_2 < 0)$ and $(\Delta_1 < 0,\ \Delta_2 > 0)$ as Models A and B, respectively. We describe the estimation procedures only for Models A and B in this paper, since $(\Delta_1 > 0,\ \Delta_2 > 0)$ and $(\Delta_1 > 0,\ \Delta_2 < 0)$ induce the similar estimation procedures to Models A and B.

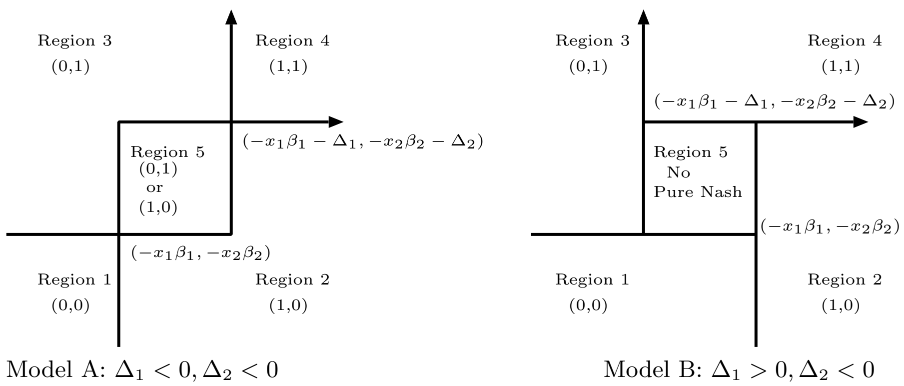

Figure 1: Pure Nash Equilibria

Figure 1 illustrates pure Nash equilibria of the entry game on the coordinates of the unobserved components (u_1, u_2). In both models, each of Regions 1 to 4 has a unique pure Nash equilibrium which is derived by the iterated elimination of strictly dominated strategies. However, Region 5 does not have a unique pure Nash equilibria; in Model A, there are two pure Nash equilibria $(y_1, y_2) = (1, 0)$ and $(0, 1)$, while in Model B, there is no pure Nash equilibrium. This non-uniqueness of Nash equilibria in Region 5 make it difficult to obtain the well-defined choice probabilities for strategy profiles.

To overcome the problem, Tamer (2003) and Ciliberto and Tamer (2009) introduced the additional model structure for the players' choice in Region 5, which is called a selection rule. Specifically in this context, they assumed that Region 5 is divided into four strategy profiles in the proportion of a vector parameter $\boldsymbol{p}_m =$

$(p_{1m}, p_{2m}, p_{3m}, p_{4m})$ such that

$$Pr(z_m = j|\boldsymbol{x}_m, \boldsymbol{p}_m) = P_j(\boldsymbol{\theta}, \boldsymbol{x}_m) + p_{jm}P_5(\boldsymbol{\theta}, \boldsymbol{x}_m), \tag{3.1}$$

$$\sum_{j=1}^{4} p_{jm} = 1, \tag{3.2}$$

where $z_m = j$ indexes the strategy profiles such that $\boldsymbol{y}_m = (y_{1m}, y_{2m}) = (0,0), (1,0), (0,1)$ and $(1,1)$ when $j = 1, 2, 3$ and 4, respectively[4], and $P_j(\boldsymbol{\theta}, \boldsymbol{x}_m)$ is the probability that $\boldsymbol{u}_m$ falls in Region j.

3.2 Bayesian estimation

3.2.1 Estimation for models with pure Nash equilibria

In this subsection, we construct a novel Bayesian estimation procedure for the entry games using hierarchical modeling. In our methodology, the sample-specific selection rule parameters is estimated using information of only one sample. It works because Bayesian estimators are well-defined even under finite samples, unlike consistent estimators. Then although the small, or actually only one, sample size might produces ambiguity of estimators as flat posterior distributions, we no longer need to employ a peculiar estimation technique like the set estimation under the Bayesian scheme.

For technical convenience, we make slight changes in the method of formulating the selection rule. First, in Model A, we can reduce the dimension of $\boldsymbol{p}_m$ to two without a loss of generality, as only two strategy profiles, $z_m = 2$ and 3, are multiple Nash equilibria in Region 5. We let p_m denote a proportion for $z_m = 2$ and $1 - p_m$ for $z_m = 3$. Second, to construct a comprehensive estimation procedure for p_m, we introduce an additional structure of the selection rule. It is established using a latent dummy variable λ_m as follows: for Model A, we assume $\lambda_m \sim Bernoulli(p_m)$, while for Model B, we assume $\boldsymbol{\lambda}_m = (\lambda_{1m}, \lambda_{2m}, \lambda_{3m}, \lambda_{4m}) \sim MN(1, \boldsymbol{p}_m)$, where MN stands for the multinomial distribution. Because of the hierarchical nature of the setting, the marginal posterior distribution for the parameter $\boldsymbol{\theta}$ remains the same as the one without λ. We call p_m and λ_m a selection proportion and a selection dummy, respectively. We assume the beta distribution as the prior distribution for p_m. To express the lack of an economic theory regarding the players' choice among multiple equilibria, we use a uniform distribution on a unit interval, which is a special case of

[4] The general definition of the index system for strategy profiles $z_m = j$, which is also applicable to games with more than 2 players, is as follows: let n_j be a binary number system representation of $j-1$ with the digit number N. Then let $y_{im} = 1$ if the i-th digit of n_j is unity, while $y_{im} = 0$ otherwise. For N-player games, the sequence of indices $j = 1, 2, ..., 2^N$ exactly covers all of the possible strategy profiles without duplication. For example, in 2-player games, $j = 1, 2, 3$ and 4 correspond to $n_j = 00, 10, 01$ and 11; therefore, $\boldsymbol{y}_m = (0,0), (1,0), (0,1)$ and $(1,1)$, respectively. For notational convenience, we have also numbered the unique Nash regions in Figure 1 in the same manner.

the beta distribution.

The estimation for the posterior distributions can be implemented by the standard Markov chain Monte Carlo (MCMC) algorithm: We initialize $\boldsymbol{\theta}$ and then repeat the following algorithm in three blocks.

1. Generate $\boldsymbol{p}|\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{z}$.
2. Generate $\boldsymbol{\lambda}|\boldsymbol{\theta}, \boldsymbol{p}, \boldsymbol{z}$.
3. Generate $\boldsymbol{\theta}|\boldsymbol{p}, \boldsymbol{\lambda}, \boldsymbol{z}$.

where $\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2, ..., \boldsymbol{p}_M)$, $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, ..., \boldsymbol{\lambda}_M)$ and $\boldsymbol{z} = (z_1, z_2, ..., z_M)$. As shown in Appendix A, we can implement the Gibbs sampler for the sampling of the latent variables, $\boldsymbol{p}$ and $\boldsymbol{\lambda}$ and we can use the Metropolis–Hastings algorithm for the model parameter $\boldsymbol{\theta}$.

Since our estimation scheme is a standard MCMC, we can conduct standard model selection techniques to decide a suitable sign condition for (Δ_1, Δ_2). In this paper, we carry out a comparison of the marginal likelihood using the procedure of Chib and Jeliazkov (2001).

3.2.2 Extension to models with mixed strategy Nash equilibrium

The availability of a standard model selection technique also encourages us to analyze various data generating processes that have not been analyzed in previous studies. In this subsection, we consider the mixed-strategy Nash equilibrium in addition to the pure Nash equilibrium. For convenience, we call the models with only the pure Nash equilibrium, which we considered in the previous section, as the pure Nash model and we call the following model which has both pure and mixed-strategy Nash equilibria, the mixed Nash model.

A mixed strategy is defined as a probability distribution over pure strategies. The payoff for a mixed strategy is generally assumed to be the expected value of the corresponding pure strategy payoffs, and the mixed-strategy Nash equilibrium is defined as a mixed-strategy profile from which no player deviates alone. The players must be indifferent between all pure strategies on which they put a positive probability in the Nash mixed strategy; otherwise, they would be better off deviating to a pure strategy.

In both Models A and B, in Regions 1 through 4, the pure Nash equilibrium strategy yields greater payoffs for both players than those yielded by any mixed strategies, as it is obtained by the iterated elimination of strictly dominated strategies. Then only Region 5 might have a mixed-strategy Nash equilibrium. Suppose a probability $\sigma^*_{im}(y_{im})$, where $\sigma^*_{im}(0) + \sigma^*_{im}(1) = 1$, comprises a Nash equilibrium mixed strategy for the player i in the market m on the pure strategy y_{im}. As mentioned above, we required that pure strategies $y_{jm} = 0$ and $y_{jm} = 1$ yield equal expected payoffs for the player j, given $\sigma^*_{im}(y_{im})$:

$$0 = \sigma^*_{im}(0)(\boldsymbol{x}'_{jm}\boldsymbol{\beta}_j + u_{jm}) + \sigma^*_{im}(1)(\boldsymbol{x}'_{jm}\boldsymbol{\beta}_j + \Delta_j + u_{jm}), \tag{3.3}$$

where the left- and right-hand sides represent the expected payoffs for pure strategies $y_{jm} = 0$ and 1, respectively. We thus obtain,

$$\sigma^*_{im}(1) \;=\; -\frac{\boldsymbol{x}'_{jm}\boldsymbol{\beta}_j + u_{jm}}{\Delta_j}, \quad \sigma^*_{im}(0) = 1 + \frac{\boldsymbol{x}'_{jm}\boldsymbol{\beta}_j + u_{jm}}{\Delta_j}. \tag{3.4}$$

In the whole area of Region 5, $(\sigma^*_{1m}(y_{1m}), \sigma^*_{2m}(y_{2m})) \in [0,1]^2$ holds for $\{y_{1m}, y_{2m}\} \in \{0,1\}^2$ and $\sigma^*_{im}(0) + \sigma^*_{im}(1) = 1$. Therefore, $(\sigma^*_{1m}(y_{1m}), \sigma^*_{2m}(y_{2m}))$ is a unique mixed-strategy Nash equilibrium in Region 5 for both Models A and B. Let $\rho_j(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m)$ for $j = 1, 2, 3$ and 4 be the joint probability for each strategy profile $(y_{1m}, y_{2m}) = (0,0), (1,0), (0,1)$ and $(1,1)$ under the mixed-strategy Nash equilibrium. We have

$$\rho_1(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m) = \sigma^*_{1m}(0)\sigma^*_{2m}(0), \qquad \rho_2(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m) = \sigma^*_{1m}(1)\sigma^*_{2m}(0), \tag{3.5}$$

$$\rho_3(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m) = \sigma^*_{1m}(0)\sigma^*_{2m}(1), \qquad \rho_4(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m) = \sigma^*_{1m}(1)\sigma^*_{2m}(1). \tag{3.6}$$

We discuss the remaining estimation procedure separately for the sign conditions for (Δ_1, Δ_2). First, we consider the Model A. In Figure 1, Region 5 has two pure Nash equilibria for Model A. To incorporate the mixed-strategy Nash equilibrium into the data generating process, we model the selection rule as the following two-step process. In the first step, we select pure or mixed Nash equilibrium. Using a selection dummy $\lambda^{Nash}_m \sim Bernoulli(p^{Nash}_m)$, we choose a pure Nash equilibrium if $\lambda^{Nash}_m = 1$; otherwise, we choose a mixed strategy. In the second step, we pick a specific strategy profile as follows: if $\lambda^{Nash}_m = 1$, one of the pure Nash equilibria, $z_m = 2$ or 3, is chosen as in the pure Nash model in Section 3.1. To avoid a notational confusion, we replace p_m and λ_m for the pure Nash models with p^{pure}_m and λ^{pure}_m. If $\lambda^{Nash}_m = 0$, one of the strategy profiles is chosen with probability $\boldsymbol{\rho}(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m) = \{\rho_1(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m), \rho_2(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m), \rho_3(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m), \rho_4(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m)\}$, corresponding to the mixed-strategy Nash equilibrium. In the posterior sampling, we integrate out $\boldsymbol{\rho}(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m)$ with respect to $\boldsymbol{u}_m$. The MCMC procedure is presented in the Appendix B.2

Next, we consider Model B. As shown in Figure 1, there is no pure Nash equilibrium in Region 5 for Model B. Because only the mixed strategy induces the feasible Nash equilibrium, we use $\rho_j(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m)$ as a Nash selection proportion p_{jm} in the choice probabilities (3.1). Then, the selection proportion is an explicit function of model parameters as in (3.5)-(3.6) instead of being a non-identified nuisance parameter. The MCMC procedure is presented in the Appendix B.1

In the classical estimation, although parameters in the mixed Nash model for Model B are point-identified, we cannot compare the pure and mixed Nash models because it is difficult to carry out the model selection between set-identified and point-identified models. However, our Bayesian approach can apply the model selection, as both models can be estimated by the standard MCMC method. This is an advantage of our Bayesian methodology.

3.3 Prediction analysis

In the analysis of the entry game, the prediction of choice probabilities often provides rich implications for empirical studies. Then we establish a prediction technique as another example to show flexibility of our Bayesian approach. The predicted posterior distribution of strategy profiles in the new market can be easily obtained in the MCMC scheme, as shown in Appendix C.

4 Empirical study of duopoly in Japanese airline market

4.1 Data

We consider Japanese airline competition as an entry game where the two players are ANA and JAL, including their affiliated companies. We define a market as a route between two airports. If a firm has at least one flight between the airports [5], we let $y_{im} = 1$. The data are constructed from the timetable (Japan Railway Company, 2007) for the period from February 1 to March 31, 2007.

When choosing samples for our empirical study, we eliminated several airports with very few flights. These are airports that generally suffer from a lack of demand for transportation and survive because of local or central government subsidies. It is true that this subsidizing policy helps people living in areas such as isolated islands where they have difficulty getting the transportation they need. However, because there is a wide range of obstacles, including the financing deficit for free parking services, it is difficult to evaluate the exact amount of the subsidy. Therefore, for our analysis of market behavior, we simply remove these small airports with one or two routes and also remove airports located on isolated islands except Naha, which is the major airport serving the southern islands scattered around the Okinawa prefecture. Overall, we have 39 airports and 741 markets for the combined two airlines.

For the explanatory variables, we first consider two factors used in Berry (1992) and Ciliberto and Tamer (2009): (1) the direct distance (**Distance**, measured in thousands of kilometers) between the airports by which we measure travel cost, and

[5]The code-sharing flights with the other carrier are also included.

(2) the product of the city populations (**Population**, measured in tens of trillions), which we use as proxy variables for market size or demand. The direct distances are calculated by the geodesic distance formula Banerjee et al. (2004, pp.17–18) [6].

For the city population, we use the population of the prefecture based on the census data collected in the year 2005[7]. The squared population (**Sq Population**) is also adopted as an explanatory variable to capture the nonlinear effect of the market sizes.

To compose a firm-market specific explanatory variable, we include the sums of the number of flights from the two airports for each firm (**Flight ANA** and **Flight JAL** for ANA and JAL, respectively, measured in groups of 10 flights). We adopt these variables to measure the marginal cost of operations. They are expected to have positive coefficients because the marginal cost would decrease as the companies have more facilities such as ticket offices.

Because the railway service has a larger share of domestic transportation than the airlines, we also construct a dummy variable for the bullet train called "Shinkansen" (**Train**). It takes the value of one if the flight can be replaced by the bullet train without a transfer and the travel time between the airport and the railway station is less than one and a half hours[8].

Table 2 shows the summary statistics for the independent variables. Furthermore, we present the entrance status for the different market sizes separated by quartiles of population in Table 3. A similar table for American airlines is provided by Ciliberto and Tamer (2009), where the number of entrants is not monotonically related to the market size. In Japanese airline markets, we can see that the large markets are especially competitive. However, an interesting finding is that the number of markets where only ANA is present is not increasing proportionally to the market size. As we shall show later, this would imply the efficiency of the strategy of ANA.

---

[6]Let θ_1, θ_2 be the latitudes in radian of the two points, and λ_1, λ_2 be the radian longitudes. The geodesic distance formula says that the distance of the two points is

$$R \arccos[\sin\theta_1 \sin\theta_2 + \cos\theta_1 \cos\theta_2 \cos(\lambda_2 - \lambda_1)], \tag{4.1}$$

where R denotes the radius of the earth. We adopt $R = 6371$ km as in Banerjee et al. (2004, p.19). In this formula, the earth is assumed to be spherical. The actual shape of the earth is ellipsoidal; however, for our purposes, approximation in the spherical model has enough accuracy. We used the coordinates data taken from 2005 Annual Statistics of Civil Aviation (Ministry of Land, Infrastructure and Transport, 2005).

[7]Because the area of the Hokkaido prefecture is much larger than other prefectures, we use sub-prefecture populations for those local airports in Hokkaido.

[8]We also considered other explanatory variables such as the number of business offices and the amount of industrial production, which are expected to describe the condition of the local economy. However, these variables were not effective in explaining market entry in the sense that the 95% credible intervals for their parameters include 0. Therefore, we removed them from our payoff function; however, the estimates of other parameters are basically unchanged.

Variable	Mean	Stdev
Distance	0.742	0.487
Population	0.824	1.395
Sq Population	2.625	10.160
Train	0.108	0.310
Flight ANA	3.887	4.236
Flight JAL	3.969	4.126

Table 2: Summary Statistics

	Market size				
	Large	Medium large	Medium small	Small	Total
No entrant	100	142	163	174	579
1 entrant	36	36	17	10	99
(ANA)*	(13)	(17)	(7)	(7)	(44)
(JAL)*	(23)	(19)	(10)	(3)	(55)
2 entrants	50	7	5	1	63
Total	186	185	185	185	741
(ANA)**	(63)	(24)	(12)	(8)	(107)
(JAL)**	(73)	(26)	(15)	(4)	(118)

Table 3: Entrance status by market size

*: The number of markets of each firm for one entrant market.

**: The total number of markets entered for each firm.

4.2 Estimation results

4.2.1 Model Selection

The marginal likelihoods for the model selection are shown in Table 4 (in a logarithmic scale) [9].

Based on the marginal likelihoods, we select a pure Nash model with $\Delta_{ANA} < 0$ and $\Delta_{JAL} > 0$ as the best model for the Japanese domestic markets. It implies that the entry of JAL reduces the payoff of ANA, while the entry of ANA increases that of JAL. This result supports the common belief about Japanese airline companies: because ANA was allowed to operate in both the main and local routes under the 45/47 system, it was able to accumulate more knowledge of the domestic market than JAL and JAS, who could contribute only to main and local routes, respectively. This superiority is of great advantage to ANA by allowing the construction of more

[9] When we use a very flat prior with the prior variance equal to 1000, the sample paths are often found to be unstable, which could be a result of the flat likelihood because the paths seem to be more stable as we increase the number of independent variables.

beneficial networks after the deregulation. To overcome its disadvantage, JAL has been challenging the routes dominated by ANA. Therefore, the routes where ANA has already been operating are favored by JAL (i.e., $\Delta_{JAL} > 0$), while the entry of JAL reduces ANA's benefit in profitable routes ($\Delta_{ANA} < 0$).

	Mixed Nash				Pure Nash			
Model	$\Delta_1 < 0$	$\Delta_1 > 0$	$\Delta_1 < 0$	$\Delta_1 > 0$	$\Delta_1 < 0$	$\Delta_1 > 0$	$\Delta_1 < 0$	$\Delta_1 > 0$
	$\Delta_2 < 0$	$\Delta_2 > 0$	$\Delta_2 > 0$	$\Delta_2 < 0$	$\Delta_2 < 0$	$\Delta_2 > 0$	$\Delta_2 > 0$	$\Delta_2 < 0$
Likelihood	-530.29	-533.36	-416.02	-418.11	-432.41	-415.65	-410.68	-428.65
Prior	-28.68	-28.75	-28.59	-28.60	-28.68	-28.60	-31.94	-32.31
Posterior	29.19	20.75	24.25	25.05	29.19	24.88	12.16	5.47
ML	-588.16	-582.86	-468.96	-471.76	-490.28	-469.13	-454.78	-466.43
(S.E.)	(0.06)	(1.00)	(0.15)	(0.10)	(0.05)	(0.10)	(0.24)	(0.36)

Table 4: Marginal likelihoods for real data (in logarithm)

This asymmetric property of duopoly competition must be emphasized as a clear difference from that of the US market, where the airlines are playing a severe substitution competition. Our result may suggest that the Japanese version of "open-sky" deregulation is still incomplete.

4.2.2 Estimation for model parameters

Table 5 shows the estimation result for the pure Nash model with ($\Delta_{JAL} < 0, \Delta_{ANA} > 0$). Before analyzing details of the estimators, we check the performance of our MCMC procedure. The last two columns in Table 5 report statistics for this purpose, inefficiency factors (IF) and p-values of the convergence diagnostics for the MCMC (CD). The 20,000 MCMC samples were generated after discarding 10,000 initial samples as the burn-in period. The inefficiency factors are 13 to 52, which implies that we would obtain the same variance of the posterior sample means from 400 uncorrelated draws, even in the worst case. In Appendix D, we also present figures for and paths of MCMC samples, which clearly show that the chain seems to mix well for all parameters. All p-values of the convergence diagnostics are greater than 0.05 and there is no evidence against convergence. The acceptance rates of the Metropolis–Hastings algorithm are high enough (0.870, 0.997, 0.645 and 0.768 for $\boldsymbol{\beta}_1$, Δ_1, $\boldsymbol{\beta}_2$ and Δ_2) and the proposal distribution seems to approximate the conditional posterior distribution well. These results indicate that our MCMC procedure works well.

Param.	Mean	Stdev	95% Interval	IF	CD
ANA					
Constant	-2.386**	0.284	(-2.962,-1.850)	24.8	0.38
Distance	-0.516*	0.222	(-0.955,-0.096)	5.7	0.79
Population	0.376**	0.216	(-0.048, 0.793)	8.9	0.92
Sq Population	-0.082**	0.025	(-0.128,-0.030)	7.6	0.72
Train	-1.230**	0.438	(-2.109,-0.383)	2.4	0.55
Flight	0.531**	0.072	(0.400, 0.680)	47.1	0.29
Δ	-6.745**	1.052	(-8.912,-4.820)	46.6	0.17
JAL					
Constant	-2.045**	0.579	(-3.506,-1.379)	51.2	0.86
Distance	-0.660	0.489	(-1.605, 0.308)	18.7	0.54
Population	0.369	0.292	(-0.205, 0.915)	14.7	0.76
Sq Population	-0.050	0.033	(-0.113, 0.011)	13.6	0.76
Train	-0.718	0.717	(-2.420, 0.627)	30.8	0.67
Flight JAL	0.035	0.067	(-0.122, 0.141)	22.5	0.82
Δ	6.534**	2.045	(3.154,10.833)	46.3	0.63

Table 5: Estimation results (Pure Nash, $\Delta_{ANA} < 0, \Delta_{JAL} > 0$)
*(**): The 95% (99%) credible interval does not include zero.

Next, we analyze estimation results for the model parameters. The first four columns of Table 5 report the posterior means, posterior standard deviations and 95% credible intervals, along with the figure of the posterior densities in Appendix D. In the payoff function for ANA, all independent variables are effective in the sense that the 95% credible intervals for the corresponding parameters do not include zero. The coefficient of Distance is estimated to be negative, as expected, implying that the large distance (or high cost) would decrease the payoff. It is somewhat surprising that the estimates for Population and Sq Population are found to be positive and negative, respectively. We suppose that the payoff would initially increase as the population increases; however, it would decrease when the demand exceeds some certain size. The negative sign of Train indicates that the existence of substitutive railway transportation decreases the benefit of the airline company. Flight ANA, which is the sum of the number of flights from the two airports, has a positive sign. As the marginal cost of light operations decreases, the payoff will increase, as expected.

However, in the payoff function for JAL, it seems that the independent variables are ineffective, as opposed to those of ANA, except for the constant term and a dummy variable for the entry of ANA, although the estimates are similar to those obtained for the payoff function of ANA. It may imply that the entry decision of JAL is to follow the market leader, ANA.

The above result is similar to the study of Jia (2008) on US retail chains, where Kmart is the leader and Wal-Mart is the follower. It would be interesting to study the competition as a sequential game where ANA plays the role of the leader and JAL is the follower. However, this topic is beyond the scope of this paper and will remain for future work.

4.3 Prediction analysis for a new airport

Next, we consider the prediction of choice probabilities for the airline companies regarding a new airport (Shizuoka airport), which was established recently in 2009. ANA and JAL decided to have their flights depart from the new airport toward two airports (Shin-Chitose and Naha for ANA, and Shin-Chitose and Fukuoka for JAL). The airport authority is requesting more routes for the airlines and four more airports, Narita, Komatsu, Matsuyama and Kagoshima, are listed as candidates on the website [10]. However, there has been a lot of discussion about whether there was sufficient demand to build this new airport. This is because Shizuoka city is located between Tokyo and Nagoya (as shown in Figure 2) where they have large airports that are accessible from Shizuoka within one hour using the Shinkansen bullet train.

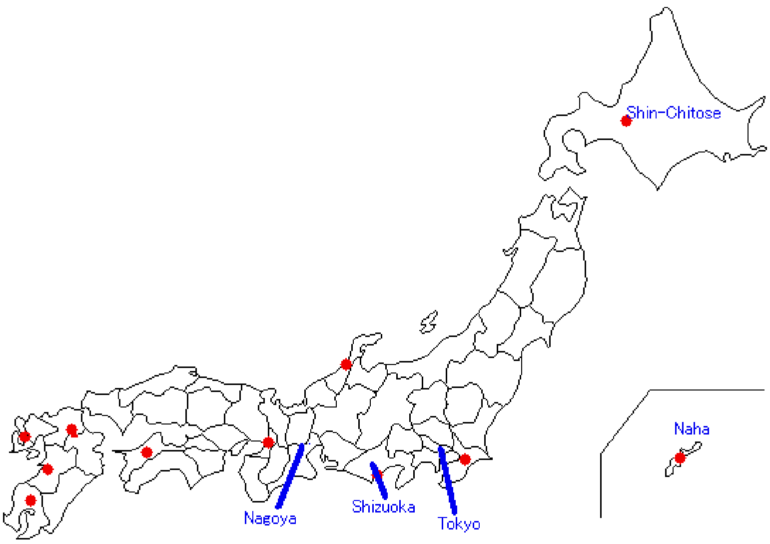

Figure 2: Map of Airports

To evaluate the need for the new airport, we predict the choice probability of no entry for ANA and JAL, *i.e.* $z_{new} = 1$ $(y_{ANA,new} = 0, y_{JAL,new} = 0)$, using the pure Nash Model with $\Delta_{ANA} < 0$ and $\Delta_{JAL} > 0$ selected in the previous subsection. As discussed in the previous section, we have a region without a pure Nash equilibrium:

$$Pr(z_{new} = 1|\boldsymbol{\theta}, \boldsymbol{x}_{new}, \boldsymbol{\lambda}_{new}) = P_1(\boldsymbol{\theta}, \boldsymbol{x}_{new}) + \lambda_{1,new} P_5(\boldsymbol{\theta}, \boldsymbol{x}_{new}), \tag{4.2}$$

[10] http://www.pref.shizuoka.jp/kuukou/contents/gaiyo/yotei.html

where $\boldsymbol{\lambda}_{new} = (\lambda_{1,new}, \ldots, \lambda_{4,new}) \sim MN(1, \boldsymbol{p}_{new})$, $\boldsymbol{p}_{new} \sim \mathcal{D}(\mathbf{1})$. Therefore, regarding the choice probability for the new airport, we focus on three probabilities: (1) the predicted probability in (C.6) or the expected value with respect to $\boldsymbol{\lambda}_{new}$ $(P_1(\boldsymbol{\theta}, \boldsymbol{x}_{new}) + P_5(\boldsymbol{\theta}, \boldsymbol{x}_{new})/4)$, (2) the lower bound $(P_1(\boldsymbol{\theta}, \boldsymbol{x}_{new}))$, and (3) the upper bound $(P_1(\boldsymbol{\theta}, \boldsymbol{x}_{new}) + P_5(\boldsymbol{\theta}, \boldsymbol{x}_{new}))$.

	(1) Expected		(2) Lower		(3) Upper	
Airport	Mean	95% Interval	Mean	95% Interval	Mean	95% Interval
Shin-Chitose	0.459	(0.364, 0.556)	0.290	(0.161, 0.420)	0.968	(0.902, 1.000)
Komatsu	0.944	(0.866, 0.987)	0.933	(0.836, 0.984)	0.980	(0.927, 1.000)
Matsuyama	0.950	(0.918, 0.974)	0.942	(0.905, 0.969)	0.976	(0.944, 0.999)
Fukuoka	0.961	(0.939, 0.979)	0.952	(0.924, 0.974)	0.989	(0.973, 1.000)
Kagoshima	0.951	(0.926, 0.971)	0.939	(0.907, 0.964)	0.986	(0.969, 1.000)
Naha	0.934	(0.906, 0.959)	0.917	(0.881, 0.949)	0.985	(0.962, 1.000)
Narita	0.837	(0.761, 0.901)	0.786	(0.684, 0.871)	0.989	(0.950, 1.000)

Table 6: Posterior means and 95% credible intervals for three probabilities of no entry $((y_{ANA,new}, y_{JAL,new}) = (0,0))$ in the routes from the new airport.

Table 6 shows the posterior means and 95% credible intervals for three (expected, lower and upper) probabilities of no entry ($z_{new} = 1$) in the routes from the new Shizuoka airport to seven airports where the airlines are planning to operate or the airport authority is requesting the operation.

Among the seven airports, the three predicted probabilities for Shin-Chitose Airport are much smaller than those for the other airports. The probabilities for Narita Airport are relatively small but exceed 0.7. This result implies that Shin-Chitose is the only promising airport for airline companies seeking routes from the new airport. While ANA or JAL might be interested also in the routes to Narita, other airports may not be very attractive [11].

Because the Nash equilibrium is a long-run concept to which the economy will converge, these prediction results indicate that there is a high probability of withdrawals from these routes for both airline companies in the future.

5 Conclusion

This paper has analyzed the duopoly competition in the Japanese domestic airline market. Our estimation result is consistent with a popular perception that ANA plays the leader's role in the domestic airline market, while JAL is a follower. Further, our prediction analysis has indicated that the new Shizuoka airport will suffer from a lack of demand in the future.

[11] We also computed these probabilities for Tokyo (Haneda) airport, and found them to be unexpectedly low. However, Tokyo is very close to Shizuoka City and does not seem suitable, as suggested by the fact that there is no flight between Tokyo and Nagoya. This prediction failure is probably because there are not many airports whose locations are similar to that of Shizuoka in our dataset.

Methodologically, we have proposed an alternative estimation method for the entry game models using Bayesian approach. Because of the flexibility in inferences, our methodology might attract empirical analysis for 2×2 games which are commonly seen in the real world.

Acknowledgment

The authors thank Amit Gandhi, Hidehiko Ichimura and Yasuyuki Sawada for helpful discussions and comments. This work is partially supported by a Grant-in-Aid for Scientific Research (No. 18330039) from the Japanese Ministry of Education, Science, Sports, Culture and Technology. The computational results are obtained by using Ox version 5.10(See Doornik (2007)).

References

V. Aguirregabiria and P. Mira. Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. *Econometrica*, 70(4):1519–1543, 2002.

S. Banerjee, B. P. Carlin, and A. E. Gelfand. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall / CRC Press, Bocaraton, Florida, 2004.

S. Berry. Estimation of a model of entry in the airline industry. *Econometrica*, 60(4): 889–917, 1992.

T.F. Bresnahan and P.C. Reiss. Empirical models of discrete games. *Journal of Econometrics*, 48:57–81, 1991.

S. Chib and L. Jeliazkov. Marginal likelihood from the Metropolis-Hastings algorithm. *Journal of American Statistical Association*, 96:270–281, 2001.

Federico Ciliberto and Ellie Tamer. Market strucuture and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828, 2009.

J. A. Doornik. *Object-Oriented Matrix Programming Using Ox*. Timberlake Consultants Press and Oxford, London, 3rd edition, 2007.

International Air Transport Association. *World Air Transport Statistics*. International Air Transport Association, 2007.

International Civil Aviation Organization. *Annual Report of the Counsil 2006*. International Civil Aviation Organization, Montreal, 2006.

Japan Railway Company. *JR Time Table, February*. Kotsushimbunsha, 2007. (in Japanese).

P Jia. What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, 76(6):1263–1316, 2008.

P. Kooreman. Estimation of econometric models of some discrete games. *Journal of Applied Econometrics*, 9(3):255–268, 1994.

Ministry of Land, Infrastructure and Transport. *Annual Statistics of Civil Aviation*. All Japan Air Transport and Service Association, 2005. (in Japanese).

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

United Nations Statistical Office. *Demographic Yearbook*. United Nation, Lake Success, NY, 2009.

A Bayesian estimation for pure Nash models

A.1 Prior distributions

We first describe the prior distributions of the parameters. For the parameter $\boldsymbol{\theta}$, we assume a normal distribution with mean $\boldsymbol{\theta}_0$ and covariance matrix Σ_0 truncated on the region R:

$$\boldsymbol{\theta} \sim \mathcal{TN}_R(\boldsymbol{\theta}_0, \Sigma_0).$$

For example, we take the region $R = (-\infty, 0) \times (-\infty, 0) \times (-\infty, \infty)^K$ for the case where $\Delta_1 < 0$ and $\Delta_2 < 0$. For p_m, we assume a beta distribution with parameters (a_{1m}, a_{2m}):

$$p_m \sim \mathcal{B}(a_{1m}, a_{2m}),$$

for the case $\Delta_1 \times \Delta_2 > 0$. As mentioned in Section 3.2.1, we use the uniform prior, which is a special case of the beta distribution with $a_{1m} = a_{2m} = 1$, for all m.

For the case where $\Delta_1 \times \Delta_2 < 0$, The prior distribution of $\boldsymbol{p}_m = (p_{1m}, \ldots, p_{4m})$ is assumed to be a Dirichlet distribution with parameter $\boldsymbol{a}_m = (a_{1m}, \ldots, a_{4m})$:

$$\boldsymbol{p}_m \sim \mathcal{D}(\boldsymbol{a}_m).$$

We use $a_{jm} = 1$ for any j and m for the hyperparameters in our empirical analysis.

A.2 Posterior distributions

For the unobserved payoff $\boldsymbol{u}_m$, we assume a standard bivariate normal distribution, $\boldsymbol{u}_m \sim$ i.i.d. $\mathcal{N}(\mathbf{0}, \mathbf{I})$, where the variances are set equal to one for identification. However, other distributions such as a logistic distribution can also be used.

A.2.1 Model A with pure Nash equilibria

In this subsection, we derive the posterior probability density and MCMC simulation procedure for the case where $\Delta_1 < 0$ and $\Delta_2 < 0$. The derivation for the case where $\Delta_1 > 0$ and $\Delta_2 > 0$ is completely analogous and is omitted. When $\boldsymbol{u}_m \sim$ i.i.d. $\mathcal{N}(\mathbf{0}, I)$, we have

$$
\begin{aligned}
P_1(\boldsymbol{\theta}, \boldsymbol{x}_m) &= \Phi(-\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1)\Phi(-\boldsymbol{x}_{2m}'\boldsymbol{\beta}_2), && \text{(A.1)}\\
P_2(\boldsymbol{\theta}, \boldsymbol{x}_m) &= \{\Phi(\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1) - \Phi(\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1 + \Delta_1)\}\Phi(-\boldsymbol{x}_{2m}'\beta_2) \\
&\quad +\Phi(\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1 + \Delta_1)\Phi(-\boldsymbol{x}_{2m}'\boldsymbol{\beta}_2 - \Delta_2), && \text{(A.2)}\\
P_3(\boldsymbol{\theta}, \boldsymbol{x}_m) &= \{\Phi(\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1) - \Phi(\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1 + \Delta_1)\}\Phi(\boldsymbol{x}_{2m}'\boldsymbol{\beta}_2 + \Delta_2) \\
&\quad +\Phi(-\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1)\Phi(\boldsymbol{x}_{2m}'\boldsymbol{\beta}_2), && \text{(A.3)}\\
P_4(\boldsymbol{\theta}, \boldsymbol{x}_m) &= \Phi(\boldsymbol{x}_{1m}'\beta_1 + \Delta_1)\Phi(\boldsymbol{x}_{2m}'\beta_2 + \Delta_2), && \text{(A.4)}\\
P_5(\boldsymbol{\theta}, \boldsymbol{x}_m) &= \{\Phi(\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1) - \Phi(\boldsymbol{x}_{1m}'\boldsymbol{\beta}_1 + \Delta_1)\}\{\Phi(\boldsymbol{x}_{2m}'\boldsymbol{\beta}_2) - \Phi(\boldsymbol{x}_{2m}'\boldsymbol{\beta}_2 + \Delta_2)\}, \\
& && \text{(A.5)}
\end{aligned}
$$

where $\Phi(\cdot)$ denotes a cumulative distribution function of a univariate standard normal distribution. Then the likelihood function $f(\boldsymbol{z}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ is given by

$$
\begin{aligned}
f(\boldsymbol{z}|\boldsymbol{\theta}, \boldsymbol{\lambda}) &= \prod_{m=1}^{M} f(z_m|\boldsymbol{\theta}, \lambda_m), \\
&= \prod_{m=1}^{M} P_1(\boldsymbol{\theta}, \boldsymbol{x}_m)^{I[z_m=1]}\{P_2(\boldsymbol{\theta}, \boldsymbol{x}_m) + \lambda_m P_5(\boldsymbol{\theta}, \boldsymbol{x}_m)\}^{I[z_m=2]} \\
&\quad \times\{P_3(\boldsymbol{\theta}, \boldsymbol{x}_m) + (1-\lambda_m)P_5(\boldsymbol{\theta}, \boldsymbol{x}_m)\}^{I[z_m=3]} P_{4m}(\boldsymbol{\theta}, \boldsymbol{x}_m)^{I[z_m=4]}. \\
& && \text{(A.6)}
\end{aligned}
$$

The posterior probability density is

$$
\pi(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{p}|\boldsymbol{z}) \propto f(\boldsymbol{z}|\boldsymbol{\theta}, \boldsymbol{\lambda})\pi(\boldsymbol{\theta}) \prod_{m=1}^{M} p_m^{(\lambda_m + a_{1m})-1}(1-p_m)^{(1-\lambda_m+a_{2m})-1}, \qquad \text{(A.7)}
$$

where $\pi(\boldsymbol{\theta})$ denotes a probability density function of the truncated normal distribution $\mathcal{TN}_R(\boldsymbol{\theta}_0, \Sigma_0)$. It is easy to see that the conditional posterior probability distributions of λ_m and p_m are

$$
\begin{aligned}
\lambda_m|\boldsymbol{\theta}, p_m, z_m &\sim Bernoulli(q_m), && \text{(A.8)}\\
p_m|\boldsymbol{\theta}, \lambda_m, z_m &\sim \mathcal{B}(a_{1m} + \lambda_m, a_{2m} + 1 - \lambda_m), && \text{(A.9)}
\end{aligned}
$$

where

$$q_m = \frac{p_m^{a_{1m}}(1-p_m)^{a_{2m}-1}f(z_m|\boldsymbol{\theta},\lambda_m=1)}{p_m^{a_{1m}}(1-p_m)^{a_{2m}-1}f(z_m|\boldsymbol{\theta},\lambda_m=1)+p_m^{a_{1m}-1}(1-p_m)^{a_{2m}}f(z_m|\boldsymbol{\theta},\lambda_m=0)}. \tag{A.10}$$

We note that $p_m = q_m$ when $z_m = 1$ or 4.

A.2.2 Model B with pure Nash equilibria

Next, we describe the MCMC implementation for the case where $\Delta_1 > 0$ and $\Delta_2 < 0$. The implementation for the case where $\Delta_1 < 0$ and $\Delta_2 > 0$ can be obtained by switching the labels for the two players. The five choice probabilities are

$$\begin{aligned}
P_1(\boldsymbol{\theta},\boldsymbol{x}_m) &= \Phi(-\boldsymbol{x}'_{1m}\boldsymbol{\beta}_1)\Phi(-\boldsymbol{x}'_{2m}\boldsymbol{\beta}_2), & \text{(A.11)}\\
P_2(\boldsymbol{\theta},\boldsymbol{x}_m) &= \Phi(\boldsymbol{x}'_{1m}\boldsymbol{\beta}_1)\Phi(-\boldsymbol{x}'_{2m}\beta_2-\Delta_2), & \text{(A.12)}\\
P_3(\boldsymbol{\theta},\boldsymbol{x}_m) &= \Phi(-\boldsymbol{x}'_{1m}\boldsymbol{\beta}_1-\Delta_1)\Phi(\boldsymbol{x}'_{2m}\boldsymbol{\beta}_2), & \text{(A.13)}\\
P_4(\boldsymbol{\theta},\boldsymbol{x}_m) &= \Phi(\boldsymbol{x}'_{1m}\beta_1+\Delta_1)\Phi(\boldsymbol{x}'_{2m}\beta_2+\Delta_2), & \text{(A.14)}\\
P_5(\boldsymbol{\theta},\boldsymbol{x}_m) &\equiv \{\Phi(\boldsymbol{x}'_{1m}\boldsymbol{\beta}_1+\Delta_1)-\Phi(\boldsymbol{x}'_{1m}\boldsymbol{\beta}_1)\}\{\Phi(\boldsymbol{x}'_{2m}\boldsymbol{\beta}_2)-\Phi(\boldsymbol{x}'_{2m}\boldsymbol{\beta}_2+\Delta_2)\}, & \text{(A.15)}
\end{aligned}$$

and the likelihood function $f(\boldsymbol{z}|\boldsymbol{\theta},\boldsymbol{\lambda})$ is given by

$$\begin{aligned}
f(\boldsymbol{z}|\boldsymbol{\theta},\boldsymbol{\lambda}) &= \prod_{m=1}^{M} f(z_m|\boldsymbol{\theta},\boldsymbol{\lambda}_m)\\
&= \prod_{m=1}^{M}\prod_{j=1}^{4}\{P_j(\boldsymbol{\theta},\boldsymbol{x}_m)+\lambda_{jm}P_5(\boldsymbol{\theta},\boldsymbol{x}_m)\}^{I[z_m=j]}. \qquad \text{(A.16)}
\end{aligned}$$

Then the posterior probability density is

$$\pi(\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{p}|\boldsymbol{z}) \propto f(\boldsymbol{z}|\boldsymbol{\theta},\boldsymbol{\lambda})\pi(\boldsymbol{\theta})\prod_{m=1}^{M}\prod_{j=1}^{4}p_{jm}^{\lambda_{jm}+a_{jm}-1}, \tag{A.17}$$

where $\pi(\boldsymbol{\theta})$ denotes a probability density function of the truncated normal distribution $\mathcal{TN}_R(\boldsymbol{\theta}_0,\Sigma_0)$. It is easy to see that the conditional posterior probability distributions of $\boldsymbol{\lambda}_m$ and $\boldsymbol{p}_m$ are

$$\begin{aligned}
\boldsymbol{\lambda}_m|\boldsymbol{\theta},\boldsymbol{p}_m,z_m &\sim MN(1,\boldsymbol{q}_m), & \text{(A.18)}\\
\boldsymbol{p}_m|\boldsymbol{\theta},\boldsymbol{\lambda}_m,z_m &\sim \mathcal{D}(\boldsymbol{a}_m+\boldsymbol{\lambda}_m), & \text{(A.19)}
\end{aligned}$$

where $\boldsymbol{q}_m = (q_{1m}, ..., q_{4m})$ such that

$$q_{jm} = \frac{p_{jm}^{a_{jm}} \left(\prod_{l \neq j} p_{lm}^{a_{lm}-1} \right) f(z_m | \boldsymbol{\theta}, \lambda_{jm} = 1, \boldsymbol{\lambda}_{m \backslash j} = \mathbf{0})}{\sum_{k=1}^{4} p_{km}^{a_{km}} \left(\prod_{l \neq k} p_{lm}^{a_{lm}-1} \right) f(z_m | \boldsymbol{\theta}, \lambda_{km} = 1, \boldsymbol{\lambda}_{m \backslash k} = \mathbf{0})}, \quad j = 1, \dots, 4, \quad \text{(A.20)}$$

where $\boldsymbol{\lambda}_{m \backslash j} = \mathbf{0}$ implies that all elements in $\boldsymbol{\lambda}_m$ are equal to zero except λ_{jm}. We implement the Metropolis-Hastings algorithm for the posterior sampling of $\boldsymbol{\theta}$.

B Posterior distribution for mixed Nash models

This appendix provides a detail of the Bayesian estimation for the models with the mixed-strategy Nash equilibrium, which is summarized in Section 3.2.2. We firstly describes Model B because it has a simpler structure, then proceeds to Model A.

B.1 Model B

As mentioned in Section 3.2.2, Model B with the mixed-strategy Nash equilibrium is equivalent to the pure Nash model when we use $\rho_j(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m)$ as the selection proportion p_{jm}. Thus the conditional likelihood function given $\rho_j(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m)$ is

$$\prod_{m=1}^{M} \prod_{j=1}^{4} [P_j(\boldsymbol{\theta}, \boldsymbol{x}_m) + \rho_j(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m) P_5(\boldsymbol{\theta}, \boldsymbol{x}_m)]^{I[z_m = j]}, \quad \text{(B.1)}$$

where $P_j(\boldsymbol{\theta}, \boldsymbol{x}_m),\ j = 1, 2, 3, 4$ and 5 is equivalent to (A.11) - (A.15).

Because the support of $\boldsymbol{u}_m$ depends on $\boldsymbol{\theta}$, we need to integrate it out to obtain the likelihood function for this model. Let $\boldsymbol{g}(\boldsymbol{u}_m | \boldsymbol{\theta}, \boldsymbol{x}_m)$ be the joint density function for (u_{1m}, u_{2m}), which is a normal density with a truncation such that they are located in Region 5. Integrating out $\boldsymbol{u}_m$ from (B.1), we have

$$f(\boldsymbol{z} | \boldsymbol{\theta}) = \prod_{m=1}^{M} \prod_{j=1}^{4} [P_j(\boldsymbol{\theta}, \boldsymbol{x}_m) + R_j(\boldsymbol{\theta}, \boldsymbol{x}_m) P_5(\boldsymbol{\theta}, \boldsymbol{x}_m)]^{I[z_m = j]}, \quad \text{(B.2)}$$

where

$$R_j(\boldsymbol{\theta}, \boldsymbol{x}_m) = \int \rho_j(\boldsymbol{\theta}, \boldsymbol{x}_m, \boldsymbol{u}_m) \boldsymbol{g}(\boldsymbol{u}_m | \boldsymbol{\theta}, \boldsymbol{x}_m) d\boldsymbol{u}_m. \quad \text{(B.3)}$$

To obtain the closed form of $R_j(\boldsymbol{\theta}, \boldsymbol{x}_m)$, it is convenient to switch the notation from $z_m = 1, 2, 3$ and 4 to $(y_{1m}, y_{2m}) = (0, 0), (1, 0), (0, 1)$ and $(1, 1)$. We define the function $r[(y_{1m}, y_{2m}) | \boldsymbol{\theta}, \boldsymbol{x}_m]$ which corresponds to $R_j(\boldsymbol{\theta}, \boldsymbol{x}_m)$ as

$$
\begin{aligned}
& r[(y_{1m}, y_{2m})|\boldsymbol{\theta}, \boldsymbol{x}_m] \\
& \overset{def}{=} \int \sigma^*_{1m}(y_{1m})\sigma^*_{2m}(y_{2m})\boldsymbol{g}(\boldsymbol{u}_m|\boldsymbol{\theta}, \boldsymbol{x}_m)d\boldsymbol{u}_m \\
& = \Big[1 - y_{1m} + \frac{(-1)^{y_{1m}}[\boldsymbol{x}'_{2m}\boldsymbol{\beta}_2 + A_{2m}(\boldsymbol{\theta}, \boldsymbol{x}_m)]}{\Delta_2}\Big]\Big[1 - y_{2m} + \frac{(-1)^{y_{2m}}[\boldsymbol{x}'_{1m}\boldsymbol{\beta}_1 + A_{1m}(\boldsymbol{\theta}, \boldsymbol{x}_m)]}{\Delta_1}\Big],
\end{aligned}
\tag{B.4}
$$

where

$$
\begin{aligned}
A_{im}(\boldsymbol{\theta}, \boldsymbol{x}_m) &= \frac{\phi(\underline{u}_{im}) - \phi(\overline{u}_{im})}{\Phi(\overline{u}_{im}) - \Phi(\underline{u}_{im})}, && \text{(B.5)} \\
\underline{u}_{im}(\boldsymbol{\theta}, \boldsymbol{x}_{im}) &= -\boldsymbol{x}'_{im}\boldsymbol{\beta}_i + \min(-\Delta_i, 0), && \text{(B.6)} \\
\overline{u}_{im}(\boldsymbol{\theta}, \boldsymbol{x}_{im}) &= -\boldsymbol{x}'_{im}\boldsymbol{\beta}_i + \max(-\Delta_i, 0). && \text{(B.7)}
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& r[(y_{1m}, y_{2m}) = (0,0)|\boldsymbol{\theta}, \boldsymbol{x}_m] = R_1(\boldsymbol{\theta}, \boldsymbol{x}_m), \quad && r[(y_{1m}, y_{2m}) = (1,0)|\boldsymbol{\theta}, \boldsymbol{x}_m] = R_2(\boldsymbol{\theta}, \boldsymbol{x}_m), \\
& r[(y_{1m}, y_{2m}) = (0,1)|\boldsymbol{\theta}, \boldsymbol{x}_m] = R_3(\boldsymbol{\theta}, \boldsymbol{x}_m), \quad && r[(y_{1m}, y_{2m}) = (1,1)|\boldsymbol{\theta}, \boldsymbol{x}_m] = R_4(\boldsymbol{\theta}, \boldsymbol{x}_m).
\end{aligned}
$$

Given the likelihood function, the posterior density is

$$
\pi(\boldsymbol{\theta}|\boldsymbol{z}) \propto f(\boldsymbol{z}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}). \tag{B.8}
$$

Because there is no latent variable, the MCMC procedure for Model B is composed only of the posterior sampling for $\boldsymbol{\theta}$. It is implemented by the Metropolis-Hastings algorithm as in the pure Nash model.

B.2 Model A

In Model A, we integrate out $\boldsymbol{u}_m$ as in Model B, while treat $\boldsymbol{\lambda}^{Nash}$ and $\boldsymbol{\lambda}^{pure}$ as latent variables in the manner similar to the pure Nash model. we obtain the likelihood function as

$$
\begin{aligned}
& f(\boldsymbol{z}|\boldsymbol{\theta}, \boldsymbol{\lambda}^{Nash}, \boldsymbol{\lambda}^{pure}) \\
& = \prod_{m=1}^{M}\prod_{j=1}^{4}\Big[P_j(\boldsymbol{\theta}, \boldsymbol{x}_m) + T_j(\boldsymbol{\theta}, \boldsymbol{x}_m, \lambda_m^{Nash}, \lambda_m^{pure})P_5(\boldsymbol{\theta}, \boldsymbol{x}_m)\Big]^{I[z_m=j]},
\end{aligned}
\tag{B.9}
$$

where

$$
\begin{aligned}
T_1(\boldsymbol{\theta}, \boldsymbol{x}_m, \lambda_m^{Nash}, \lambda_m^{pure}) &= (1-\lambda_m^{Nash})R_1(\boldsymbol{\theta}, \boldsymbol{x}_m), && \text{(B.10)}\\
T_2(\boldsymbol{\theta}, \boldsymbol{x}_m, \lambda_m^{Nash}, \lambda_m^{pure}) &= \lambda_m^{Nash}\lambda_m^{pure} + (1-\lambda_m^{Nash})R_2(\boldsymbol{\theta}, \boldsymbol{x}_m), && \text{(B.11)}\\
T_3(\boldsymbol{\theta}, \boldsymbol{x}_m, \lambda_m^{Nash}, \lambda_m^{pure}) &= \lambda_m^{Nash}(1-\lambda_m^{pure}) + (1-\lambda_m^{Nash})R_3(\boldsymbol{\theta}, \boldsymbol{x}_m), && \text{(B.12)}\\
T_4(\boldsymbol{\theta}, \boldsymbol{x}_m, \lambda_m^{Nash}, \lambda_m^{pure}) &= (1-\lambda_m^{Nash})R_4(\boldsymbol{\theta}, \boldsymbol{x}_m), && \text{(B.13)}
\end{aligned}
$$

and the definition of R_j is same as (B.3).

We adopt beta prior distributions for the selection proportions:

$$
\begin{aligned}
p_m^{Nash} &\sim \mathcal{B}(a_{1m}^{Nash}, a_{2m}^{Nash}), && \text{(B.14)}\\
p_m^{pure} &\sim \mathcal{B}(a_{1m}^{pure}, a_{2m}^{pure}). && \text{(B.15)}
\end{aligned}
$$

Then we obtain the joint posterior density as

$$
\begin{aligned}
&\pi(\boldsymbol{\theta}, \boldsymbol{\lambda}^{Nash}, \boldsymbol{\lambda}^{pure}, \boldsymbol{p}^{Nash}, \boldsymbol{p}^{pure}|\boldsymbol{z})\\
&\propto f(\boldsymbol{z}|\boldsymbol{\theta}, \boldsymbol{\lambda}^{Nash}, \boldsymbol{\lambda}^{pure})\pi(\boldsymbol{\theta}) \prod_{m=1}^{M} \left(p_m^{Nash}\right)^{(\lambda_m^{Nash}+a_{1m}^{Nash})-1} \left(1-p_m^{Nash}\right)^{(1-\lambda_m^{Nash}+a_{2m}^{Nash})-1}\\
&\quad \times \prod_{m:\lambda_m^{Nash}=1} (p_m^{pure})^{(\lambda_m^{pure}+a_{1m}^{pure})-1} (1-p_m^{pure})^{(1-\lambda_m^{pure}+a_{2m}^{pure})-1}. \qquad \text{(B.16)}
\end{aligned}
$$

For the latent variables, the conditional posterior distributions are

$$
\begin{aligned}
\lambda_m^{Nash}|\boldsymbol{\theta}, p_m^{Nash}, \lambda_m^{pure}, z_m &\sim Bernoulli(q_m^{Nash}), && \text{(B.17)}\\
p_m^{Nash}|\boldsymbol{\theta}, \lambda_m^{Nash}, z_m &\sim \mathcal{B}(a_{1m}^{Nash}+\lambda_m^{Nash}, a_{2m}^{Nash}+1-\lambda_m^{Nash}), && \text{(B.18)}\\
\lambda_m^{pure}|\boldsymbol{\theta}, \lambda_m^{Nash}, p_m^{pure}, z_m &\sim Bernoulli(1, q_m^{pure}), && \text{(B.19)}\\
p_m^{pure}|\boldsymbol{\theta}, \lambda_m^{pure}, z_m &\sim \mathcal{B}(a_{1m}^{pure}+\lambda_m^{pure}, a_{2m}^{pure}+1-\lambda_m^{pure}), && \text{(B.20)}
\end{aligned}
$$

where

$$
\begin{aligned}
q_m^{Nash} &= (p_m^{Nash})^{a_{1m}^{Nash}}(1-p_m^{Nash})^{a_{2m}^{Nash}-1} f(z_m|\boldsymbol{\theta},\lambda_m^{Nash}=1,\lambda_m^{pure}) \\
&\quad /\Big[(p_m^{Nash})^{a_{1m}^{Nash}}(1-p_m^{Nash})^{a_{2m}^{Nash}-1} f(z_m|\boldsymbol{\theta},\lambda_m^{Nash}=1,\lambda_m^{pure}) \\
&\quad +(p_m^{Nash})^{a_{1m}^{Nash}-1}(1-p_m^{Nash})^{a_{2m}^{Nash}} f(z_m|\boldsymbol{\theta},\lambda_m^{Nash}=0,\lambda_m^{pure})\Big],
\end{aligned} \tag{B.21}
$$

$$
\begin{aligned}
q_m^{pure} &= (p_m^{pure})^{a_{1m}^{pure}}(1-p_m^{pure})^{a_{2m}^{pure}-1} f(z_m|\boldsymbol{\theta},\lambda_m^{pure},\lambda_m^{pure}=1) \\
&\quad /\Big[(p_m^{pure})^{a_{1m}^{pure}}(1-p_m^{pure})^{a_{2m}^{pure}-1} f(z_m|\boldsymbol{\theta},\lambda_m^{Nash},\lambda_m^{pure}=1) \\
&\quad +(p_m^{pure})^{a_{1m}^{pure}-1}(1-p_m^{pure})^{a_{2m}^{pure}} f(z_m|\boldsymbol{\theta},\lambda_m^{Nash},\lambda_m^{pure}=0)\Big],
\end{aligned} \tag{B.22}
$$

$$
f(z_m|\boldsymbol{\theta},\lambda_m^{Nash},\lambda_m^{pure}) = \prod_{j=1}^{4}[P_j(\boldsymbol{\theta},\boldsymbol{x}_m)+P_5(\boldsymbol{\theta},\boldsymbol{x}_m)T_j(\boldsymbol{\theta},\boldsymbol{x}_m,\lambda_m^{Nash},\lambda_m^{pure})]^{I[z_m=j]}. \tag{B.23}
$$

The model parameters $\boldsymbol{\theta}$ is estimated via the Metropolis-Hastings algorithm as in Model B. In summary, we implement the MCMC procedure for Model A as follows:

- For $m = 1, 2, ..., M$,
  - Generate $p_m^{Nash}|\boldsymbol{\theta},\lambda_m^{pure},z_m$.
  - Generate $\lambda_m^{Nash}|\boldsymbol{\theta},\lambda_m^{Nash},p_m^{Nash},z_m$.
  - Generate $p_m^{pure}|\boldsymbol{\theta},\lambda_m^{pure},z_m$.
  - Generate $\lambda_m^{pure}|\boldsymbol{\theta},\lambda_m^{Nash},p_m^{pure},z_m$.
- Generate $\boldsymbol{\theta}|\boldsymbol{\lambda}^{Nash},\boldsymbol{\lambda}^{pure},\boldsymbol{z}$.

C Posterior predictive probability

In the analysis of the entry game, the prediction of choice probabilities often provides rich implications for empirical studies. Using the MCMC samples $\boldsymbol{\theta}^{(r)}$ $(r = 1, \ldots, R)$ from the posterior distribution, it is easy to find the probability of entries $\boldsymbol{y}_{M+1}$ for the new market $M+1$. In this appendix, we present a methodology of the prediction analysis associated with our Bayesian estimation for Model A and B with pure Nash equilibria. We can also conduct the prediction analysis in the similar manner for the models with the mixed strategy Nash equilibria.

For Model A, the posterior predictive probability mass function is

$$
\begin{aligned}
& f(z_{M+1}|\boldsymbol{z}) \\
= & \sum_{\lambda_{M+1}=0}^{1} \int f(z_{M+1}|\boldsymbol{\theta}, \lambda_{M+1}) \frac{p_{M+1}^{(\lambda_{M+1}+a_{1,M+1})-1}(1-p_{M+1})^{(1-\lambda_{M+1}+a_{2,M+1})-1}}{B(a_{1,M+1}, a_{2,M+1})} \\
& \times \pi(\boldsymbol{\theta}|\boldsymbol{z}) d\boldsymbol{\theta} dp_{M+1}, \\
= & \sum_{\lambda_{M+1}=0}^{1} \int f(z_{M+1}|\boldsymbol{\theta}, \lambda_{M+1}) \frac{B(\lambda_{M+1}+a_{1,M+1}, 1-\lambda_{M+1}+a_{2,M+1})}{B(a_{1,M+1}, a_{2,M+1})} \pi(\boldsymbol{\theta}|\boldsymbol{z}) d\boldsymbol{\theta}, \\
= & w_1 \int f(z_{M+1}|\boldsymbol{\theta}, \lambda_{M+1}=1) \pi(\boldsymbol{\theta}|\boldsymbol{z}) d\boldsymbol{\theta} + w_2 \int f(z_{M+1}|\boldsymbol{\theta}, \lambda_{M+1}=0) \pi(\boldsymbol{\theta}|\boldsymbol{z}) d\boldsymbol{\theta},
\end{aligned} \tag{C.1}
$$

where $B(a_{1,M+1}, a_{2,M+1})$ is a beta function which is required as a normalizing constant for $\mathcal{B}(a_{1,M+1}, a_{2,M+1})$, and

$$
w_j = \frac{a_{j,M+1}}{a_{1,M+1}+a_{2,M+1}}. \tag{C.2}
$$

For each $\boldsymbol{\theta}^{(r)}$, we compute:

$$
\hat{f}(z_{M+1}|\boldsymbol{z}) = \frac{w_1}{R} \sum_{r=1}^{R} f(z_{M+1}|\boldsymbol{\theta}^{(r)}, \lambda=1) + \frac{w_2}{R} \sum_{r=1}^{R} f(z_{M+1}|\boldsymbol{\theta}^{(r)}, \lambda=0). \tag{C.3}
$$

For Model B, the posterior predictive probability mass function is:

$$
\begin{aligned}
& f(z_{M+1}|\boldsymbol{z}) \\
= & \sum_{\boldsymbol{\lambda}_{M+1}} \int f(z_{M+1}|\boldsymbol{\theta}, \boldsymbol{\lambda}_{M+1}) \frac{\prod_{j=1}^{4} p_{j,M+1}^{\lambda_{j,M+1}+a_{j,M+1}-1}}{D(\boldsymbol{a}_{M+1})} \pi(\boldsymbol{\theta}|\boldsymbol{z}) d\boldsymbol{\theta} d\boldsymbol{p}_{M+1} \\
= & \sum_{j=1}^{4} w_j \int f(z_{M+1}|\boldsymbol{\theta}, \lambda_{j,M+1}=1, \boldsymbol{\lambda}_{M+1\backslash j}=\mathbf{0}) \pi(\boldsymbol{\theta}|\boldsymbol{z}) d\boldsymbol{\theta},
\end{aligned} \tag{C.4}
$$

where $D(\boldsymbol{a}_{M+1})$ is a normalizing constant for $\mathcal{D}(\boldsymbol{a}_{M+1})$ and the summation is taken over the set $\lambda_{j,M+1} = 1$, $\boldsymbol{\lambda}_{M+1\backslash j} = \mathbf{0}$ for $j = 1, \ldots, 4.$, and

$$
w_j = \frac{a_{j,M+1}}{\sum_{k=1}^{4} a_{j,M+1}}. \tag{C.5}
$$

For each $\boldsymbol{\theta}^{(r)}$, we compute:

$$\begin{aligned}
&\hat{f}(z_{M+1}|\boldsymbol{z}) \\
&= \sum_{j=1}^{4} \frac{w_j}{R} \sum_{r=1}^{R} f(z_{1,M+1}|\boldsymbol{\theta}^{(r)}, \lambda_{j,M+1} = 1, \boldsymbol{\lambda}_{M+1\backslash j} = \mathbf{0}).
\end{aligned} \tag{C.6}$$

D Estimated posterior densities and sample paths

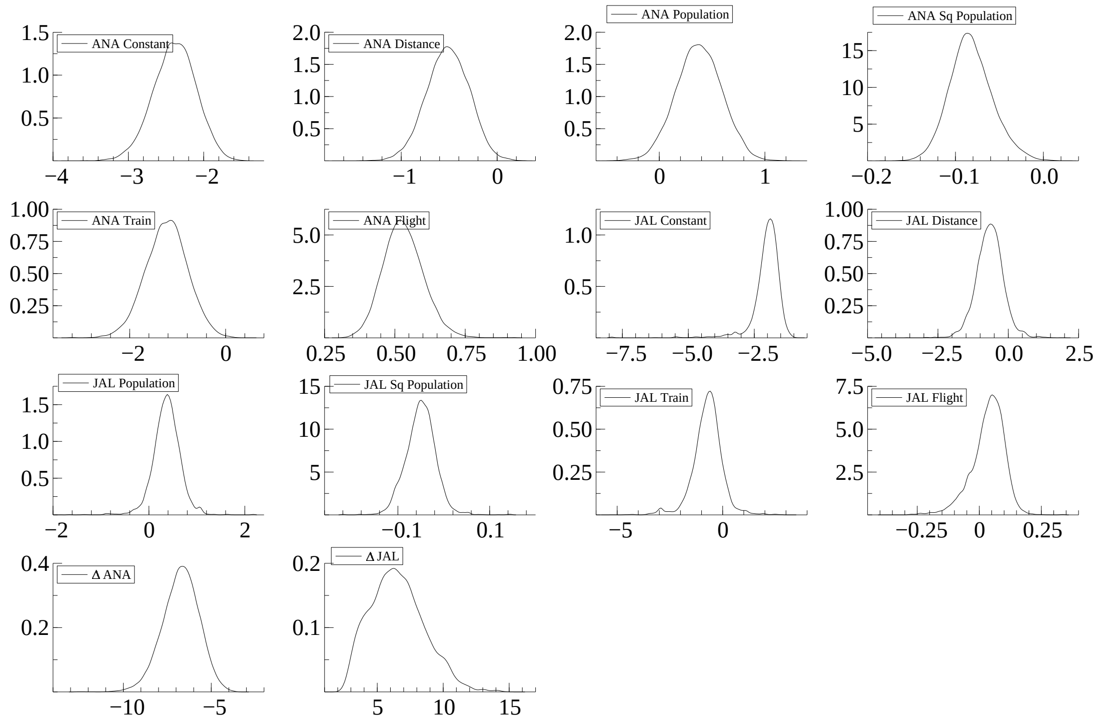

Figure 3: Posterior densities of parameters, $\boldsymbol{\beta}$ and $\boldsymbol{\Delta}$ (Pure Nash, $\Delta_{ANA} < 0, \Delta_{JAL} > 0$).

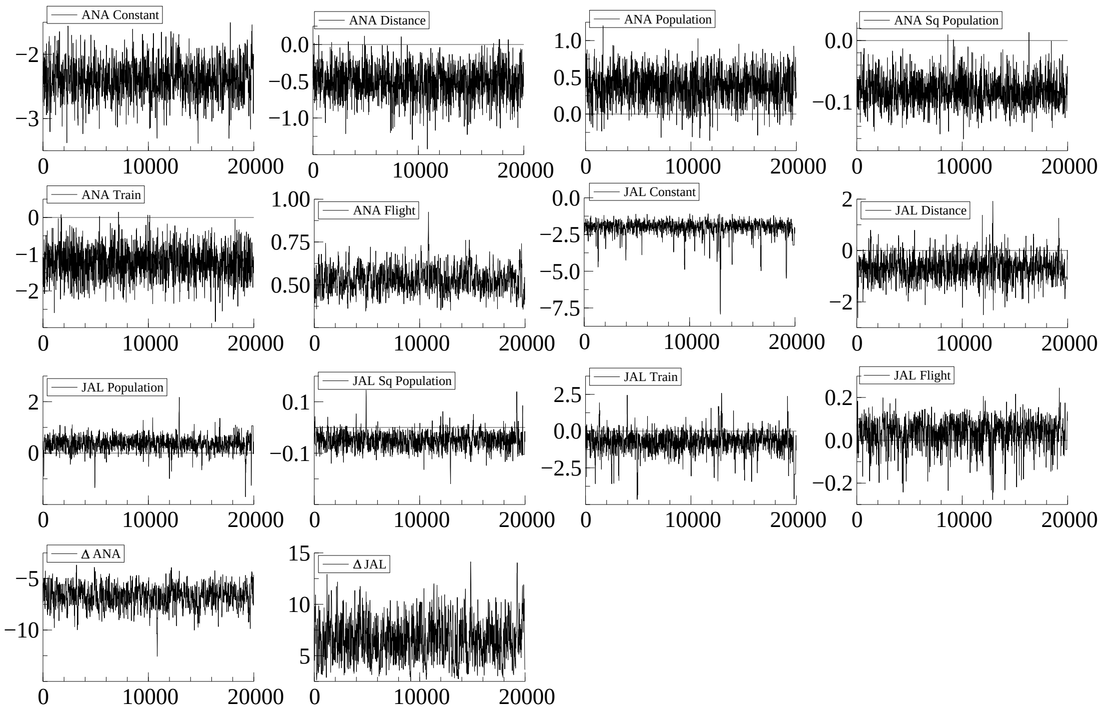

Figure 4: Sample paths for parameters, $\boldsymbol{\beta}$ and $\boldsymbol{\Delta}$ (Pure Nash, $\Delta_{ANA} < 0, \Delta_{JAL} > 0$)