Identification of Structural Parameters in Dynamic Discrete Choice Games with Fixed Effects Unobserved Heterogeneity*

Victor Aguirregabiria †
University of Toronto
CEPR

Jiaying Gu ‡
University of Toronto

Pedro Mira §
CEMFI

May 10, 2023

Abstract

In the structural estimation of dynamic discrete choice games, misspecification of unobserved heterogeneity can lead to significant biases in two types of structural parameters: those related to dynamic state dependence, such as costs of adjustment or switching, and those capturing strategic interactions among players, such as competition and peer effects. We investigate the identification of these parameters within models that incorporate market unobserved heterogeneity with a fixed effect structure. This structure assumes a nonparametric distribution conditional on the initial values of the state variables. To tackle this identification problem, we extend the existing *Sufficient Statistics - Conditional Maximum Likelihood* approach to dynamic games with multiple equilibria, allowing for partial identification. We provide identification results for different types of games, taking into consideration various factors such as simultaneous versus sequential moves, myopic versus forward-looking players, and one-direction versus two-direction strategic interactions.

Keywords: Panel data; Dynamic discrete choice games; Fixed effects; Unobserved heterogeneity; Structural state dependence; Identification; Sufficient statistics.

JEL codes: C23; C25; C41; C51; C61.

---

*We thank comments from seminar participants at the International Panel Data Conference in Vilnius, IAAE annual conference in Nicosia, the 3rd conference on Structural Dynamic Models in Chicago, and the Banff Empirical Microeconomics Workshop.

†Department of Economics, University of Toronto. 150 St. George Street, Toronto, ON, M5S 3G7, Canada, victor.aguirregabiria@utoronto.ca.

‡Department of Economics, University of Toronto. 150 St. George Street, Toronto, ON, M5S 3G7, Canada, jiaying.gu@utoronto.ca.

§CEMFI. Casado del Alisal, 5, Madrid, Spain. mira@cemfi.es.

1 Introduction

Dynamic games are valuable tools for analyzing economic and social phenomena involving intertemporal agent interactions. The structural estimation of dynamic games has received considerable attention, particularly in the study of oligopoly competition dynamics (Ericson and Pakes, 1995) with empirical applications across various industries.[1] Furthermore, econometric models of dynamic games have been applied to study dynamic interactions within households (Eckstein and Lifshitz, 2015), long-term care decisions (Sovinsky and Stern, 2016), electoral competition (Sieg and Yoon, 2017), and the ratification of international treaties (Wagner, 2016) among other topics. Moreover, a substantial body of literature exists on dynamic discrete choice models with social interactions, wherein agents do not exhibit forward-looking behavior (Brock and Durlauf, 2007, Blume, Brock, Durlauf, and Ioannides, 2011).

In dynamic games, the model predictions heavily rely on two types of structural parameters: those that capture dynamic state dependence, encompassing factors like costs of switching, adjustment, investment, or entry and exit (referred to as the *dynamic* part of the parameters), and those that represent the impact of other players' actions on a player's payoff, arising from competition, spillovers, peer effects, or social interactions (referred to as the *game* part of the structure). The identification of these parameters critically depends on the model's assumptions concerning the stochastic properties of variables known to the players but unobservable to the researcher, which we can refer to as the specification of *unobserved heterogeneity*.

In dynamic models, it is widely acknowledged that neglecting or misspecifying persistent unobserved heterogeneity can result in significant biases when estimating structural parameters that capture genuine dynamics (Heckman, 1981). The presence of spurious dynamics due to unobserved heterogeneity can become entangled with true dynamics arising from state dependence. Similarly, within the literature on game estimation, it is well-known that disregarding correlated unobserved heterogeneity across players can lead to substantial biases in estimating structural parameters that capture strategic or social interactions among players (Bresnahan and Reiss, 1991, Blume, Brock, Durlauf, and Ioannides, 2011). This common unobserved heterogeneity can become confounded with strategic, social, or peer effects.

In this paper, we investigate the identification of dynamic games in empirical applications

[1] Recent applications include industries such as automobiles (Hashmi and Biesebroeck, 2016), airlines (Aguirregabiria and Ho, 2012), pharmaceuticals (Ching, 2010, Gallant, Hong, and Khwaja, 2018), procurement auctions (Jofre-Bonet and Pesendorfer, 2003, Groeger, 2014), construction materials (Ryan, 2012, Collard-Wexler, 2013), hotels (Suzuki, 2013), microprocessors (Goettler and Gordon, 2011), hard drives (Igami, 2017), commercial radio (Sweeting, 2013, Jeziorski, 2014), movies (Einav, 2010, Takahashi, 2015), medical services (Dunne, Klimek, Roberts, and Xu, 2013), shipbuilding (Kalouptsidi, 2014), fishing (Huang and Smith, 2014), and retail stores (Aguirregabiria and Mira, 2007, Igami and Yang, 2016), among others.

where players are observed playing the game across a large number of markets and a small number of periods. We study the identification of these models when there exists time-invariant unobserved heterogeneity at the market or player-market level which follows a nonparametric distribution, thus adhering to a fixed effects panel data model.

We expand the application of the fixed effect conditional likelihood method, initially introduced by Cox (1958), Rasch (1961), Andersen (1970), and Chamberlain (1980), to dynamic discrete choice games. This method involves deriving sufficient statistics for the incidental parameters (representing the fixed effects) and maximizing the likelihood function of the data conditional on these statistics. One notable advantage of this approach is its robustness against misspecification of the distribution of unobserved heterogeneity, ensuring the robust estimation of structural parameters. Furthermore, this method offers computational simplicity, adding to its appeal.

For those versions of the model where the conditional likelihood method fails to identify all the structural parameters in our model, we employ a functional differencing method as proposed by Bonhomme, (2012). Specifically, we utilize a variant of the functional differencing technique recently introduced by Dobronyi, Gu, and Kim (2021) which shares similarities with the approach adopted by Honoré and Weidner (2020). This approach is based on deriving a comprehensive set of moment conditions and moment inequalities that are implied by the fixed effects dynamic model. Through our analysis, we demonstrate that this methodology successfully identifies certain crucial parameters that remain unidentified when employing the conditional likelihood method. By incorporating the functional differencing method into our study, we enhance the identification of important structural parameters that would have otherwise been overlooked.

Our paper contributes to the literature on the identification and estimation of dynamic games with unobserved heterogeneity. All previous studies in this field adopted a random effects approach, employing a finite mixture specification of the unobserved heterogeneity, and imposing restrictions on the initial conditions, e.g., Aguirregabiria and Mira, (2007), Kasahara and Shimotsu (2009), Arcidiacono and Miller (2011), among others. In contrast, our research focuses on the identification of structural parameters without enforcing restrictions on the distribution of the unobserved heterogeneity or the initial conditions. By relaxing these assumptions, we provide a more flexible framework for studying dynamic games with unobserved heterogeneity.

This paper also contributes to the literature on the identification and estimation of structural dynamic discrete choice models with fixed effects. We build upon and extend recent work by Aguirregabiria, Gu, and Luo (2021) who investigate the identification of single-agent dynamic structural models. Extending the identification to games with multiple equilibria is not a straightforward task because these game models do not yield a unique prediction for the probability of a choice history; instead, they provide bounds. However, we develop a method for obtaining sufficient statistics for the contribution of the incidental parameters to these bounds. Moreover, we show that this approach leads to the partial identification of the structural parameters. To the best of our knowledge, our paper represents the first attempt to combine the fixed effects - sufficient statistics approach with bounds and partial identification.

Our paper relates to Honoré and Kyriazidou (2019) and Honoré and De Paula (2021) who present identification results for some panel data bivariate dynamic logit models. We extend their findings by delving into models that incorporate contemporaneous effects between dependent variables, forward-looking players, and multiple equilibria.

The rest of the paper is organized as follows. Section 2 describes the model and assumptions. Section 3 presents our identification results. We distinguish two versions of the model depending on whether players are *myopic* (section 3.1) or *forward-looking* (section 3.2). In section 4, we illustrate our identification results with an empirical application. We summarize and conclude in section 5.

2 Model

2.1 Framework

In our study, we specifically focus on two-player binary choice games. However, in Section 2.5, we discuss various extensions that encompass scenarios with more than two choices or involve more than two players. For our analysis, we assign indices i and j to represent the two players, such that $i, j \in \{1, 2\}$. To capture the temporal dimension, we consider discrete-time and utilize the index t, which ranges from 1 to T, to represent different periods. The game between the two players takes place within a designated *market*. The definition of a market may vary depending on the specific empirical application and can refer to a geographic location, a school, a family, an industry, an election, and so on. To denote different markets, we employ the index m, with m belonging to the set $\{1, 2, ..., M\}$. For the moment, for notational simplicity, we omit the market subindex.

In each period t, the players in the game make a binary decision, which we represent using the variables $y_{1t} \in \{0, 1\}$ and $y_{2t} \in \{0, 1\}$. The objective of each player is to maximize their expected and discounted intertemporal payoffs. This is expressed as: $\mathbb{E}_t \left[\sum_{s=0}^{\infty} \delta_i^s \, U_{i,t+s}\right]$. Here, $\delta_i \in [0, 1]$ represents the discount factor of player i in market m. U_{it} represents the one-period payoff for player i. The utility function has the following structure.

$$U_{it} \;=\; u_i \left(y_{it}, \; y_{jt}, \; y_{i,t-1}, \; y_{j,t-1}\right) + \varepsilon_{it} \left(y_{it}\right). \tag{1}$$

$u_i(\cdot)$ is a utility function that depends on the current and previous actions of the two players. The arguments (y_{it}, y_{jt}) capture contemporaneous strategic effects between the players, indicating how the choice of one player, j, may influence the payoff of the other player, i, in the same period. The arguments $(y_{i,t-1}, y_{j,t-1})$ capture state dependence with respect to the lagged value of the players' actions. They represent factors such as adjustment costs or switching costs, which influence a player's current decision based on their previous action. The variables $\varepsilon_{it}(0)$ and $\varepsilon_{it}(1)$ are observable to the players but unobservable to the researcher. They are independently and identically distributed over (i, m, t, y_i), following a type I Extreme Value distribution. These unobservable terms capture the random shocks or idiosyncratic components that affect the players' payoffs and choices in each period.

We consider games of complete information. Following the majority of the empirical literature on dynamic discrete games, we assume that players' decisions are derived from a *Markov Perfect Equilibrium* (MPE). This assumption implies that players' strategies solely depend on state variables that are relevant to their payoffs. In any given period t, player i bases her action on the variables known to her which have an impact on her own payoff or the payoffs of other players at period t. The vector of state variables that are relevant to the payoffs in this game is denoted as $(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_t)$, where $\mathbf{y}_{t-1} \equiv (y_{1,t-1}, y_{2,t-1})$ and $\boldsymbol{\varepsilon}_t \equiv (\varepsilon_{1t}(0), \varepsilon_{1t}(1), \varepsilon_{2t}(0), \varepsilon_{2t}(1))$. A strategy function for player i can be represented as $\sigma_i(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_t)$.

In this dynamic game, a Markov Perfect Equilibrium (MPE) consists of a pair of strategy functions, one for each player, such that a player's strategy maximizes her intertemporal payoff at any state of the game while taking the other player's strategy function as given. Let $V_i^{\sigma}(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_t)$ represents player i's value function for a given strategy of player j. The decision problem for player i can be formulated using the following Bellman equation:[2]

$$V_i^{\sigma}(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_t) = \max_{y_{it} \in \{0,1\}} \left\{ u_i\left(y_{it}, y_{jt}, \mathbf{y}_{t-1}\right) + \varepsilon_{it}(y_{it}) + \delta_i \int V_i^{\sigma}(\mathbf{y}_t, \boldsymbol{\varepsilon}_{t+1})\ g(\boldsymbol{\varepsilon}_{t+1})\ d\boldsymbol{\varepsilon}_{t+1} \right\} \tag{2}$$

The integral accounts for the expectation over $\boldsymbol{\varepsilon}_{t+1}$, and $g(\boldsymbol{\varepsilon}_{t+1})$ is the density function of $\boldsymbol{\varepsilon}_{t+1}$.

The model can be characterized by the following system of best-response equations:

$$\begin{aligned} y_{1t} &= 1\left\{ \widetilde{u}_1\left(y_{2t}, \mathbf{y}_{t-1}\right) + \widetilde{V}_1\left(y_{2t}\right) - \varepsilon_{1t} \geq 0 \right\} \\ y_{2t} &= 1\left\{ \widetilde{u}_2\left(y_{1t}, \mathbf{y}_{t-1}\right) + \widetilde{V}_2\left(y_{1t}\right) - \varepsilon_{2t} \geq 0 \right\} \end{aligned} \tag{3}$$

with $\varepsilon_{it} \equiv \varepsilon_{it}(0) - \varepsilon_{it}(1)$. Here, $\widetilde{u}_i\left(y_{jt}, \mathbf{y}_{t-1}\right)$ represents the utility difference $u_i\left(1, y_{jt}, \mathbf{y}_{t-1}\right) -$

[2]This Bellman equation accounts for the fact that the other player's action, y_{jt}, is equal to $\sigma_j(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_t)$, which is known to player i given the state $(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_t)$.

$u_i\left(0, y_{jt}, \mathbf{y}_{t-1}\right)$. The term $\widetilde{V}_i\left(y_{jt}\right)$ captures the difference in continuation values:

$$\widetilde{V}_i\left(y_{jt}\right) \equiv \delta_i \int \left(V_i^{\sigma}(1, y_{jt}, \boldsymbol{\varepsilon}_{t+1}) - V_i^{\sigma}(0, y_{jt}, \boldsymbol{\varepsilon}_{t+1})\right) g(\boldsymbol{\varepsilon}_{t+1})\, d\boldsymbol{\varepsilon}_{t+1} \tag{4}$$

Given $(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_t)$, the model assumes that the realized values (y_{1t}, y_{2t}) represent a solution to the system of equations presented in (3).

2.2 Structural and incidental parameters

Let us now provide the specification of the utility function u_i. To differentiate between incidental and interest parameters, we explicitly introduce the market subindex m. The parameters that vary across markets are considered unrestricted and are treated as fixed effects or incidental parameters. Our focus lies in examining the identification of parameters that vary across players but are assumed to be constant across markets.

The utility difference $\widetilde{u}_{im}(y_{jmt}, \mathbf{y}_{m,t-1}) \equiv u_{im}(1, y_{jmt}, \mathbf{y}_{m,t-1}) - u_{im}(0, y_{jmt}, \mathbf{y}_{m,t-1})$ has the following structure:

$$\widetilde{u}_{imt} \;=\; \alpha_{im} + \beta_i\; y_{im,t-1} + \gamma_i\; y_{jmt} + \lambda_i\; y_{jm,t-1}. \tag{5}$$

Here, α_{im} captures unobservable market and player characteristics that are not observable to the researcher. To account for these unobservable factors, we define the vector $\boldsymbol{\alpha}_m \equiv (\alpha_{1m}, \alpha_{2m})$, which represents the fixed effects specific to market m. These fixed effects are referred to as the incidental parameters of the model. The term $\beta_i y_{im,t-1}$ captures state dependence with respect to the lagged value of the player's action. It incorporates factors such as adjustment costs or switching costs, influencing a player's current decision based on their previous action. The term $\gamma_i y_{jmt}$ captures contemporaneous strategic effects between the players' actions. It accounts for how the choice of one player, j, may influence the payoff of the other player, i, in the same period. The term $\lambda_i y_{jm,t-1}$ represents state dependence with respect to the lagged value of the other player's action. It captures the dynamic strategic interactions between the two players, reflecting how a player's previous choice may have an impact on the other player's payoff. Together, these components form the structure of the model, incorporating fixed effects, contemporaneous strategic effects, and state dependence, to capture the dynamics of the two-player binary choice game.

By substituting the expression for the utility difference from equation (5) into the best response equations in (3), we obtain the system of equations defining the econometric model in this paper:

$$
\begin{cases} y_{1mt} & = & 1\left\{\alpha_{1m} + \beta_1 \ y_{1m,t-1} + \gamma_1 \ y_{2mt} + \lambda_1 \ y_{2m,t-1} + \widetilde{V}_{1m}\left(y_{2mt}\right) - \varepsilon_{1mt} \geq 0\right\} \\ \\ y_{2mt} & = & 1\left\{\alpha_{2m} + \beta_2 \ y_{2m,t-1} + \gamma_2 \ y_{1mt} + \lambda_2 \ y_{1m,t-1} + \widetilde{V}_{2m}\left(y_{1mt}\right) - \varepsilon_{2mt} \geq 0\right\} \end{cases} \tag{6}
$$

It is important to note that the continuation values $\widetilde{V}_{im}(0)$ and $\widetilde{V}_{im}(1)$ are also incidental parameters since they are functions of $\boldsymbol{\alpha}_m$

2.3 Multiple equilibria and probabilities of game outcomes

The model exhibits two forms of the multiple equilibrium problem. First, given the model's primitives, there can exist multiple strategy functions, denoted as $\sigma_{1m}(.)$ and $\sigma_{2m}(.)$, that satisfy the system of best response restrictions characterizing the MPE of the model. Second, even if we fix the continuation value functions, $\widetilde{V}_{1m}(.)$ and $\widetilde{V}_{2m}(.)$, there are specific combinations of the state variables, $\mathbf{y}_{m,t-1}$ and $\boldsymbol{\varepsilon}_{mt}$, for which the model generates multiple predictions regarding the best response values of (y_{1mt}, y_{2mt}). This issue resembles the problem of multiple equilibria observed in static games of complete information, as documented in seminal studies such as Bresnahan and Reiss (1991) and Tamer (2003).

The model implies a partition of the space of the unobservables $(\varepsilon_{1t}, \varepsilon_{2t})$ such that each region in the partition corresponds to a prediction (or multiple predictions) about players' choices. The form of this partition depends on the sign of the parameters γ_1 and γ_2. For the sake of concreteness, here we assume that players' decisions are strategic substitutes such that $\gamma_1 \leq 0$ and $\gamma_2 \leq 0$. Figure 1 represents the threshold values for ε_{1t} and ε_{2t} that define this partition. We use this figure to describe the regions in the space of $(\varepsilon_{1t}, \varepsilon_{2t})$ associated with different outcomes (y_{1t}, y_{2t}).

For each player i, we can define two threshold values of variable ε_{it}: a lower threshold e_{it}^L and an upper threshold e_{it}^U with the following definitions:

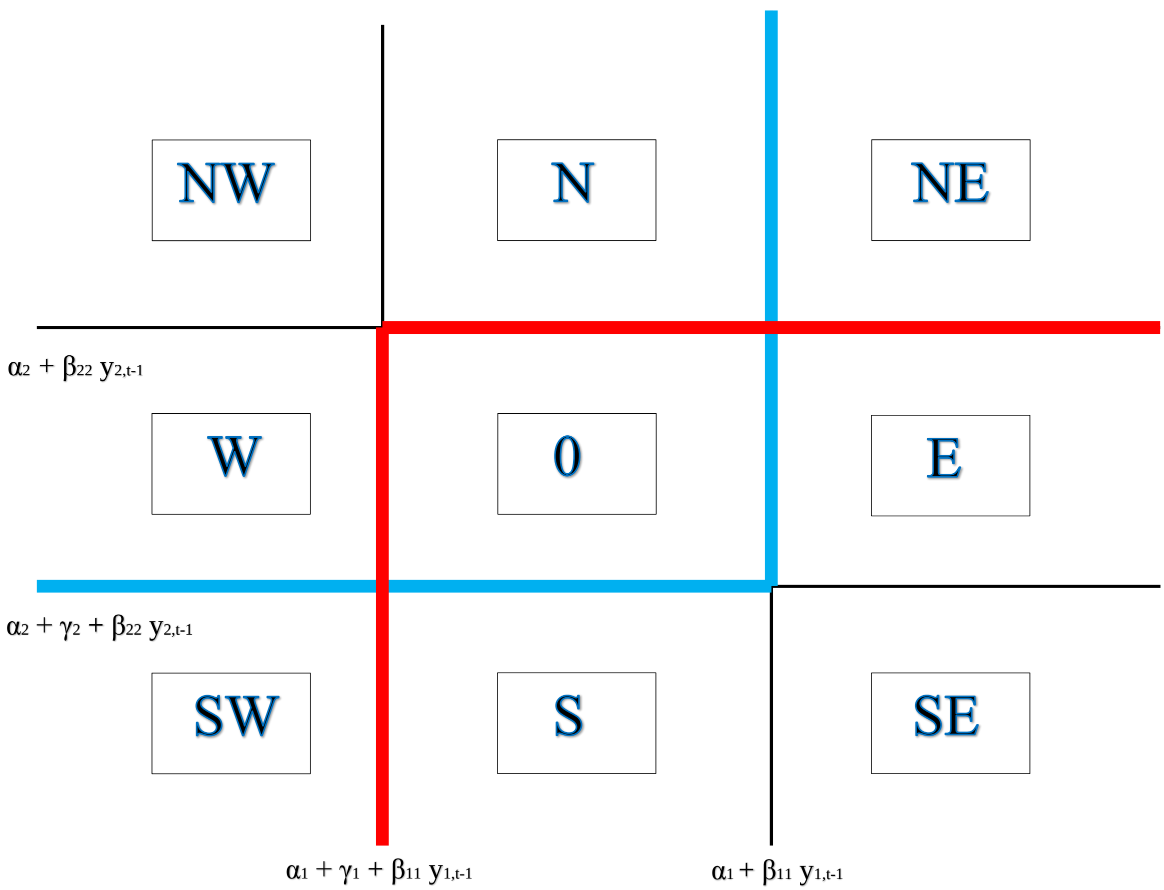
$$
\begin{cases} e_{it}^L & \equiv & \alpha_i + \widetilde{V}_i(1) + \gamma_i + \beta_i \ y_{i,t-1} + \lambda_i \ y_{j,t-1} \\ \\ e_{it}^U & \equiv & \alpha_i + \widetilde{V}_i(0) + \beta_i \ y_{i,t-1} + \lambda_i \ y_{j,t-1} \end{cases} \tag{7}
$$

These four lines in the space of $(\varepsilon_{1t}, \varepsilon_{2t})$: two vertical lines associated to values e_{1t}^L and e_{1t}^U, and two horizontal lines associated to values e_{2t}^L and e_{2t}^U. These four lines divide the space of $(\varepsilon_{1t}, \varepsilon_{2t})$ into nine quadrangles. It is convenient to label these quadrangles using the cardinal directions, i.e., Northwest (NW), North (N), Northeast (NE), etc.

The outcome of the game is $(y_{1t}, y_{2t}) = (1,1)$ if and only if $\varepsilon_{1t} \leq e^L_{1t}$ and $\varepsilon_{2t} \leq e^L_{2t}$ which corresponds to the Southwest (SW) quadrangle. Similarly, the outcome of the game is $(y_{1t}, y_{2t}) = (0,0)$ if and only if $\varepsilon_{1t} \geq e^U_{1t}$ and $\varepsilon_{2t} \geq e^U_{2t}$ which corresponds to the Northeast (NE) quadrangle. Therefore, the model provides unique predictions for the probabilities $\mathbb{P}((y_{1t}, y_{2t}) = (1,1)|\mathbf{y}_{t-1}, \boldsymbol{\alpha})$ and $\mathbb{P}((y_{1t}, y_{2t}) = (0,0)|\mathbf{y}_{t-1}, \boldsymbol{\alpha})$. That is,

$$\begin{cases} \mathbb{P}(0,0 \mid \mathbf{y}_{t-1}, \boldsymbol{\alpha}) = \dfrac{1}{1+\exp\{e^U_{1t}\}} \dfrac{1}{1+\exp\{e^U_{2t}\}} \\ \\ \mathbb{P}(1,1 \mid \mathbf{y}_{t-1}, \boldsymbol{\alpha}) = \dfrac{\exp\{e^L_{1t}\}}{1+\exp\{e^L_{1t}\}} \dfrac{\exp\{e^L_{2t}\}}{1+\exp\{e^L_{2t}\}} \end{cases} \quad (8)$$

Figure 1: Regions in the Space of $(\varepsilon_1, \varepsilon_2)$



The quadrangle in the center of Figure 1 – labeled as O – is associated with two possible outcomes or equilibria of the game: $(y_{1t}, y_{2t}) = (1,0)$ and $(y_{1t}, y_{2t}) = (0,1)$. This region with multiple equilibria implies that the model does not have unique predictions on the probabilities $\mathbb{P}(0,1|\mathbf{y}_{t-1}, \boldsymbol{\alpha})$ and $P(1,0|\mathbf{y}_{t-1}, \boldsymbol{\alpha})$. However, the model establishes bounds on the values of these probabilities.

The upper bound to the probability of outcome $(1,0)$ is given by the region up and to the left of the blue right angle: quadrangles NW, N, W, and 0. The upper bound to the probability of outcome $(0,1)$ is associated with the region down and to the right of the red right angle: quadrangles 0, E, S, and SE. These bounds have a logit structure: they are the product of two logit probabilities:

$$\begin{cases} U(0,1|\mathbf{y}_{t-1}, \boldsymbol{\alpha}) & \equiv & \dfrac{1}{1+\exp\{e_{1t}^L\}} \dfrac{\exp\{e_{2t}^U\}}{1+\exp\{e_{2t}^U\}} \\ \\ U(1,0|\mathbf{y}_{t-1}, \boldsymbol{\alpha}) & \equiv & \dfrac{\exp\{e_{1t}^U\}}{1+\exp\{e_{1t}^U\}} \dfrac{1}{1+\exp\{e_{2t}^L\}} \end{cases} \tag{9}$$

The lower bound to the probability of outcome $(0,1)$ is given by regions $\{E, E, SE\}$: the upper bound excluding quadrangle O where multiple equilibria exist. Unfortunately, this sharp lower bound does not have a logit structure. Without the logit structure, it is not possible to derive sufficient statistics for $\boldsymbol{\alpha}$ (Chamberlain, 2010). For this reason, we use non-sharp lower bounds which have a logit structure. We consider two different lower bounds: a lower bound using region $\{E, SE\}$, and another lower bound using region $\{S, SE\}$. These regions imply the following (logit) lower bounds for the probability of outcome $(0,1)$:

$$\begin{cases} L^{\{E,SE\}}(0,1|\mathbf{y}_{t-1}, \boldsymbol{\alpha}) & \equiv & \dfrac{1}{1+\exp\{e_{1t}^U\}} \dfrac{\exp\{e_{2t}^U\}}{1+\exp\{e_{2t}^U\}} \\ \\ L^{\{S,SE\}}(0,1|\mathbf{y}_{t-1}, \boldsymbol{\alpha}) & \equiv & \dfrac{1}{1+\exp\{e_{1t}^L\}} \dfrac{\exp\{e_{2t}^L\}}{1+\exp\{e_{2t}^L\}} \end{cases} \tag{10}$$

The lower bound to the probability of outcome $(1,0)$ is given by regions $\{NW, N, W\}$: the region for the upper bound excluding quadrangle O. Similarly, as before, the probability of this sharp region does not have a logit structure. We use non-sharp lower bounds with a logit structure: a lower bound using region $\{NW, W\}$, and another lower bound using region $\{NW, N\}$.

These regions imply the following (logit) lower bounds for the probability of outcome $(1,0)$:

$$
\begin{cases}
L^{\{NW,SW\}}(1,0|\mathbf{y}_{t-1},\boldsymbol{\alpha}) \equiv \dfrac{\exp\left\{e_{1t}^{L}\right\}}{1+\exp\left\{e_{1t}^{L}\right\}}\dfrac{1}{1+\exp\left\{e_{2t}^{L}\right\}} \\[2ex]
L^{\{NW,N\}}(1,0|\mathbf{y}_{t-1},\boldsymbol{\alpha}) \equiv \dfrac{\exp\left\{e_{1t}^{U}\right\}}{1+\exp\left\{e_{Ut}^{L}\right\}}\dfrac{1}{1+\exp\left\{e_{2t}^{U}\right\}}
\end{cases}
\tag{11}
$$

2.4 Different versions of the model

We classify different versions of the model based on three criteria that are essential for our identification results.

a. Myopic versus forward-looking players: A player is considered myopic if her discount factor δ_i is zero, indicating a lack of forward-looking behavior. On the other hand, if a player has a non-zero discount factor, she is classified as forward-looking. In the case of myopic players, the continuation value $\widetilde{V}_i$ is zero.

b. Contemporaneous strategic interactions: The presence and nature of contemporaneous strategic interactions in the game depend on the values of parameters γ_1 and γ_2. Specifically, the game can exhibit two-direction interactions if both γ_1 and γ_2 are non-zero. If one of these parameters is zero, the interactions become one-directional. Lastly, if both parameters are zero, there are no contemporaneous strategic interactions between the players.

b. Sequential versus simultaneous moves: Players can make their choices, y_{1t} and y_{2t}, either simultaneously or sequentially in each period t. The choice between sequential and simultaneous moves affects the set of equilibria that the model exhibits. The models without or with one-direct contemporaneous interactions have a unique equilibrium, such that having sequential or simultaneous moves does not affect the equilibrium outcome. In contrast, this assumption matters in a game with two-direction interactions. In the case of simultaneous moves, the model has multiple equilibria as described earlier. In a game with sequential moves, a unique equilibrium exists. For instance, in a sequential move game where player 1 moves first, if we are in region 0 of Figure 1, the outcome $(1,0)$ is the only equilibrium (known as Subgame Perfect Nash equilibrium) for that specific region of $(\varepsilon_{1t}, \varepsilon_{2t})$. Player 1 knows that if she chooses $y_{1t} = 1$, player 2 will choose $y_{2t} = 0$, and if she chooses $y_{1t} = 0$, player 2 will choose $y_{2t} = 1$. Consequently, player 1's choice determines the selection of either equilibrium $(0,1)$ or $(1,0)$. Player 1 chooses the equilibrium that maximizes her profit. With $\gamma_1 \leq 0$, the equilibrium that yields the highest profit for player 1 is $(1,0)$. Therefore, in the sequential move game, region 0 in the $(\varepsilon_{1t}, \varepsilon_{2t})$ space is uniquely associated with outcome $(1,0)$.

By considering these criteria, we can distinguish different versions of the model and explore their implications for the identification of the structural parameters.

Table 1
Different Models and Summary of Identification Results

No Contemporaneous Interactions $\gamma_1 = \gamma_2 = 0$	**One-Direction Interactions** $\gamma_1 = 0$	**Two-Direction Interactions Sequential Move**	**Two-Direction Interactions Simultaneous Move**
MYOPIC PLAYERS			
Point iden. β_1, β_2, λ_1, λ_2	Point iden. β_1, β_2, λ_1, γ_2	Point iden. β_1, β_2 Partial iden. γ_1, γ_2	Partial iden. β_1, β_2, γ_1, γ_2
FORWARD-LOOKING PLAYERS			
Point iden. β_1, β_2	Point iden. β_1, β_2 Partial iden. γ_2	Point iden. β_1, β_2	Partial iden. β_1, β_2

Table 1 provides an overview of the various versions of the model examined in this paper and summarizes the identification results. The table highlights several key patterns emerging from our findings. First, we observe that point identification of the dynamic parameters β_1 and β_2 is more general compared to the identification of strategic interactions γ_1 and γ_2. This implies that we can obtain precise estimates for the dynamic parameters in a broader range of scenarios. Second, we can achieve point identification only by imposing restrictions related to myopic behavior, sequential moves, or the nature of strategic interactions. Third, it is noteworthy that under the assumption of sequential moves, we achieve point identification for the dynamic parameters β without the need to restrict players' discount factors δ or strategic parameters γ. This finding suggests that the sequential move framework alone can provide valuable information for identifying dynamic parameters.

Overall, Table 1 highlights the varying identification outcomes across different versions of the model and emphasizes the significance of incorporating specific assumptions to achieve precise estimates for the parameters of interest.

2.5 Extensions of the model

a. Strictly exogenous state variables $\mathbf{x}_{it}$.

b. Duration dependence.

c. Multinomial choice.

d. More than two players.

3 Identification

The sampling framework involves a random sample of M markets. Within each market, the data consists of the observed sequence of choices made between periods 1 and T, as well as the initial conditions (y_{1m0}, y_{2m0}). The number of markets M is large and T is small. To simplify notation, we will omit the market subindex m for the remainder of this section. To represent the complete history of choices within a market, we use the vector $\widetilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, ..., T)$.

We use $\boldsymbol{\theta}$ to represent the vector of structural parameters $(\beta_1, \beta_2, \gamma_1, \gamma_2, \lambda_1, \lambda_2)$, and $\boldsymbol{\alpha}$ to represent the incidental parameters or fixed effects. The model is a fixed effects model in the sense that the joint probability distribution of the incidental parameters and the initial conditions (y_{10}, y_{20}) is nonparametrically specified. We are interested in the identification of the vector of structural parameters $\boldsymbol{\theta}$

3.1 Myopic players

3.1.1 Model with no contemporaneous strategic interactions

Consider the myopic model (i.e., $\widetilde{V}_{1t} = \widetilde{V}_{2t} = 0$) under the condition that $\gamma_1 = \gamma_2 = 0$. The best response equations for this model are:

$$\left\{ \begin{array}{lcl} y_{1t} & = & 1\{\alpha_1 + \beta_1 \; y_{1t-1} + \lambda_1 \; y_{2t-1} - \varepsilon_{1t} \geq 0\} \\ \\ y_{2t} & = & 1\{\alpha_2 + \beta_2 \; y_{2t-1} + \lambda_2 \; y_{1t-1} - \varepsilon_{2t} \geq 0\} \end{array} \right. \tag{12}$$

This is an autoregressive bivariate logit model. Narendranthan, Nickell, and Metcalf (1985) consider this model in their study of the joint dynamics of unemployment and sickness. They present a proof for the identification of the parameters using the same conditional likelihood approach as in our paper.[3] Consequently, the identification of this model is a well-established

[3] See Honoré and Kyriazidou (2019) and Honoré and De Paula (2021) for their recent analysis of this model.

result in the literature. We include this result as it serves as a straightforward example for introducing notation and ensuring comprehensiveness.

The model implies the following expression for the probability of a market history:

$$\mathbb{P}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right) = \prod_{i=1}^{2} \prod_{t=1}^{T} \frac{\exp\{\ y_{it}\left[\alpha_i + \beta_i\ y_{it-1} + \lambda_i\ y_{jt-1}\right]\ \}}{1 + \exp\left\{\alpha_i + \beta_i\ y_{it-} + \lambda_i\ y_{jt-1}\right\}}\ p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) \tag{13}$$

where $p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right)$ represents the probability of the initial condition given $\boldsymbol{\alpha}$. The logit model possesses a crucial property that enables an additive separability of the log-likelihood with respect to the incidental parameters $\boldsymbol{\alpha}$ and the structural parameters $\boldsymbol{\theta}$. We now illustrate this separability and its significance in the identification results.

Define, for $i \in \{1,2\}$, function $\sigma_{\boldsymbol{\alpha} i}\left(y_1, y_2\right) \equiv -\ln[1 + \exp\left\{\alpha_i + \beta_i\ y_1 + \lambda_i\ y_2\right\}]$, and let $\sigma_{\boldsymbol{\alpha}}\left(y_1, y_2\right) \equiv \sigma_{\alpha 1}\left(y_1, y_2\right) + \sigma_{\boldsymbol{\alpha} 2}\left(y_1, y_2\right)$. Given a choice history $\widetilde{\mathbf{y}}$, define the statistics:

- $T_i^{(1)} \equiv \sum_{t=1}^{T} y_{it}$ is the number of times that player i chooses alternative 1.
- $T^{(y_1,y_2)} \equiv \sum_{t=1}^{T} 1\{(y_{1t}, y_{2t}) = (y_1, y_2)\}$ is the number of times the two players choose (y_1, y_2).
- $C^{(y_1,y_2)} \equiv \sum_{t=1}^{T} 1\{(y_{1t}, y_{2,t-1}) = (y_1, y_2)\}$ is the number of times that player 1 chooses alternative y_1 given that player 2 chose alternative y_2 at previous period.

Then, the logarithm of the probability of a market history can be written as:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) &= \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} \\
&+ \sum_{y_1, y_2} \sigma_{\boldsymbol{\alpha}}\left(y_1, y_2\right)\ \left[T^{(y_1,y_2)} + 1\{(y_{10}, y_{20}) = (y_1, y_2)\} - 1\{(y_{1T}, y_{2T}) = (y_1, y_2)\}\right] \\
&+ \beta_1\ C_{11} + \lambda_1\ C_{12} + \lambda_2\ C_{21} + \beta_2\ C_{22}
\end{aligned} \tag{14}$$

Or using a more compact representation:

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) = \mathbf{s}\left(\widetilde{\mathbf{y}}\right)'\ \mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}\left(\widetilde{\mathbf{y}}\right)'\ \boldsymbol{\theta} \tag{15}$$

where $\mathbf{s}\left(\widetilde{\mathbf{y}}\right)$ and $\mathbf{c}\left(\widetilde{\mathbf{y}}\right)$ are vectors of statistics, $\mathbf{g}_{\boldsymbol{\alpha}}$ is a vector of functions of the incidental parameters, and $\boldsymbol{\theta}$ is the vector of structural parameters $(\beta_1, \lambda_1, \lambda_2, \beta_2)'$. More specifically,[4]

$$\left\{ \begin{array}{rcllll} \mathbf{s}(\widetilde{\mathbf{y}}) & = & [1,\ y_{10},\ y_{20},\ y_{10}y_{20} & ;\ 1, y_{1T}, y_{2T}, y_{1T}y_{2T} & ;\ T,\ T_1^{(1)},\ T_2^{(1)},\ T^{(1,1)}\Big]' \\ \mathbf{g}_{\boldsymbol{\alpha}} & = & [\ln \mathbf{p}^*_{\boldsymbol{\alpha}} + \sigma^*_{\boldsymbol{\alpha}} & ;\ -\sigma^*_{\boldsymbol{\alpha}} & ;\ \sigma^*_{\boldsymbol{\alpha}} + (0, \alpha_1, \alpha_2, 0)]' \\ \mathbf{c}(\widetilde{\mathbf{y}}) & = & [C_{11},\ C_{12},\ C_{21},\ C_{22}]' & & \\ \boldsymbol{\theta} & = & [\beta_1,\ \lambda_1,\ \lambda_2,\ \beta_2]' & & \end{array} \right. \tag{16}$$

Given (15) and (16) we can establish the following identification result.

PROPOSITION 1. In the myopic dynamic game without contemporaneous interactions in equation (12), the structural parameters β_1, β_2, λ_1, and λ_2 are point identified when $T \geq 3$. ■

Proof. Given that, $\mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \boldsymbol{\alpha}, \boldsymbol{\theta}) = \frac{\mathbb{P}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})}{\mathbb{P}(\mathbf{s}(\widetilde{\mathbf{y}}) \mid \boldsymbol{\alpha}, \boldsymbol{\theta})}$, and using the additive structure in equation (15), we have that:

$$\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \beta) \quad = \quad \mathbf{c}(\widetilde{\mathbf{y}})' \boldsymbol{\theta} - \ln \left(\textstyle\sum_{\widetilde{\mathbf{y}}' : \mathbf{s}(\widetilde{\mathbf{y}}') = \mathbf{s}(\widetilde{\mathbf{y}})} \exp \{ \mathbf{c}(\widetilde{\mathbf{y}}')' \boldsymbol{\theta}) \} \right) \tag{17}$$

This equation implies that vector $\mathbf{s}(\widetilde{\mathbf{y}})$ is a sufficient statistic for $\boldsymbol{\alpha}$. Furthermore, there are pairs of market histories, say A and B, with $\mathbf{s}(A) = \mathbf{s}(B)$ and $\mathbf{c}(A) \neq \mathbf{c}(B)$ that identify the structural parameters of the model.

Suppose that $T = 3$, let $\boldsymbol{y}_t \equiv (y_{1t}, y_{2t})$, and consider the following pair of histories: $A = \{\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\}$ and $B = \{\boldsymbol{y}_0, \boldsymbol{b}, \boldsymbol{a}, \boldsymbol{y}_3\}$. We first verify that histories A and B have the same sufficient statistic $\mathbf{s}$. It is clear that the two histories have the same initial condition $\boldsymbol{y}_0$, and last period choices, $\boldsymbol{y}_3$. And it is also clear that the frequency of choices in $\{\boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\}$ is the same as in $\{\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{y}_3\}$ such that $T^{(y_1,y_2)}(A) = T^{(y_1,y_2)}(B)$ for any pair $(y_1, y_2) \in \{0,1\}^2$. Therefore, $\mathbf{s}(A) = \mathbf{s}(B)$. Now, for $\boldsymbol{a} \neq \boldsymbol{b}$ we have that $\mathbf{c}(A) \neq \mathbf{c}(B)$ and the difference between the log-probabilities of these histories identifies parameters of interest. Note that,

$$\left\{ \begin{array}{rcl} C_{11}(A) - C_{11}(B) & = & (a_1 - b_1)\ \ (y_{10} - y_{13}) \\ C_{12}(A) - C_{12}(B) & = & (a_1 - b_1) y_{20} - (a_2 - b_2) y_{13} + a_2 b_1 - a_1 b_2 \\ C_{21}(A) - C_{21}(B) & = & (a_2 - b_2) y_{10} - (a_1 - b_1) y_{23} + a_1 b_2 - a_2 b_1 \\ C_{22}(A) - C_{22}(B) & = & (a_2 - b_2)\ \ (y_{20} - y_{23}) \end{array} \right. \tag{18}$$

Using the expressions in (18), Table 2 presents four pairs of histories, with each pair identifying one of the structural parameters. The corresponding parameter that is identified by $\ln \mathbb{P}(A) -$

[4]The functions of incidental parameters $\ln \mathbf{p}^*_{\boldsymbol{\alpha}}$ and $\sigma^*_{\boldsymbol{\alpha}}$ have the following definition $\ln \mathbf{p}^*_{\boldsymbol{\alpha}} \equiv [\ln p_{\alpha}(0,0)$, $\ln p_{\boldsymbol{\alpha}}(1,0) - \ln p_{\boldsymbol{\alpha}}(0,0)$, $\ln p_{\boldsymbol{\alpha}}(0,1) - \ln p_{\boldsymbol{\alpha}}(0,0)$, $\ln p_{\alpha}(1,1) - \ln p_{\boldsymbol{\alpha}}(0,1) - \ln p_{\boldsymbol{\alpha}}(1,0) + \ln p_{\boldsymbol{\alpha}}(0,0)]$, and $\sigma^*_{\boldsymbol{\alpha}} \equiv [\sigma_{\boldsymbol{\alpha}}(0,0)$, $\sigma_{\boldsymbol{\alpha}}(1,0) - \sigma_{\boldsymbol{\alpha}}(0,0)$, $\sigma_{\boldsymbol{\alpha}}(0,1) - \sigma_{\boldsymbol{\alpha}}(0,0)$, $\sigma_{\boldsymbol{\alpha}}(1,1) - \sigma_{\boldsymbol{\alpha}}(0,1) - \sigma_{\boldsymbol{\alpha}}(1,0) + \sigma_{\boldsymbol{\alpha}}(0,0)]$.

$\ln \mathbb{P}(B)$. In cases 1 and 2, we identify the parameter β_i by keeping constant the choice of the other player – $j \neq i$ – and comparing the frequency of the history where player i "switches" – $(0,1,0,1)$ – with the frequency of the history where she "stays" – $(0,0,1,1)$. In cases 3 and 4, we compare the probability of history $(0,0,0,1)$ for player i when the other player chooses alternative 1 at period $t = 2$ – $(0,0,1,0)$ – and when this choice is at period $t = 1$ – $(0,1,0,0)$. There are other values for $\boldsymbol{y}_0$, $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{y}_3$ that identify linear combinations of the several parameters in $\boldsymbol{\theta}$. ■

Table 2
Myopic dynamic game without contemporaneous effects
Examples of histories and identified parameters with T=3
$A = \{\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\}$; $B = \{\boldsymbol{y}_0,\ \boldsymbol{b},\ \boldsymbol{a},\ \boldsymbol{y}_3\}$

	$\boldsymbol{y}_0$	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{y}_3$	$\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$
Case 1:	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	β_1
Case 2:	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	β_2
Case 3:	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	λ_1
Case 4:	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	λ_2

3.1.2 Myopic players with one-direction strategic interactions

Now, we relax the condition of no contemporaneous strategic interactions and allow γ_2 to be different to zero: there is a contemporaneous effect of y_1 on y_2. We still keep the restriction $\gamma_1 = 0$ – no contemporaneous effect of y_2 on y_1, and include the restriction $\lambda_2 = 0$. That is, the model is defined by the following best response functions:

$$\begin{cases} y_{1t} & = & 1\{\alpha_1 + \beta_1 \; y_{1t-1} + \lambda_1 \; y_{2t-1} - \varepsilon_{1t} \geq 0\} \\ \\ y_{2t} & = & 1\{\alpha_2 + \gamma_2 \; y_{1t} + \beta_2 \; y_{2t-1} - \varepsilon_{2t} \geq 0\} \end{cases} \tag{19}$$

The log-probability of the market history $\widetilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, .., T)$ has the following structure:

$$\begin{aligned} \ln \mathbb{P}(\widetilde{\mathbf{y}}|\boldsymbol{\alpha}, \boldsymbol{\theta}) &= \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) + \alpha_1 \; T_1^{(1)} + \alpha_2 \; T_2^{(1)} + \sum_{t=1}^{T} \sigma_{\boldsymbol{\alpha}1}(y_{1t-1}, y_{2t-1}) + \sigma_{\boldsymbol{\alpha}2}(y_{1t}, y_{2t-1}) \\ \\ &+ \beta_1 \; C_{11} + \lambda_1 \; C_{12} + \beta_2 \; C_{22} + \gamma_2 \; T^{(1,1)} \end{aligned} \tag{20}$$

with $\sigma_{\boldsymbol{\alpha}1}(y_1, y_2) \equiv -\ln[1+\exp\{\alpha_1+\beta_1 y_1+\lambda_1 y_2\}]$ and $\sigma_{\boldsymbol{\alpha}2}(y_1, y_2) \equiv -\ln[1+\exp\{\alpha_2+\gamma_2 y_1+\beta_2 y_2\}]$. By comparing equations (20) and (14), we can find two important differences. Firstly, equation (20) includes the term $\gamma_2 T^{(1,1)}$, which is absent in (14). Secondly, in equation (20), the term that depends on incidental parameters includes not only the sum $\sum_{t=1}^{T} \sigma_{\boldsymbol{\alpha}1}(y_{1,t-1}, y_{2t-1})$ – which is also present in equation (14) – but also the sum $\sum_{t=1}^{T} \sigma_{\boldsymbol{\alpha}2}(y_{1t}, y_{2t-1})$, which did not appear in equation (14). These differences have implications for parameter identification.

Similarly as for the previous model, we can rewrite the right hand side of equation (20) as $\mathbf{s}(\widetilde{\mathbf{y}})' \; \mathbf{g}_\alpha + \mathbf{c}(\widetilde{\mathbf{y}})' \; \boldsymbol{\theta}^*$, but now the vectors of statistics $\mathbf{s}(\widetilde{\mathbf{y}})$ and $\mathbf{c}(\widetilde{\mathbf{y}})$, and the vector of identified parameters $\boldsymbol{\theta}^*$ are different. More specifically,[5]

$$\begin{cases} \mathbf{s}(\widetilde{\mathbf{y}}) & = & [1, y_{10}, y_{20}, y_{10}y_{20} \quad ; \quad 1, y_{1T}, y_{2T}, y_{1T}y_{2T} \quad ; \quad T, \; , T_1^{(1)}, \; T_2^{(1)}, \; T^{(1,1)} \quad ; \quad C_{12}]' \\ \mathbf{g}_{\boldsymbol{\alpha}} & = & \left[\ln \mathbf{p}_{\boldsymbol{\alpha}}^* + \sigma_{\boldsymbol{\alpha}}^* \quad ; \quad -\sigma_{\boldsymbol{\alpha}}^* \; ; \quad ; \quad \begin{matrix} \sigma_{\boldsymbol{\alpha}}^* + \\ (0, \alpha_1, \alpha_2, -\Delta\sigma_{\alpha 2} + \gamma_2) \end{matrix} \quad ; \quad \begin{matrix} \Delta\sigma_{\alpha 2} \\ +\lambda_1 \end{matrix} \right]' \\ \mathbf{c}(\widetilde{\mathbf{y}}) & = & [C_{11}, \; C_{22}]' \\ \boldsymbol{\theta}^* & = & [\beta_1, \; \beta_2]' \end{cases} \tag{21}$$

where $\ln \mathbf{p}_{\boldsymbol{\alpha}}^*$ and $\sigma_{\boldsymbol{\alpha}}^*$ have the same definition as in section 2.2 above, and $\Delta\sigma_{\alpha i}$ is the incidental parameter $\sigma_{\alpha i}(1,1) - \sigma_{\alpha i}(0,1) - \sigma_{\alpha i}(1,0) + \sigma_{\alpha i}(0,0)$.

There are some fundamental differences with respect to the model without contemporaneous

[5] For this derivation, it is helpful to write $\sum_{t=1}^{T} \sigma_{\boldsymbol{\alpha}2}(y_{1t}, y_{2t-1})$ as $\left[\sum_{t=1}^{T}(1-y_{1t})(1-y_{2t-1})\right] \; \sigma_{\boldsymbol{\alpha}2}(0,0)+$ $\left[\sum_{t=1}^{T} y_{1t}(1-y_{2t-1})\right] \; \sigma_{\boldsymbol{\alpha}2}(1,0)+ \; \left[\sum_{t=1}^{T}(1-y_{1t})y_{2t-1}\right] \; \sigma_{\boldsymbol{\alpha}2}(0,1)+ \; \left[\sum_{t=1}^{T} y_{1t}y_{2t-1}\right] \; \sigma_{\boldsymbol{\alpha}2}(1,1)$. Note that this expression is equal to $T \; \sigma_{\boldsymbol{\alpha}2}(0,0) + T_1^{(1)} \; [\sigma_{\boldsymbol{\alpha}2}(1,0) - \sigma_{\boldsymbol{\alpha}2}(0,0)] + \left[T_2^{(1)} + y_{20} - y_{2T}\right] \; [\sigma_{\alpha 2}(0,1) - \sigma_{\boldsymbol{\alpha}2}(0,0)] + C_{12}$ $[\sigma_{\boldsymbol{\alpha}2}(1,1) - \sigma_{\boldsymbol{\alpha}2}(1,0) - \sigma_{\boldsymbol{\alpha}2}(0,1) + \sigma_{\boldsymbol{\alpha}2}(0,0)]$.

strategic interactions. First, the statistic C_{12} and the structural parameter λ_1 appear in the log-probability of a choice history through the term C_{12} $(\Delta\sigma_{\alpha 2}+\lambda_1)$. Without further restrictions we have that the incidental parameter $\Delta\sigma_{\alpha 2}$ is not zero. This implies that this sufficient statistics approach cannot identify parameter λ_1. Second, the statistic $T^{(1,1)}$ and the structural parameter γ_2 appear through the term $T^{(1,1)}$ $(\Delta\sigma_{\alpha 1} + \gamma_2)$. Without further restrictions, the sufficient statistics approach does not identify parameter γ_2.

Proposition 2 establishes the point identification of dynamic parameters β_1 and β_2 in this model.

PROPOSITION 2. In the myopic dynamic game with one-direction contemporaneous interactions as in equation (19), the parameters β_1 and β_2 are point identified when $T \geq 3$. ■

Proof. Consider the same framework as in the proof of Proposition 1: $T = 3$ and the pair of histories $A = \{\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\}$ and $B = \{\boldsymbol{y}_0, \boldsymbol{b}, \boldsymbol{a}, \boldsymbol{y}_3\}$. In the proof of Proposition 1, we showed that these histories have the same value for the statistics $\boldsymbol{y}_0$, $\boldsymbol{y}_3$, and $T^{(y_1,y_2)}$. Now, in this model with a contemporaneous effect, the sufficient statistic includes C_{12}, so we need to impose additional conditions on histories A and B such that $C_{12}(A) = C_{12}(B)$. In the histories in Table 2, we have that $C_{12}(A) = C_{12}(B)$ for cases 1 and 2. Therefore, these two pairs of market histories still identify the parameters β_1 and β_2, respectively, in this dynamic game. We present this result in Table 3. ■

Table 3
Myopic Game with One-Direction Strategic Interactions
Examples of histories and identified parameters with T=3
$A = \{\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\}$; $B = \{\boldsymbol{y}_0,\ \boldsymbol{b},\ \boldsymbol{a},\ \boldsymbol{y}_3\}$ with $C_{12}(A) = C_{12}(B)$

	$\boldsymbol{y}_0$	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{y}_3$	$\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$
Case 1:	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	β_1
Case 2:	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	β_2

As explained above, the no identification of the parameters γ_2 and λ_1 is due to the fact that they appear in the log-probability of a choice history only through the terms $T^{(1,1)}$ $(\Delta\sigma_{\alpha 1} + \gamma_2)$

and C_{12} $(\Delta\sigma_{\alpha 2} + \lambda_1)$, respectively. This feature of the model also provides conditions for the identification of these parameters. The parameter γ_2 is identified if and only if $\Delta\sigma_{\alpha 1}$ is equal to zero (or a constant) for any possible value of the incidental parameter α_1. Remember that $\Delta\sigma_{\alpha 1}$ is defined as $\sigma_{\alpha 1}(1,1) - \sigma_{\alpha 1}(0,1) - \sigma_{\alpha 1}(1,0) + \sigma_{\alpha 1}(0,0)$, and $\sigma_{\alpha 1}(y_1, y_2)$ is defined as $-\ln[1 + \exp\{\alpha_1 + \beta_1\ y_1 + \lambda_1\ y_2\}]$. Taking this into account we have that $\Delta\sigma_{\alpha 1} = 0$ if and only if $\beta_1 = 0$ or/and $\lambda_1 = 0$. Following a similar argument, we have the parameter λ_1 is identified if and only if $\Delta\sigma_{\alpha 2} = 0$ for any value of α_2, and this is the case if and only if $\gamma_2 = 0$ or/and $\beta_2 = 0$. We summarize these identification results in the following Proposition 3.

PROPOSITION 3. Consider the myopic dynamic game with one-direction strategic interactions described in equation (19). Using a sufficient-statistic approach: (A) a necessary and sufficient condition for the identification of parameter γ_2 is that $\beta_1 = 0$ or/and $\lambda_1 = 0$; (B) similarly, a necessary and sufficient condition for the identification of parameter λ_1 is that $\gamma_2 = 0$ or/and $\beta_2 = 0$. ■

Functional differencing approach

So far, our investigation has focused on identification results using a *conditional likelihood–sufficient statistics* approach. However, recent studies by Honoré and Weidner (2020) and Dobronyi, Gu, and Kim (2021) have employed a functional differencing approach inspired by Bonhomme (2012) to establish parameter identification in dynamic logit models that cannot be identified using the conditional likelihood method. We now adopt the same approach, specifically following the methodology outlined in Dobronyi, Gu, and Kim (2021). Our particular interest lies in identifying the parameters γ_2 and λ_1 without imposing the restrictions stated in Proposition 3.

The functional differencing approach, when employed without any further constraints, falls short of achieving (point) identification for the parameters γ_2 and λ_1. However, by introducing the condition that the fixed effects α_{1m} and α_{2m} are identical for both players, the functional differencing approach successfully resolves the identification problem for these parameters. It is important to note that the conditional likelihood approach, even with this additional restriction, does not lead to the identification of γ_2 and λ_1, as evidenced by Proposition 4.

PROPOSITION 4. Consider the myopic dynamic game with one-direction strategic interactions as described in equation (19) where the fixed effects of the two players are restricted to be the same: $\alpha_{1m} = \alpha_{2m}$. The functional differenting approach implies moment conditions that point identify all the structural parameters, β_1, β_2, λ_1, and γ_2. ■

Proof: In Appendix A.1.

3.1.3 Myopic players with two-direction strategic interactions

Consider the game with two-direction contemporaneous interactions such that $\gamma_1 \neq 0$ $\gamma_2 \neq 0$. In this model, we eliminate the lagged strategic interactions between players such that $\lambda_1 = \lambda_2 = 0$.

$$\begin{cases} y_{1t} & = & 1\{\alpha_1 + \gamma_1 \ y_{2t} + \beta_1 \ y_{1t-1} - \varepsilon_{1t} \geq 0\} \\ \\ y_{2t} & = & 1\{\alpha_2 + \gamma_2 \ y_{1t} + \beta_2 \ y_{2t-1} - \varepsilon_{2t} \geq 0\} \end{cases} \tag{22}$$

For the rest of this section, we assume that the researcher knows the sign of parameters γ_1 and γ_2. For concreteness, we consider that $\gamma_1 \leq 0$ and $\gamma_2 \leq 0$.

Two versions of the model are distinguished based on whether players move sequentially or simultaneously. The difference between the sequential and the simultaneous move games is in the set of equilibria. In the simultaneous move game, there is a quadrangle in the space of $(\varepsilon_{1t}, \varepsilon_{2t})$ for which outcomes $(0,1)$ and $(1,0)$ are Nash equilibria. This quadrangle is:

$$\{e_{1t}^{L} \ < \varepsilon_{1t} \leq e_{1t}^{U} \qquad \& \qquad e_{2t}^{L} \ < \varepsilon_{2t} \leq e_{2t}^{U}\} \tag{23}$$

where $e_{it}^{L} \equiv \alpha_i + \gamma_i + \beta_i y_{i,t-1}$, and $e_{1t}^{U} \equiv \alpha_i + \beta_i y_{i,t-1}$. In the sequential move game, where player 1 moves first, the outcome $(1,0)$ is the unique equilibrium (Subgame Perfect Nash equilibrium) associated with this region of $(\varepsilon_{1t}, \varepsilon_{2t})$. Player 1 is aware that if she chooses $y_{1t} = 1$, player 2 will choose $y_{2t} = 0$, and if she chooses $y_{1t} = 0$, then player 2 will choose $y_{2t} = 1$. Therefore, player 1's decision determines which of the two Nash equilibria, $(0,1)$ or $(1,0)$, is selected. Player 1 selects the equilibrium that maximizes its profit. Considering $\gamma_1 \leq 0$, the Nash equilibrium with the highest profit for player 1 is $(1,0)$.

A. Sequential move

Suppose that player 1 moves first. The probability for outcomes $(0,0)$, $(1,0)$, and $(1,1)$ can be represented using the following product of logit probabilities:

$$\mathbb{P}(y_{1t}, y_{2t}; \boldsymbol{\alpha}) = \Lambda([2y_{1t} - 1][\alpha_1 + \gamma_1 y_{2t} + \beta_1 y_{1t-1}]) \ \Lambda([2y_{2t} - 1][\alpha_2 + \gamma_2 y_{1t} + \beta_2 y_{2t-1}]) \tag{24}$$

In contrast, the probability of outcome $(0,1)$ cannot be represented as the product of logits. This has implications for the derivation of a sufficient statistic for $\boldsymbol{\alpha}$.

Let $\widetilde{\mathbf{y}}$ be a choice history where every period's outcome is an element of $\{(0,0),(1,0),(1,1)\}$, i.e., it does not include outcome $(0,0)$. For this sequential move game, the log-probability of this choice history has the following structure:

$$\begin{aligned}\ln \mathbb{P}(\widetilde{\mathbf{y}}|\boldsymbol{\alpha},\boldsymbol{\theta}) &= \ln p_\alpha(y_{10},y_{20}) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} + \sum_{t=1}^{T} \sigma_{\alpha_1}(y_{1t-1},y_{2t}) + \sigma_{\alpha_2}(y_{1t},y_{2t-1}) \\ &+ \beta_1\ C_{11} + \beta_2\ C_{22} + (\gamma_1+\gamma_2)\ T^{(1,1)}\end{aligned} \tag{25}$$

with $\sigma_{\alpha_i}(y_{i,t-1},y_{jt}) = -\ln[1+\exp\{\alpha_i+\beta_i y_{i,t-1}+\gamma_i y_{jt}\}]$. Similarly as for the previous models, we can rewrite the right hand side of equation (25) as $\mathbf{s}(\widetilde{\mathbf{y}})'\ \mathbf{g}_\alpha + \mathbf{c}(\widetilde{\mathbf{y}})'\ \boldsymbol{\theta}^*$, where now the vectors of statistics and parameters have the following form:

$$\begin{aligned}\ln \mathbb{P}(\widetilde{\mathbf{y}}|\boldsymbol{\alpha},\boldsymbol{\theta}) &= \left[1, y_{10}, y_{20}, y_{10}y_{20},\ y_{1T}, y_{2T}, y_{1T}y_{2T},\ T, T_1^{(1)}, T_2^{(1)},\ C_{12}, C_{21}\right]'\ \mathbf{g}_{\boldsymbol{\alpha}} \\ &+ \beta_1\ C_{11} + \beta_2\ C_{22} + (\gamma_1+\gamma_2)\ T^{(1,1)}\end{aligned} \tag{26}$$

A preliminary examination of equation (26) might suggest that the parameter $\gamma_1+\gamma_2$ is identified, as it appears alongside the statistic $T^{(1,1)}$, which is not included in the vector of sufficient statistics $\mathbf{s}(\widetilde{\mathbf{y}})$. However, due to the nature of the choice history $\widetilde{\mathbf{y}}$, where $(1-y_{1t})y_{2t}=0$ for every t, it follows that $\sum_t y_{2t} = \sum_t y_{1t}y_{2t}$, implying that $T_2^{(1)} = T^{(1,1)}$. Consequently, given $\mathbf{s}(\widetilde{\mathbf{y}})$, the statistic $T^{(1,1)}$ lacks variation, thereby rendering this approach of using sufficient statistics insufficient for identifying $\gamma_1+\gamma_2$. Nonetheless, parameters β_1 and β_2 can still be identified. This assertion is formalized in the following proposition.

PROPOSITION 5. In the myopic dynamic game without two-direction interactions and sequential move, the structural parameters β_1 and β_2 are point identified when $T \geq 3$. ■

Proof. The identification of β_1 can be established by considering two distinct choice histories (y_0,y_1,y_2,y_3) with $T=3$. Let $A = \begin{pmatrix}0&0&1&1\\0&0&0&0\end{pmatrix}$ and $B = \begin{pmatrix}0&1&0&1\\0&0&0&0\end{pmatrix}$. It can be readily verified that A and B yield identical values for the sufficient statistics y_0, y_3, $T_1^{(1)}$, $T_2^{(1)}$, C_{21}, and C_{12}. Additionally, they share the same value for the statistic C_{22}. However, while $C_{11}(A)=1$, we have $C_{11}(B)=0$. This indicates that statistic $\ln P(A) - \ln P(B)$ identifies β_1.

To identify β_2, let's consider two choice histories: $A = \begin{pmatrix}1&1&1&1\\0&0&1&1\end{pmatrix}$ and $B = \begin{pmatrix}1&1&1&1\\0&1&0&1\end{pmatrix}$. In these histories, the values of the sufficient statistics y_0, y_3, $T_1^{(1)}$, $T_2^{(1)}$, C_{21}, and C_{12} are identical, as well as the value of C_{11}. However, it is noteworthy that $C_{22}(A)=0$ while $C_{22}(B)=1$. Consequently, statistic $\ln P(A) - \ln P(B)$ identifies β_2. ■

Partial Identification of γ_1, γ_2

For $(y_{1t}, y_{2t}) = (0, 1)$, we can consider the following logit form lower bound:

$$\ln \mathbb{P}((0,1)|\alpha, \boldsymbol{\theta}) \geq \sigma_{\alpha_1}(y_{1t-1}, 1) + \sigma_{\alpha_2}(1, y_{2t-1}) + \alpha_2 + \beta_2 y_{2t-1} + \gamma_2$$

or

$$\ln \mathbb{P}((0,1)|\alpha, \boldsymbol{\theta}) \geq \sigma_{\alpha_1}(y_{1t-1}, 0) + \sigma_{\alpha_2}(0, y_{2t-1}) + \alpha_2 + \beta_2 y_{2t-1}$$

and the following logit form upper bound:

$$\ln \mathbb{P}((0,1)|\alpha, \boldsymbol{\theta}) \leq \sigma_{\alpha_1}(y_{1t-1}, 1) + \sigma_{\alpha_2}(0, y_{2t-1}) + \alpha_2 + \beta_2 y_{2t-1}$$

This gives us access to use Proposition 8 in the paper to partially identify γ_1, γ_2.

3.1.4 Sharp Identified set

Since the model is complete, the sharp identified set can be written as the collection of $\boldsymbol{\theta} = (\beta_1, \beta_2, \gamma_1, \gamma_2)$ such that for each value of (y_{10}, y_{20}), there exists a distribution G (allowed to vary over (y_{10}, y_{20})) such that, for all $\widetilde{\mathbf{y}} \in \{0,1\}^T$,

$$\mathbb{P}(\widetilde{\mathbf{y}}|y_{10}, y_{20}) = \int \mathbb{L}(\widetilde{\mathbf{y}}|\alpha_1, \alpha_2, \boldsymbol{\theta}) dG(\alpha_1, \alpha_2|y_{10}, y_{20})$$

where $\mathbb{L}$ is the likelihood function given $\alpha_1, \alpha_2, \boldsymbol{\theta}$. Since model is complete, with the logit distribution assumption, we have a likelihood funciton for given values of $(\alpha_1, \alpha_2, \boldsymbol{\theta})$. If we are willing to take a fixed group of (α_1, α_2), we can use linear program to numerically compute the identified set for $\boldsymbol{\theta}$. The approach taken in Dobroyni, Gu and Kim (2021) can in principle be used to derive all moment equality conditions available from the model for $\boldsymbol{\theta}$. For example, we can write the model as

$$\mathcal{P} = H(\boldsymbol{\theta})\tilde{m}$$

where $\mathcal{P}$ is the 2^T choice probability vector and $H(\boldsymbol{\theta})$ is a matrix that only involves parameters, and $\tilde{m}$ are a vector of moments of $(A_1, A_2) := (\exp(\alpha_1), \exp(\alpha_2))$ (i.e. entries of $\tilde{m_{\boldsymbol{\theta}}}$ takes the form $\int A_1^j A_2^k dG(A_1, A_2, \boldsymbol{\theta})$ for some measure of G). The left null space of $H(\boldsymbol{\theta})$ provides all moment equality conditions available in the model for $\boldsymbol{\theta}$.

B. Simultaneous move: Conditional likelihood - Bounds approach

Here we concentrate on the (point) identification of the switching cost parameters – β_1 and β_2 – and on the partial identification of all the parameters.

The following Lemma 1 presents a property that plays a key role in our sufficient statistics - bounds approach.

LEMMA 1. Suppose that the log-probability of a market history has lower and upper bounds with the following structure: the lower bound is $\ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbf{s}_L(\widetilde{\mathbf{y}})' \mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}_L(\widetilde{\mathbf{y}})' \boldsymbol{\theta}$ *and the upper bound is* $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbf{s}_U(\widetilde{\mathbf{y}})' \mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}_U(\widetilde{\mathbf{y}})' \boldsymbol{\theta}$*, where* $\mathbf{s}_L(\widetilde{\mathbf{y}})$*,* $\mathbf{s}_U(\widetilde{\mathbf{y}})$*,* $\mathbf{c}_L(\widetilde{\mathbf{y}})$*, and* $\mathbf{c}_U(\widetilde{\mathbf{y}})$ *are vectors of statistics, and* $\mathbf{g}_{\boldsymbol{\alpha}}$ *is a vector of incidental parameters. Given this structure, the logarithm of the probability of a market history* $\widetilde{\mathbf{y}}$ *(unconditional on* $\boldsymbol{\alpha}$*) has the following bounds:*

$$h(\mathbf{s}_L(\widetilde{\mathbf{y}})) + \mathbf{c}_L(\widetilde{\mathbf{y}})' \beta \ \leq \ \ln \mathbb{P}(\widetilde{\mathbf{y}}) \ \leq \ h(\mathbf{s}_U(\widetilde{\mathbf{y}})) + \mathbf{c}_U(\widetilde{\mathbf{y}})' \boldsymbol{\theta} \tag{27}$$

where $h(\mathbf{s})$ *is a function (described in the Appendix) that depends on the vector of statistics* $\mathbf{s}$ *and on the probability distribution of the incidental parameters* $\boldsymbol{\alpha}$*. Given two different histories, say* A *and* B*.*

(i) If $\mathbf{s}_L(A) = \mathbf{s}_U(B)$ *and* $\mathbf{c}_L(A) \neq \mathbf{c}_U(B)$*, we have that:*

$$[\mathbf{c}_L(A) - \mathbf{c}_U(B)]' \boldsymbol{\theta} \ \leq \ \ln \mathbb{P}(A) - \ln \mathbb{P}(B) \tag{28}$$

(ii) If $\mathbf{s}_U(A) = \mathbf{s}_L(B)$ *and* $\mathbf{c}_U(A) \neq \mathbf{c}_L(B)$*, we have that:*

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \ \leq \ [\mathbf{c}_U(A) - \mathbf{c}_L(B)]' \boldsymbol{\theta} \tag{29}$$

These inequalities imply partial identification of some structural parameters. ■

Lemma 1 does not imply that $\mathbf{s}_L(\widetilde{\mathbf{y}})$ or $\mathbf{s}_U(\widetilde{\mathbf{y}})$ – or even the union of these two vectors of statistics – are sufficient statistics for the incidental parameters in the probability $\mathbb{P}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$. In general, this is not true for this model. However, the vectors $\mathbf{s}_L(\widetilde{\mathbf{y}})$ and $\mathbf{s}_U(\widetilde{\mathbf{y}})$ are sufficient statistics for the the incidental parameters in the lower and in the upper bounds of this probability, respectively. This property – together with the condition that there are histories with $\mathbf{s}_L(A) = \mathbf{s}_U(B)$ and with $\mathbf{c}_L(A) \neq \mathbf{c}_U(B)$ – allow us to obtain partial identification of the structural parameters.

The rest of this section describes the derivation of the expressions for the bounds, $\ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbf{s}_L(\widetilde{\mathbf{y}})' \mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}_L(\widetilde{\mathbf{y}})' \boldsymbol{\theta}$ and $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbf{s}_U(\widetilde{\mathbf{y}})' \mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}_U(\widetilde{\mathbf{y}})' \beta$, and our (set) identification results.

Given a market history $\widetilde{\mathbf{y}}$, we can construct a lower bound and an upper bound for the log-probability of this history $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta})$. These bounds are:

$$\begin{cases} \ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) & \equiv & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) + \sum_{t=1}^{T} \ln L(\mathbf{y}_t|\mathbf{y}_{t-1}; \boldsymbol{\alpha}, \boldsymbol{\theta}) \\ \\ \ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) & \equiv & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) + \sum_{t=1}^{T} \ln U(\mathbf{y}_t|\mathbf{y}_{t-1}; \boldsymbol{\alpha}, \boldsymbol{\theta}) \end{cases} \tag{30}$$

For outcomes $(0,0)$ and $(1,1)$, the upper bounds and the lower bounds are the same and they are the probabilities in equation (8). For outcomes $(0,1)$ and $(1,0)$, the upper bounds $U(\mathbf{y}_t|\mathbf{y}_{t-1}; \boldsymbol{\alpha}, \beta)$ are the ones in equation (9), and the lower bounds $L(\mathbf{y}_t|\mathbf{y}_{t-1}; \boldsymbol{\alpha}, \boldsymbol{\theta})$ come from equations (10) and (11).

Lemma 2 presents bounds for the log-probability of a market history in our model, shows that these bounds have the structure in Lemma 1, and provides the specific form of the vectors of statistics $\mathbf{s}_L$, $\mathbf{s}_U$, $\mathbf{c}_U$, and $\mathbf{c}_U$.

LEMMA 2. For the myopic complete information dynamic game with contemporaneous effects in equation (22), the log-probability of a market history has lower bounds $\ln \mathbb{P}_{L\{E,W\}}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta})$ *and* $\ln \mathbb{P}_{L\{S,N\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$ *and upper bound* $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$ *which have the following expressions:*

$$\begin{array}{rcl} \ln \mathbb{P}_{L\{E,W\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) & = & \mathbf{s}^1(\widetilde{\mathbf{y}})' \, \mathbf{g}^1_{\boldsymbol{\alpha}} + \left[T_1^{(1)}, \, T_1^{(1)}, \, C_{11}, \, C_{12} \right] \; \mathbf{g}^2_{\boldsymbol{\alpha}} \\ & + & C_{11} \; \beta_1 + C_{22} \; \beta_2 + T_1^{(1)} \; \gamma_1 + T^{(1,1)} \; \gamma_2 \\ \\ \ln \mathbb{P}_{L\{S,N\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) & = & \mathbf{s}^1(\widetilde{\mathbf{y}})' \, \mathbf{g}^1_{\boldsymbol{\alpha}} + \left[T_2^{(1)}, \, T_2^{(1)}, \, C_{21}, \, C_{22} \right] \; \mathbf{g}^2_{\boldsymbol{\alpha}} \\ & + & C_{11} \; \beta_1 + C_{22} \; \beta_2 + T^{(1,1)} \; \gamma_1 + T_2^{(1)} \; \gamma_2 \\ \\ \ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) & = & \mathbf{s}^1(\widetilde{\mathbf{y}})' \, \mathbf{g}^1_{\boldsymbol{\alpha}} + \left[T_2^{(1)}, \, T_1^{(1)}, \, C_{21}, \, C_{12} \right] \; \mathbf{g}^2_{\boldsymbol{\alpha}} \\ & + & C_{11} \; \beta_1 + C_{22} \; \beta_2 + T^{(1,1)} \; [\gamma_1 + \gamma_2] \end{array} \tag{31}$$

where $\mathbf{g}^1_{\boldsymbol{\alpha}}$ *and* $\mathbf{g}^2_{\boldsymbol{\alpha}}$ *are vectors of incidental parameters which are defined in the Appendix, and the vector of statistics* $\mathbf{s}^1(\widetilde{\mathbf{y}})$ *consists of* T, y_{10}, y_{20}, y_{1T}, y_{2T}, $T_1^{(1)}$, *and* $T_2^{(1)}$. ■

Combining the general identification approach in Lemma 1 with the specific expressions for the bounds in Lemma 2, we can obtain the following identification results in Proposition 6.

PROPOSITION 6. Consider the myopic complete information dynamic game with contemporaneous effects in equation (22). Define the vector of statistics $\mathbf{s}^1(\widetilde{\mathbf{y}}) \equiv [T, \, y_{10}, \, y_{20}, \, y_{1T}, \, y_{2T}, \, T_1^{(1)}, \, T_2^{(1)}]$. *Let A and B be two market histories such that* $\mathbf{s}^1(A) = \mathbf{s}^1(B)$ *and* $T_1^{(1)} = T_2^{(1)}$. *Let* $\Delta(A, B, \beta_1, \beta_2)$ *be* $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) - [C_{11}(A) - C_{11}(B)] \, \beta_1 - [C_{22}(A) - C_{22}(B)] \, \beta_2$.

(i) If $C_{12}(A) = C_{12}(B)$ and $C_{11}(A) = C_{21}(B)$, then:

$$\Delta(A, B, \beta_1, \beta_2) \quad \geq \quad \left[T_1^{(1)}(A) - T^{(1,1)}(B)\right] \ \gamma_1 + \left[T^{(1,1)}(A) - T^{(1,1)}(B)\right] \ \gamma_2 \tag{32}$$

(ii) If $C_{12}(A) = C_{12}(B)$ and $C_{21}(A) = C_{11}(B)$, then:

$$\Delta(A, B, \beta_1, \beta_2) \quad \leq \quad \left[T^{(1,1)}(A) - T_1^{(1)}(B)\right] \ \gamma_1 + \left[T^{(1,1)}(A) - T^{(1,1)}(B)\right] \ \gamma_2 \tag{33}$$

(iii) If $C_{21}(A) = C_{21}(B)$ and $C_{22}(A) = C_{12}(B)$, then:

$$\Delta(A, B, \beta_1, \beta_2) \quad \geq \quad \left[T^{(1,1)}(A) - T^{(1,1)}(B)\right] \ \gamma_1 + \left[T_2^{(1)}(A) - T^{(1,1)}(B)\right] \ \gamma_2 \tag{34}$$

(iv) If $C_{21}(A) = C_{21}(B)$ and $C_{12}(A) = C_{22}(B)$, then:

$$\Delta(A, B, \beta_1, \beta_2) \quad \leq \quad \left[T^{(1,1)}(A) - T^{(1,1)}(B)\right] \ \gamma_1 + \left[T^{(1,1)}(A) - T_2^{(1)}(B)\right] \ \gamma_2 \tag{35}$$

Based on these inequalities, we can find pairs of market histories – A and B – that set identify the parameters β_1, β_2, γ_1, and γ_2. ■

The following examples present specific pairs of market histories that point identify the switching cost parameters and set identify the strategic interaction parameters.

EXAMPLE 8. Consider the pair of histories $A = [(0,0),\ (0,0),\ (1,1),\ (1,1)]$ and $B = [(0,0),\ (0,1),\ (1,0),\ (1,1)]$. These histories have the same value for the vector of statistics $s^1(\widetilde{\mathbf{y}}) = [T,\ y_{10},\ y_{20},\ y_{1T},\ y_{2T},\ T_1^{(1)},\ T_2^{(1)}]$. These histories also satisfy the condition $T_1^{(1)} = T_2^{(1)}$. Note that $C_{11}(A) - C_{11}(B) = 0$ and $C_{22}(A) - C_{22}(B) = 1$ such that $\Delta(A, B, \beta_1, \beta_{22}) = \ln \mathbb{P}(A) - \ln \mathbb{P}(B) - \beta_2$. We now check conditions (i) to (iv) in Proposition 6.

Condition (i) holds because $C_{12}(A) = C_{12}(B) = 1$ and $C_{11}(A) = C_{21}(B) = 1$. It implies:

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \geq \beta_{22} + \gamma_1 + \gamma_2 \tag{36}$$

Condition (ii) holds because $C_{12}(A) = C_{12}(B) = 1$ and $C_{21}(A) = C_{11}(B) = 1$. It implies:

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \leq \beta_{22} + \gamma_1 \tag{37}$$

Condition (iii) holds because $C_{21}(A) = C_{21}(B) = 1$ and $C_{22}(A) = C_{12}(B) = 1$. It implies:

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \geq \beta_{22} + \gamma_1 + \gamma_2 \tag{38}$$

Note that – for this example – this inequality is equivalent to the one provided by condition (i).

Condition (iv) does not hold because $C_{12}(A) = 1 \neq 0 = C_{22}(B)$. ■

We can also consider the mirror version of the pair of histories in Example 8. That is, consider $A = [(0,0),\ (0,0),\ (1,1),\ (1,1)]$ and $B = [(0,0),\ (1,0),\ (0,1),\ (1,1)]$. It is simple to show that this pair of histories imply the inequalities $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \geq \beta_1 + \gamma_1 + \gamma_2$ and $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \leq \beta_1 + \gamma_2$

These two examples may leave the impression that conditions (i) and (iii) generate always the same lower bound. This is not the case. For instance, consider the pairs of histories $A = [(0,0),\ (0,0),\ (0,1),\ (1,0)]$ and $B = [(0,0),\ (0,1),\ (0,0),\ (1,0)]$. For these histories, we have that $C_{12}(A) = 1 \neq 0 = C_{12}(B)$, and this implies that both condition (i) and condition (ii) fail. But condition (iii) and (iv) are satisfied and imply informative bounds on the parameters.

3.2 Forward-looking players

3.2.1 Forward-looking players with one-direction strategic interactions

Consider the complete information game in equation (**??**). It is convenient to represent this model as follows:

$$y_{it} \quad = \quad 1\left\{\widetilde{\alpha}_i + \beta_i\ y_{i,t-1} + \widetilde{\gamma}_{i\alpha}\ y_{jt} - \varepsilon_{it} \geq 0\right\} \tag{39}$$

where $\widetilde{\alpha}_i \equiv \alpha_i + \widetilde{v}_{i\alpha}(0)$, and $\widetilde{\gamma}_{i\alpha} \equiv \gamma_i + \widetilde{v}_{i\alpha}(1) - \widetilde{v}_{i\alpha}(0)$. Given this representation, it should be clear that it is not possible to point identify parameters γ_1 and γ_2 because they always appear together with the incidental parameters $\widetilde{v}_{i\alpha}(1) - \widetilde{v}_{i\alpha}(0)$.

Our purpose here is to study: (1) the point identification of the switching cost parameters β_1 and β_2; (2) the partial identification of parameters γ_1 and γ_2; and (3) whether there are triangular models – in the spirit of the models we studied in section 2.3 but now with forward-looking players – where the γ parameters are point identified.

We start here with a forward-looking, complete information, triangular dynamic game. Consider a version of the model with $\lambda_1 = \gamma_1 = 0$. Under these restrictions, the player 1's payoff does not depend on past, present, or future decisions of player 2. Therefore, the decision problem for player 1 is a single-agent problem, and it can represented as:

$$y_{1t} \quad = \quad 1\left\{\ \varepsilon_{1t} \leq \ \ \alpha_1 + \beta_1\ y_{1t-1} + \widetilde{v}_{1\alpha}\ \right\} \tag{40}$$

This identification of this forward-looking dynamic logit model – with fixed effects unobserved heterogeneity – has been established in Aguirregabiria, Gu, and Luo (2019). In this model: the incidental parameter is $\alpha_1 + \widetilde{v}_{1\alpha}$; the vector of sufficient statistics is $\mathbf{s}(\widetilde{\mathbf{y}}) = [y_{10},\ y_{1T},\ T_1^{(1)}]$;

and the structural parameter β_1 is identified from the maximization of the condtional likelihood function.

We now establish the point identification of parameters β_1 and β_2. The best response of player 2 in this triangular model is:

$$y_{2t} \;\;=\;\; 1\left\{\; \varepsilon_{2t} \leq \;\; \widetilde{\alpha}_2 + \lambda_2 \; y_{1,t-1} + \beta_2 \; y_{2,t-1} + \widetilde{\gamma}_{2\alpha} \; y_{1t} \;\right\} \tag{41}$$

where $\widetilde{\alpha}_2 \equiv \alpha_2 + \widetilde{v}_{2\alpha}(0)$, and $\widetilde{\gamma}_{2\alpha} \equiv \gamma_2 + \widetilde{v}_{2\alpha}(1) - \widetilde{v}_{2\alpha}(0)$. Given equations (40) and (41), the log-probability of the market history $\widetilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, .., T)$ has the following structure:

$$\begin{aligned}\ln \mathbb{P}\left(\widetilde{\mathbf{y}}|\boldsymbol{\alpha}, \beta\right) \;\;=&\;\; \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) + \alpha_1 \; T_1^{(1)} + \alpha_2 \; T_2^{(1)} + \widetilde{\gamma}_{2\alpha} \; T^{(1,1)} + \sum_{t=1}^{T} \sigma_{\boldsymbol{\alpha}1}\left(y_{1t-1}\right) + \sigma_{\boldsymbol{\alpha}2}\left(y_{1t}, y_{2t-1}\right) \\ &+ \;\; \beta_1 \; C_{11} + \beta_2 \; C_{22}\end{aligned} \tag{42}$$

where $\sigma_{\boldsymbol{\alpha}1}(y_1) \equiv -\ln[1 + \exp\{\alpha_1 + \beta_1 \; y_1\}]$ and $\sigma_{\boldsymbol{\alpha}2}(y_1, y_2) \equiv -\ln[1 + \exp\{\alpha_2 + \widetilde{\gamma}_{2\alpha} \; y_1 + \beta_2 \; y_2\}]$. We can rewrite this equation for the log-probability of a market history as $\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$ as $\mathbf{s}\left(\widetilde{\mathbf{y}}\right)'$ $\mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}\left(\widetilde{\mathbf{y}}\right)' \; \beta^*$, with

$$\begin{cases} \mathbf{s}\left(\widetilde{\mathbf{y}}\right)' \;=\; [1,\; y_{10},\; y_{20},\; y_{10}y_{20} \quad ; \quad 1,\; y_{1T},\; y_{2T},\; y_{1T}y_{2T} \quad ; \quad T,\; T_1^{(1)},\; T_2^{(1)},\; T^{(1,1)} \quad ; \quad C_{12}] \\ \mathbf{c}\left(\widetilde{\mathbf{y}}\right)' \;=\; [C_{11},\; C_{22}] \\ \boldsymbol{\theta}^{*\prime} \;=\; [\beta_1,\; \beta_2] \end{cases} \tag{43}$$

PROPOSITION 7. For the forward-looking dynamic game with one-direction strategic interactions as described in equations (40) and (41): (A) The vector $\mathbf{s}\left(\widetilde{\mathbf{y}}\right) = [1,\; y_{10},\; y_{20},\; y_{10}y_{20},$ $y_{1T},\; y_{2T},\; y_{1T}y_{2T},\; T,\; T_1^{(1)},\; T_2^{(1)},\; T^{(1,1)},\; C_{12}]'$ *is a minimal sufficient statistic for* $\boldsymbol{\alpha}$ *such that* $\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{u}\left(\widetilde{\mathbf{y}}\right), \alpha, \boldsymbol{\theta}\right)$ *does not depend on* $\boldsymbol{\alpha}$. *(B)* $\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{c}\left(\widetilde{\mathbf{y}}\right), \beta\right) = \mathbf{s}\left(\widetilde{\mathbf{y}}\right)' \boldsymbol{\theta}^* - \ln(\sum_{\widetilde{\mathbf{y}}' : \mathbf{s}(\widetilde{\mathbf{y}}') = \mathbf{s}(\widetilde{\mathbf{y}})} \exp\left\{\mathbf{c}(\widetilde{\mathbf{y}}')'\boldsymbol{\theta}^*\right\}$ *with* $\mathbf{c}\left(\widetilde{\mathbf{y}}\right) = [C_{11},\; C_{22}]'$ *and* $\boldsymbol{\theta}^* = [\beta_1,\; \beta_2]'$. *(C) For* $T \geq 3$, *there are histories* $\widetilde{\mathbf{y}}$ *such that* $\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{s}\left(\widetilde{\mathbf{y}}\right), \beta\right)$ *identifies the vector of parameters of interest* $\boldsymbol{\theta}^*$. ■

EXAMPLE 10. The same histories in Example 3 that – in the myopic, complete information, triangular model – identify parameters β_1 and β_2, still identify these parameters in the forward-looking version of the model. More specifically: the pair of histories $A = \{(0,0),\; (0,0),\; (1,0),\; (1,0)\}$ and $B = \{(0,0),\; (1,0),\; (0,0),\; (1,0)\}$ identifies β_1; and the pair of histories $A = \{(0,0),\; (0,0),\; (0,1),\; (0,1)\}$ and $B = \{(0,0),\; (0,1),\; (0,0),\; (0,1)\}$ identifies β_2. ■

3.2.2 Forward-looking players with two-direction strategic interactions

We consider now the forward-looking dynamic game where we do not restrict the parameters γ_1 or γ_2 to be zero.

$$\begin{cases} y_{1t} & = & 1\left\{ \ \varepsilon_{1t} \leq \ \ \widetilde{\alpha}_1 + \beta_1 \ y_{1,t-1} + \widetilde{\gamma}_{1\alpha} \ y_{2t} \ \right\} \\ \\ y_{2t} & = & 1\left\{ \ \varepsilon_{2t} \leq \ \ \widetilde{\alpha}_2 + \beta_2 \ y_{2,t-1} + \widetilde{\gamma}_{2\alpha} \ y_{1t} \ \right\} \end{cases} \tag{44}$$

where $\widetilde{\alpha}_i \equiv \alpha_i + \widetilde{v}_{i\alpha}(0)$, and $\widetilde{\gamma}_{i\alpha} \equiv \gamma_i + \widetilde{v}_{i\alpha}(1) - \widetilde{v}_{i\alpha}(0)$.

The model has a similar structure as the myopic. The main difference is that now the random variables $(\widetilde{\gamma}_{1\alpha}, \widetilde{\gamma}_{2\alpha})$ replace the parameters (γ_1, γ_2). Therefore, the expressions of the lower and upper bounds for the log-probability of a market history are very similar to the ones in Lemma 2 for the myopic model, but replacing (γ_1, γ_2) with $(\widetilde{\gamma}_{1\alpha}, \widetilde{\gamma}_{2\alpha})$. Though this different is coneptually simple, it has substantial implications on the identification of the γ parameters. More specifically, we cannot point identify the switching cost parameters. Proposition 8 establsihes that these parameters are partially identified.

Proposition 8. Consider the forward-looking complete information dynamic game with contemporaneous effects in equation (44). Under conditions $\beta_1 \geq 0$, $\beta_2 \geq 0$, $\tilde{\gamma}_{1\alpha} \leq 0$, and $\tilde{\gamma}_{2\alpha} \leq 0$, there are market histories that provide informative bounds on the parameters β_1 and β_2. These parameters are partially identified. ■

Proof of Proposition 8. Denote the following terms:

$$\begin{aligned} \sigma_{\alpha_1}(y_{1t-1}, y_{2t}) &= -\ln\{1 + \exp(\tilde{\alpha}_1 + \beta_1 y_{1t-1} + \tilde{\gamma}_{1\alpha} y_{2t})\} \\ \sigma_{\alpha_2}(y_{1t}, y_{2t-1}) &= -\ln\{1 + \exp(\tilde{\alpha}_2 + \beta_2 y_{2t-1} + \tilde{\gamma}_{2\alpha} y_{1t}\} \\ \Delta\sigma_{\alpha_1}(1,0) &= \sigma_{\alpha_1}(1,0) - \sigma_{\alpha_1}(0,0) \\ \Delta\sigma_{\alpha_1}(0,1) &= \sigma_{\alpha_1}(0,1) - \sigma_{\alpha_1}(0,0) \\ \Delta\sigma_{\alpha_2}(1,0) &= \sigma_{\alpha_2}(1,0) - \sigma_{\alpha_2}(0,0) \\ \Delta\sigma_{\alpha_2}(0,1) &= \sigma_{\alpha_2}(0,1) - \sigma_{\alpha_2}(0,0) \\ \Delta^2\sigma_{\alpha_1} &= \sigma_{\alpha_1}(1,1) - \sigma_{\alpha_1}(1,0) - \sigma_{\alpha_1}(0,1) + \sigma_{\alpha_1}(0,0) \\ \Delta^2\sigma_{\alpha_2} &= \sigma_{\alpha_2}(1,1) - \sigma_{\alpha_2}(1,0) - \sigma_{\alpha_2}(0,1) + \sigma_{\alpha_2}(0,0) \end{aligned}$$

Under the conditions of Proposition 8, we have (i) $\Delta\sigma_{\alpha_1}(1,0) \leq 0$, (ii) $\Delta\sigma_{\alpha_1}(0,1) \geq 0$, (iii) $\Delta\sigma_{\alpha_2}(1,0) \geq 0$ and $\Delta\sigma_{\alpha_2}(0,1) \leq 0$, (iv) $\Delta^2\sigma_{\alpha_1} \geq 0$ and, (v) $\Delta^2\sigma_{\alpha_2} \geq 0$. For each choice history

$\tilde{y}$, we have the following lower and upper bound:

$$
\begin{aligned}
\ln P_U(\tilde{y}|\alpha,\beta) &= \ln P_\alpha(y_{10},y_{20}) + s^1(\tilde{y})' g^1_\alpha + [y_{10}-y_{1T}, y_{20}-y_{2T}, T^{(1,1)}, T^{(1,1)}]' g^2_\alpha + [T^{(1)}_2, T^{(1)}_1, C_{21}, C_{12}] g^3_\alpha \\
&\quad + C_{11}\beta_1 + C_{22}\beta_2 \\
\ln P_{L\{E,W\}}(\tilde{y}|\alpha,\beta) &= \ln P_\alpha(y_{10},y_{20}) + s^1(\tilde{y})' g^1_\alpha + [y_{10}-y_{1T}, y_{20}-y_{2T}, T^{(1)}_1, T^{(1,1)}]' g^2_\alpha + [T^{(1)}_1, T^{(1)}_1, C_{11}, C_{12}] g^3_\alpha \\
&\quad + C_{11}\beta_1 + C_{22}\beta_2 \\
\ln P_{L\{S,N\}}(\tilde{y}|\alpha,\beta) &= \ln P_\alpha(y_{10},y_{20}) + s^1(\tilde{y})' g^1_\alpha + [y_{10}-y_{1T}, y_{20}-y_{2T}, T^{(1,1)}, T^{(1)}_2]' g^2_\alpha + [T^{(1)}_2, T^{(1)}_2, C_{21}, C_{22}] g^3_\alpha \\
&\quad + C_{11}\beta_1 + C_{22}\beta_2 \\
lnP_{L\{E,N\}}(\tilde{y}|\alpha,\beta) &= \ln P_\alpha(y_{10},y_{20}) + s^1(\tilde{y})' g^1_\alpha + [y_{10}-y_{1T}, y_{20}-y_{2T}, T^{(1,1)}, T^{(1,1)}]' g^2_\alpha \\
&\quad + [T^{(1,1)}, T^{(1,1)}, R^{(1,1)}_1, R^{(1,1)}_2] g^3_\alpha \\
&\quad + C_{11}\beta_1 + C_{22}\beta_2 \\
\ln P_{L\{S,W\}}(\tilde{y}|\alpha,\beta) &= \ln P_\alpha(y_{10},y_{20}) + s^1(\tilde{y})' g^1_\alpha + [y_{10}-y_{1T}, y_{20}-y_{2T}, T^{(1)}_1, T^{(1)}_2]' g^2_\alpha \\
&\quad + [T^{(1)}_1 + T^{(1)}_2 - T^{(1,1)}, T^{(1)}_1 + T^{(1)}_2 - T^{(1,1)}, C_{11} + C_{21} - R^{(1,1)}_1, C_{12} + C_{22} - R^{(1,1)}_2] g^3_\alpha \\
&\quad + C_{11}\beta_1 + C_{22}\beta_2
\end{aligned}
$$

where $s^1(\tilde{y}) = [T, T^{(1)}_1, T^{(1)}_2]$, $g^1_\alpha = [\sigma_{\alpha_1}(0,0) + \sigma_{\alpha_2}(0,0), \alpha_1 + \Delta\sigma_{\alpha_1}(1,0), \alpha_2 + \Delta\sigma_{\alpha_2}(0,1)]'$, $g^2_\alpha = [\Delta\sigma_{\alpha_1}(1,0), \Delta\sigma_{\alpha_2}(0,1), \tilde{\gamma}_{1\alpha}, \tilde{\gamma}_{2\alpha}]'$, and $g^3_\alpha = [\Delta\sigma_{\alpha_1}(0,1), \Delta\sigma_{\alpha_2}(1,0), \Delta^2\sigma_{\alpha_1}, \Delta^2\sigma_{\alpha_2}]'$

The grouping of the g^j_α with $j = \{1,2,3\}$ terms are such that terms in g^1_α can be any sign for $\{\alpha_1,\alpha_2\} \in \mathbb{R}^2$. Terms in g^2_α are all negative. And all terms in g^3_α are positive.

We first present bounds constructed using the differences of the logrithm of the probability of a pair of choice history that satisfy certain conditions, i.e. $\ln \frac{P(A)}{P(B)} = \ln P(A) - \ln P(B)$. We then generalize to bounds constructed from $\frac{\sum_{\lambda\in S^U} P(\lambda)}{\sum_{\lambda'\in S^L} P(\lambda')}$, where the set S^U and S^L are some set of choice histories (not necessarily a singleton) that satisfy certain conditions. We focus on upper bound, because the result of lower bound from such sequences are providing symmetric information (i.e. the lower bound of $\frac{P(A)}{P(B)}$ is providing equivalent information from the upper bound of $\frac{P(B)}{P(A)}$).

For a pair of choice histroies A and B, define

$$\Delta(A,B,\beta_1,\beta_2) = \ln P(A) - \ln P(B) - [C_{11}(A) - C_{11}(B)]\beta_1 - [C_{22}(A) - C_{22}(B)]\beta_2$$

Define the statistics $s^1(\tilde{y}) = [T, T^{(1)}_1, T^{(1)}_2]$

[1] Using upper bound and $L\{E,W\}$:

$$\Delta(A,B,\beta_1,\beta_2) \le 0$$

provided the following conditions hold: (i) $y_{10}(A) = y_{20}(B)$, (ii) $y_{20}(A) - y_{20}(B)$, (iii) $s^1(A) = s^1(B)$, (iv) element-wise, $[y_{10}(A) - y_{1T}(A), y_{20}(A) - y_{2T}(A), T^{(1,1)}(A), T^{(1,1)}(A)] - [y_{10}(B) - y_{1T}(B), y_{20}(B) - y_{2T}(B), T_1^{(1)}(B), T^{(1,1)}(B)] \geq 0$, (v) element-wise, $[T_2^{(1)}(A), T_1^{(1)}(A), S_{21}(A), S_{12}(A)] - [T_1^{(1)}(B),\ T_1^{(1)}(B), S_{11}(B), S_{12}(B)] \leq 0$.

[2] Using upper bound and $L\{S, N\}$:

$$\Delta(A, B, \beta_1, \beta_2) \leq 0$$

provided the following conditions hold: (i) $y_{10}(A) = y_{20}(B)$, (ii) $y_{20}(A) - y_{20}(B)$, (iii) $s^1(A) = s^1(B)$, (iv) element-wise, $[y_{10}(A) - y_{1T}(A), y_{20}(A) - y_{2T}(A), T^{(1,1)}(A), T^{(1,1)}(A)] - [y_{10}(B) - y_{1T}(B), y_{20}(B) - y_{2T}(B), T^{(1,1)}(B), T_2^{(1)}(B)] \geq 0$, (v) element-wise, $[T_2^{(1)}(A), T_1^{(1)}(A), C_{21}(A), C_{12}(A)] - [T_2^{(1)}(B),\ T_2^{(1)}(B), C_{21}(B), C_{22}(B)] \leq 0$.

[3] Using upper bound and $L\{E, N\}$:

$$\Delta(A, B, \beta_1, \beta_2) \leq 0$$

provided the following conditions hold: (i) $y_{10}(A) = y_{20}(B)$, (ii) $y_{20}(A) - y_{20}(B)$, (iii) $s^1(A) = s^1(B)$, (iv) element-wise, $[y_{10}(A) - y_{1T}(A), y_{20}(A) - y_{2T}(A), T^{(1,1)}(A), T^{(1,1)}(A)] - [y_{10}(B) - y_{1T}(B), y_{20}(B) - y_{2T}(B), T^{(1,1)}(B), T^{(1,1)}(B)] \geq 0$, (v) element-wise, $[T_2^{(1)}(A), T_1^{(1)}(A), C_{21}(A), C_{12}(A)] - [T^{(1,1)}(B), T^{(1,1)}(B), R_1^{(1,1)}(B), R_2^{(1,1)}(B)] \leq 0$.

[4] Using upper bound and $L\{S, W\}$:

$$\Delta(A, B, \beta_1, \beta_2) \leq 0$$

provided the following conditions hold: (i) $y_{10}(A) = y_{20}(B)$, (ii) $y_{20}(A) - y_{20}(B)$, (iii) $s^1(A) = s^1(B)$, (iv) element-wise, $[y_{10}(A) - y_{1T}(A), y_{20}(A) - y_{2T}(A), T^{(1,1)}(A), T^{(1,1)}(A)] - [y_{10}(B) - y_{1T}(B), y_{20}(B) - y_{2T}(B), T_1^{(1)}(B), T_2^{(1)}(B)] \geq 0$, (v) element-wise, $[T_2^{(1)}(A), T_1^{(1)}(A), S_{21}(A), S_{12}(A)] - [T_1^{(1)}(B) + T_2^{(1)}(B) - T^{(1,1)}(B), T_1^{(1)}(B) + T_2^{(1)}(B) - T^{(1,1)}(B), C_{11}(B) + C_{21}(B) - R_1^{(1,1)}(B), C_{12}(B) + C_{22}(B) - R_2^{(1,1)}(B)] \leq 0$.

For each combination of the upper and lower bound, the conditions (i) and (ii) imposed on A and B makes sure to cancel out $\ln P_\alpha(y_{10}, y_{20})$, and condition (iii) makes sure to cancel the terms in front of g_α^1 that we can not determine its sign and condition. Condition (iv) takes advantage of the fact that all elements in $g_\alpha^2 \leq 0$ under the conditions of Proposition 8 those terms can be replaced by 0 in the upper bound of $\ln P(A) - \ln P(B)$. Finally, condition (v) takes advantage of the fact that all elements in $g_\alpha^3 \geq 0$ under the conditions of Proposition 8 such that those terms can be replaced by 0 in the upper bound of $\ln P(A) - \ln P(B)$. ■

EXAMPLE 11. $A = [(0,1),(1,1),(0,0),(0,0)]$ and $B = [(0,1),(1,0),(0,1),(0,0)]$. For this pair, we have $T_1^{(1)}(A) = T_1^{(1)}(B) = T_2^{(1)}(A) = T_2^{(1)}(B) = 1$, $T^{(1,1)}(B) = 0$, $T^{(1,1)}(A) = 1$, $C_{11}(A) = C_{11}(B) = 0$, $C_{22}(B) = 0 \neq 1 = C_{22}(A)$, and $C_{12}(A) = C_{12}(B) = 1$ and $C_{21}(B) = 1$ and $C_{21}(A) = 0$. Therefore $\ln P(A) - \ln P(B) \leq \beta_2$

Other example. $A = [(1,0),(1,1),(0,0),(0,0)]$ and $B = [(1,0),(0,1),(1,0),(0,0)]$. For this pair, we have $T_1^{(1)}(A) = T_1^{(1)}(B) = T_2^{(1)}(A) = T_2^{(1)}(B) = 1$, $T^{(1,1)}(B) = 0$ and $T^{(1,1)}(A) = 1$, $C_{11}(B) = 0 \neq 1 = C_{11}(A)$, $C_{22}(A) = C_{22}(B) = 0$, $C_{12}(B) = 1$, $C_{12}(A) = 0$, $C_{21}(A) = 1 = C_{21}(B)$, which leads to $\ln P(A) - \ln P(B) \leq \beta_1$. ■

4 Empirical application

TBW

5 Conclusions

TBW

A APPENDIX

A.1 Proof of Proposition 4

Consider the model with myopic players and one-direction strategic interactions where we assume there is only market level unobserved heterogeneity, i.e. $\alpha_1 = \alpha_2$. Consider the case $y_0 = \{y_{10}, y_{20}\} = \{0,0\}$ and $T = 2$, such that we have 16 possible choice histories. The choice probabilities conditional on the market level unobserved heterogeneity can be represented using the expressions in the following table:

$\{y_{11}, y_{21}\}$	$\{y_{12}, y_{22}\}$	$P(\widetilde{\mathbf{y}})$	$\{y_{11}, y_{21}\}$	$\{y_{12}, y_{22}\}$	$P(\widetilde{\mathbf{y}})$
$\{0,0\}$	$\{0,0\}$	$\left(\frac{1}{1+A}\right)^4$	$\{0,0\}$	$\{1,0\}$	$\frac{1}{1+A}\ \frac{1}{1+A}\ \frac{A}{1+A}\ \frac{1}{1+AC}$
$\{0,1\}$	$\{0,0\}$	$\frac{1}{1+A}\ \frac{A}{1+A}\ \frac{1}{1+AB_{12}}\ \frac{1}{1+AB_{22}}$	$\{0,1\}$	$\{1,0\}$	$\frac{1}{1+A}\ \frac{A}{1+A}\ \frac{AB_{12}}{1+AB_{12}}\ \frac{1}{1+ACB_{22}}$
$\{1,0\}$	$\{0,0\}$	$\frac{A}{1+A}\ \frac{1}{1+A}\ \frac{1}{1+AB_{11}}\ \frac{1}{1+A}$	$\{1,0\}$	$\{1,0\}$	$\frac{A}{1+A}\ \frac{1}{1+AC}\ \frac{AB_{11}}{1+AB_{11}}\ \frac{1}{1+A}$
$\{1,1\}$	$\{0,0\}$	$\frac{A}{1+A}\ \frac{AC}{1+AC}\ \frac{1}{AB_{11}B_{12}}\ \frac{1}{AB_{22}}$	$\{1,1\}$	$\{1,0\}$	$\frac{A}{1+A}\ \frac{AC}{1+AC}\ \frac{AB_{11}B_{12}}{1+AB_{11}B_{12}}\ \frac{1}{ACB_{22}}$
$\{0,0\}$	$\{0,1\}$	$\left(\frac{1}{1+A}\right)^3\ \frac{A}{1+A}$	$\{0,0\}$	$\{1,1\}$	$\frac{1}{1+A}\ \frac{1}{1+A}\ \frac{A}{1+A}\ \frac{AC}{1+AC}$
$\{0,1\}$	$\{0,1\}$	$\frac{1}{1+A}\ \frac{A}{1+A}\ \frac{1}{1+AB_{12}}\ \frac{AB_{22}}{1+AB_{22}}$	$\{0,1\}$	$\{1,1\}$	$\frac{1}{1+A}\ \frac{A}{1+A}\ \frac{AB_{12}}{1+AB_{12}}\ \frac{ACB_{22}}{1+ACB_{22}}$
$\{1,0\}$	$\{0,1\}$	$\frac{A}{1+A}\ \frac{1}{1+AC}\ \frac{1}{1+AB_{11}}\ \frac{A}{1+A}$	$\{1,0\}$	$\{1,1\}$	$\frac{A}{1+A}\ \frac{1}{1+AC}\ \frac{AB_{11}}{1+AB_{11}}\ \frac{AC}{1+AC}$
$\{1,1\}$	$\{0,1\}$	$\frac{A}{1+A}\ \frac{AC}{1+AC}\ \frac{1}{1+AB_{11}B_{12}}\ \frac{AB_{22}}{1+AB_{22}}$	$\{1,1\}$	$\{1,1\}$	$\frac{A}{1+A}\ \frac{AC}{1+AC}\ \frac{AB_{11}B_{12}}{1+AB_{11}B_{12}}\ \frac{ACB_{22}}{1+ACB_{22}}$

where $A = \exp(\alpha)$, $B_{11} = \exp(\beta_1)$, $B_{12} = \exp(\lambda_1)$, $B_{22} = \exp(\beta_2)$ and $C = \exp(\gamma_2)$.

Define $g(\alpha, \boldsymbol{\theta})$ as the minimum common denominator (MCD) of all the ratios in the table above. It is simple to verify that this MCD has the following expression.

$$g(\alpha, \boldsymbol{\theta}) \equiv (1+A)^4\ (1+AC)^2\ (1+AB_{11})\ (1+AB_{12})\ (1+AB_{22})\ (1+AB_{11}B_{12})\ (1+ACB_{22}) \quad (45)$$

By definition of MCD, we have that $P(\widetilde{\mathbf{y}}|\alpha, \boldsymbol{\theta})\ g(\alpha, \boldsymbol{\theta})$ is a polynomial function of A with its coefficients being polynomials of $(B_{11}, B_{12}, B_{22}, C)$. It is also straightforward to verify that for any $\alpha \in \mathbb{R}$, we have that $1/g(\alpha, \boldsymbol{\theta}) \in (0,1]$. Taking these properties into account, we can write:

$$P(\widetilde{\mathbf{y}} \mid \boldsymbol{\theta}, y_0) \;=\; \int P(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}, y_0)\; f(\alpha \mid y_0)\; d\alpha \;=\; \int P(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}, y_0)\; g(\alpha, \boldsymbol{\theta})\; q(\alpha \mid \boldsymbol{\theta}, y_0)\; d\alpha \quad (46)$$

where f is the distribution of the market level fixed effect, and $q(\alpha \mid \boldsymbol{\theta}, y_0) = \frac{f(\alpha | y_0)}{g(\alpha, \boldsymbol{\theta})}$. Function $q(\alpha \mid \boldsymbol{\theta}, y_0)$ is a positive Borel measure on the support $[0, \infty)$. Though q is not a probability measure, it is simple to construct a probability measure by dividing q by its integral over α, $\int q(\alpha \mid \boldsymbol{\theta}, y_0)\, d\alpha$. Since $1/g(\alpha, \boldsymbol{\theta})$ is finite everywhere on the support of α, this integral exists and is finite.

Given equation (46) and after some calculations, we can write the following system of 16 equations relating probabilities of choice histories with the vector of parameters $\boldsymbol{\theta}$ and moments in the distribution of A. Using matrix notation, this system is:

$$\mathbf{P}_{\widetilde{\mathbf{y}}} \;=\; \mathbf{G}(\boldsymbol{\theta})\; \mathbf{m}_A \quad (47)$$

where $\mathbf{P}_{\widetilde{\mathbf{y}}}$ is the 16×1 vector with the empirical probabilities of all the possible choice histories; $\mathbf{G}(\boldsymbol{\theta})$ is a 16×12 matrix with its elements only involving $\{B_{11}, B_{12}, B_{22}, C\}$; and $\mathbf{m}_A$ is a 12×1 vector with the power moments of the measure q, that is:

$$\mathbf{m}_A \;\equiv\; \int \begin{pmatrix} 1 & A & A^2 & \ldots & A^{11} \end{pmatrix}' \; q(\alpha \mid \boldsymbol{\theta})\; d\alpha \quad (48)$$

Given the system of equations in (47), we can construct a moment condition for $\boldsymbol{\theta}$ by finding a vector $\mathbf{v} \in \mathbb{R}^{16}$ – that may depend on $\boldsymbol{\theta}$ – such that $\mathbf{v}'G(\boldsymbol{\theta}) = 0$. By definition, the collection of vectors $\mathbf{v}$ satisfying this condition is nothing but the elements in the left null space of the matrix $\mathbf{G}(\boldsymbol{\theta})$. Hence, we can just take all elements in a basis that spans the left null space of $\mathbf{G}(\boldsymbol{\theta})$.

In our specific case here with $T = 2$, the rank of $\mathbf{G}(\boldsymbol{\theta})$ is 4, hence the dimension of the left null space of $\mathbf{G}(\boldsymbol{\theta})$ is 4, and we can find 4 linearly independent moment conditions for $\boldsymbol{\theta}$. In particular, two of them take the form:

$$\begin{array}{rcl} -B_{11}\; P_{(1,0),(0,1)} + P_{(1,0),(1,0)} & = & 0 \\ -C\; P_{(1,0),(0,0)} - B_{11}\; C\; P_{(1,0),(0,1)} + C\; P_{(0,0),(1,0)} + P_{(0,0),(1,1)} & = & 0 \end{array} \quad (49)$$

These two moment conditions identify B_{11} and C, or what is equivalent, β_1 and γ_2. We have two more moment conditions for the identification of λ_1 and β_2. They have a more complivated form which is the following:

$$\begin{aligned} & \frac{B_{22}(C-1)}{B_{22}-C}\ (P_{(1,0),(0,0)} - P_{(0,1),(0,0)}) - B_{22}\ P_{(1,1),(0,0)} \\ + \quad & \frac{B_{12}C - B_{22}C + B_{22}^2 - B_{12}B_{22}C}{B_{22}(B_{22}-C)}\ P_{(0,1),(0,1)} \\ + \quad & \frac{B_{11}B_{22}(C-1)}{B_{22}-C}\ P_{(1,0),(0,1)} - B_{11}B_{12}\ P_{(1,1),(0,1)} + P_{(0,1),(1,0)} + P_{(0,1),(1,1)} \quad = 0 \end{aligned} \tag{50}$$

and

$$\begin{aligned} & \frac{B_{11}B_{12}(C-1)^2}{C(B_{22}-C)(B_{11}B_{12}-B_{22}C)(B_{22}-1)}(P_{(1,0),(0,0)} - P_{(0,1),(0,0)}) - \frac{B_{11}B_{12}(C-1)}{B_{22}C^2 - B_{11}B_{12}C}P_{(1,1),(0,0)} \\ - \quad & \frac{B_{11}B_{12}(B_{12}C - B_{22}C - B_{22}^3C + B_{22}^2 + B_{22}^2C^2 - 2B_{12}B_{22}C + B_{12}B_{22}^2C)}{B_{22}^2C(B_{22}-C)(B_{11}B_{12}-B_{22}C)(B_{22}-1)}P_{(0,1),(0,1)} \\ + \quad & \frac{B_{11}^2B_{12}(C-1)^2}{C(B_{22}-C)(B_{11}B_{12}-B_{22}C)(B_{22}-1)}P_{(1,0),(0,1)} \\ - \quad & \frac{B_{11}B_{12}}{B_{22}C}P_8 + \frac{B_{11}(B_{22}C-1)(B_{12}-B_{22}C)}{B_{22}C(B_{11}B_{12}-B_{22}C)(B_{22}-1)}P_{(0,1),(1,0)} + P_{(1,1),(1,0)} \ = \ 0 \end{aligned} \tag{51}$$

A.2 Proof of Lemma 1

Given the structure for the lower bound – $\ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbf{s}'_L\mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}'_L\boldsymbol{\theta}$ – and for the upper bound – $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbf{s}'_U\mathbf{g}_\alpha + \mathbf{c}'_U\boldsymbol{\theta}$ – we have that:

$$\exp\{\mathbf{s}'_L\mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}'_L\boldsymbol{\theta}\} \le \mathbb{P}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}) \le \exp\{\mathbf{s}'_U\mathbf{g}_{\boldsymbol{\alpha}} + \mathbf{c}'_U\boldsymbol{\theta}\} \tag{A.1.1}$$

Integrating the inequalities in (A.1.1) over the distribution of $\boldsymbol{\alpha}$ we have that the inequalities still hold and they take the following form:

$$\left[\int \exp\{\mathbf{s}'_L\mathbf{g}_{\boldsymbol{\alpha}}\} f(\alpha)\ d\boldsymbol{\alpha}\right]\ \exp\{\mathbf{c}'_L\boldsymbol{\theta}\} \le \mathbb{P}(\widetilde{\mathbf{y}}) \le \left[\int \exp\{\mathbf{s}'_U\mathbf{g}_{\boldsymbol{\alpha}}\} f(\boldsymbol{\alpha})\ d\alpha\right]\ \exp\{\mathbf{c}'_U\boldsymbol{\theta}\} \tag{A.1.2}$$

Define $h(\mathbf{s})$ as $\ln\left[\int \exp\{\mathbf{s}'\mathbf{g}_{\boldsymbol{\alpha}}\}\ f(\boldsymbol{\alpha})\ d\boldsymbol{\alpha}\right]$. Then, we have that:

$$h(\mathbf{s}_L) + \mathbf{c}_L'\beta \leq \ln \mathbb{P}(\widetilde{\mathbf{y}}) \leq h(\mathbf{s}_U) + \mathbf{c}_U'\boldsymbol{\theta} \tag{A.1.3}$$

A.3 Proof of Lemma 2

For the derivations below, we use the following definitions: $\sigma_{\alpha 1}(y_{1t-1}, y_{2t}) \equiv -\ln[1+\exp\{\alpha_1+\beta_1\ y_{1t-1}+\gamma_1 y_{2t}\}]$ and $\sigma_{\alpha 2}(y_{1t}, y_{2t-1}) \equiv -\ln[1+\exp\{\alpha_2+\beta_2\ y_{2t-1}+\gamma_2 y_{1t}\}]$, and

$$\begin{aligned} \mathbf{s}^1(\widetilde{\mathbf{y}})'\,\mathbf{g}^1_{\boldsymbol{\alpha}} \ &\equiv\ \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) + T\ \left[\sigma_{\alpha 1}(0,0) + \sigma_{\alpha 2}(0,0)\right] \\ &+\ (y_{10} - y_{1T})\ \Delta\sigma_{\alpha 1}(1,0) + (y_{20} - y_{2T})\ \Delta\sigma_{\alpha 2}(0,1) \\ &+\ T_1^{(1)}\ \left[\alpha_1 + \Delta\sigma_{\alpha 1}(0,1)\right] + T_2^{(1)}\ \left[\alpha_2 + \Delta\sigma_{\alpha 2}(0,1)\right] \end{aligned} \tag{A.1}$$

where $\Delta\sigma_{\alpha 1}(1,0) \equiv \sigma_{\alpha 1}(1,0) - \sigma_{\alpha 1}(0,0)$; and $\Delta\sigma_{\alpha 2}(0,1) \equiv \sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0)$.

We also define the vector of incidental parameters:

$$\mathbf{g}^2_{\boldsymbol{\alpha}} \ \equiv\ \Big[\ \Delta\sigma_{\alpha 1}(0,1),\ \Delta\sigma_{\alpha 2}(1,0),\ \Delta^2\sigma_{\alpha 1},\ \Delta^2\sigma_{\alpha 2}\ \Big]' \tag{A.2}$$

where $\Delta\sigma_{\alpha 1}(0,1) \equiv \sigma_{\alpha 1}(0,1) - \sigma_{\alpha 1}(0,0)$; $\Delta\sigma_{\alpha 2}(1,0) \equiv \sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,0)$; $\Delta^2\sigma_{\alpha 1} \equiv \sigma_{\alpha 1}(1,1) - \sigma_{\alpha 1}(1,0) - \sigma_{\alpha 1}(0,1) + \sigma_{\alpha 1}(0,0)$; and $\Delta^2\sigma_{\alpha 2} \equiv \sigma_{\alpha 2}(1,1) - \sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,1) + \sigma_{\alpha 2}(0,0)$.

And the statistics $R_1^{(1,1)} \equiv \sum_{t=1}^{T} y_{1t-1}\ y_{1t}\ y_{2t}$ and $R_2^{(1,1)} \equiv \sum_{t=1}^{T} y_{2t-1}\ y_{1t}\ y_{2t}$.

(a) Lower Bound $\ln \mathbb{P}_{L\{E,W\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$. To obtain this lower bound, we use the bounds $L^{\{E,SE\}}(0,1|\mathbf{y}_{t-1};\boldsymbol{\alpha}) \equiv [1-\Lambda(\alpha_1+\beta_1\ y_{1t-1})]\ \Lambda(\alpha_2+\beta_2\ y_{2t-1})$ and $L^{\{W,NW\}}(1,0|\mathbf{y}_{t-1};\boldsymbol{\alpha}) \equiv \Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1)\ [1-\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)]$ for the choice probabilities. Then,

$$\begin{aligned} \ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \ &\geq\ \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\ &+\ \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t})\ \ \left(\ln\left[1-\Lambda(\alpha_1+\beta_1\ y_{1t-1})\right] + \ln\left[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})\right]\right) \\ &+\ \sum_{t=1}^{T} (1-y_{1t})\, y_{2t}\ \ \left(\ln\left[1-\Lambda(\alpha_1+\beta_1\ y_{1t-1})\right] + \ln\Lambda(\alpha_2+\beta_2\ y_{2t-1})\right) \\ &+\ \sum_{t=1}^{T} y_{1t}(1-y_{2t})\ \ \left(\ln\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1) + \ln\left[1-\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)\right]\right) \\ &+\ \sum_{t=1}^{T} y_{1t}y_{2t}\ \ \left(\ln\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1) + \ln\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)\right) \end{aligned} \tag{A.3}$$

Using the definitions $\sigma_{\alpha 1}(y_{1t-1}, y_{2t})$ and $\sigma_{\alpha 2}(y_{1t}, y_{2t-1})$, we have:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \;\geq\;& \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) \\
+& \sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \;\left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})\right] \\
+& \sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \;\left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_2 + \beta_2 \; y_{2t-1}\right] \\
+& \sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \;\left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_1 \; y_{1t-1} + \gamma_1\right] \\
+& \sum_{t=1}^{T} y_{1t} y_{2t} \;\left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_1 \; y_{1t-1} + \gamma_1 + \alpha_2 + \beta_2 \; y_{2t-1} + \gamma_2\right]
\end{aligned} \tag{A.4}$$

Grouping terms, we have:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \;\geq\;& \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) \\
+& \sum_{t=1}^{T} \left(1 - y_{1t}\right) \;\left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})\right] \\
+& \sum_{t=1}^{T} y_{2t} \;\left[\alpha_2 + \beta_2 \; y_{2t-1}\right] \\
+& \sum_{t=1}^{T} y_{1t} \;\left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_1 \; y_{1t-1} + \gamma_1\right] \\
+& \sum_{t=1}^{T} y_{1t} y_{2t} \;\left[\gamma_2\right]
\end{aligned} \tag{A.5}$$

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$,$T^{(1,1)}$, C_{11}, and C_{12}, we have the following expression for the lower bound $\ln \mathbb{P}_{L\{E,W\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \;\geq\;& \ln \mathbb{P}_{L\{E,W\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right) \\
\\
\equiv\;& \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) + T \;\left[\sigma_{\alpha 1}(0,0) + \sigma_{\alpha 2}(0,0)\right] \\
+& \left(y_{10} - y_{1T}\right) \Delta\sigma_{\alpha 1}(1,0) + \left(y_{20} - y_{2T}\right) \Delta\sigma_{\alpha 2}(0,1) \\
+& T_1^{(1)} \;\left[\alpha_1 + \Delta\sigma_{\alpha 1}(1,0)\right] + T_2^{(1)} \;\left[\alpha_2 + \Delta\sigma_{\alpha 2}(0,1)\right] \\
+& T_1^{(1)} \;\Delta\sigma_{\alpha 1}(0,1) + T_1^{(1)} \;\Delta\sigma_{\alpha 2}(1,0) \\
+& C_{11} \;\Delta^2\sigma_{\alpha 1} + C_{12} \;\Delta^2\sigma_{\alpha 2} \\
+& C_{11} \;\beta_1 + C_{22} \;\beta_2 + T_1^{(1)} \;\gamma_1 + T^{(1,1)} \;\gamma_2
\end{aligned} \tag{A.6}$$

Finally, using the definitions of $\mathbf{s}^1\left(\widetilde{\mathbf{y}}\right)' \mathbf{g}_{\boldsymbol{\alpha}}^1$ and $\mathbf{g}_{\boldsymbol{\alpha}}^2$, we get:

$$\begin{aligned}
\ln \mathbb{P}_{L\{E,W\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right) \;=\;& \mathbf{s}^1\left(\widetilde{\mathbf{y}}\right)' \mathbf{g}_{\boldsymbol{\alpha}}^1 + \left[T_1^{(1)},\, T_1^{(1)},\, C_{11},\, C_{12}\right] \;\mathbf{g}_{\boldsymbol{\alpha}}^2 \\
+& C_{11} \;\beta_1 + C_{22} \;\beta_2 + T_1^{(1)} \;\gamma_1 + T^{(1,1)} \;\gamma_2
\end{aligned} \tag{A.7}$$

(b) Lower Bound $\ln \mathbb{P}_{L\{S,N\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$. To obtain this lower bound, we use the bounds $L^{\{S,SE\}}(0,1|\mathbf{y}_{t-1};\boldsymbol{\alpha}) \equiv [1-\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1)]\ \Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)$ and $L^{\{N,NW\}}(1,0|\mathbf{y}_{t-1};\boldsymbol{\alpha}) \equiv \Lambda(\alpha_1+\beta_1\ y_{1t-1})$ $[1-\Lambda(\alpha_2+\beta_2\ y_{2t-1})]$. Then,

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \geq\ & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
+ & \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \ \ (\ln[1-\Lambda(\alpha_1+\beta_1\ y_{1t-1})] + \ln[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})]) \\
+ & \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \ (\ln[1-\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1)] + \ln\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)) \\
+ & \sum_{t=1}^{T} y_{1t}(1-y_{2t}) \ \ (\ln\Lambda(\alpha_1+\beta_1\ y_{1t-1}) + \ln[1-\Lambda(\alpha_2+\beta_2\ y_{2t-1})]) \\
+ & \sum_{t=1}^{T} y_{1t}y_{2t} \ \ (\ln\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1) + \ln\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2))
\end{aligned} \tag{A.8}$$

Using the definitions of $\sigma_{\alpha 1}(y_{1t-1}, y_{2t})$ and $\sigma_{\alpha 2}(y_{1t}, y_{2t-1})$, we have:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \geq\ & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
+ & \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1})] \\
+ & \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_2+\beta_2\ y_{2t-1}+\gamma_2] \\
+ & \sum_{t=1}^{T} y_{1t}(1-y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1}) + \alpha_1+\beta_1\ y_{1t-1}] \\
+ & \sum_{t=1}^{T} y_{1t}y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_1+\beta_1\ y_{1t-1}+\gamma_1+\alpha_2+\beta_2\ y_{2t-1}+\gamma_2]
\end{aligned} \tag*{A.9}$$

Grouping terms, we have:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \geq\ & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
+ & \sum_{t=1}^{T} (1-y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1})] \\
+ & \sum_{t=1}^{T} y_{1t} \ \ [\alpha_1+\beta_1\ y_{1t-1}] \\
+ & \sum_{t=1}^{T} y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_2+\beta_2\ y_{2t-1}+\gamma_2] \\
+ & \sum_{t=1}^{T} y_{1t}y_{2t} \ \ [\gamma_1]
\end{aligned} \tag{A.10}$$

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$,$T^{(1,1)}$, C_{11}, and C_{12}, we have the following expression for the lower bound $\ln \mathbb{P}_{L\{S,N\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$:

$$\begin{array}{rcl} \ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) & \geq & \ln \mathbb{P}_{L\{S,N\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right) \\ \\ & \equiv & \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) + T \;\; \left[\sigma_{\alpha 1}(0,0) + \sigma_{\alpha 2}(0,0)\right] \\ & + & \left(y_{10} - y_{1T}\right) \; \Delta\sigma_{\alpha 1}(1,0) + \left(y_{20} - y_{2T}\right) \; \Delta\sigma_{\alpha 2}(0,1) \\ & + & T_1^{(1)} \;\; \left[\alpha_1 + \Delta\sigma_{\alpha 1}(1,0)\right] + T_2^{(1)} \;\; \left[\alpha_2 + \Delta\sigma_{\alpha 2}(0,1)\right] \\ & + & T_2^{(1)} \;\; \left[\Delta\sigma_{\alpha 1}(0,1) + \Delta\sigma_{\alpha 2}(1,0)\right] \\ & + & C_{21} \; \Delta^2\sigma_{\alpha 1} + C_{22} \; \Delta^2\sigma_{\alpha 2} \\ & + & C_{11} \; \beta_1 + C_{22} \; \beta_2 + T^{(1,1)} \; \gamma_1 + T_2^{(1)} \; \gamma_2 \end{array} \tag{A.11}$$

Finally, using the definitions of $\mathbf{s}^1\left(\widetilde{\mathbf{y}}\right)' \mathbf{g}_{\boldsymbol{\alpha}}^1$ and $\mathbf{g}_{\boldsymbol{\alpha}}^2$, we get:

$$\begin{array}{rcl} \ln \mathbb{P}_{L\{S,N\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right) & = & \mathbf{s}^1\left(\widetilde{\mathbf{y}}\right)' \mathbf{g}_{\boldsymbol{\alpha}}^1 + \left[T_2^{(1)},\, T_2^{(1)},\, C_{21},\, C_{22}\right] \; \mathbf{g}_{\boldsymbol{\alpha}}^2 \\ & + & C_{11} \; \beta_1 + C_{22} \; \beta_2 + T^{(1,1)} \; \gamma_1 + T_2^{(1)} \; \gamma_2 \end{array} \tag{A.12}$$

(c) Lower Bound $\ln \mathbb{P}_{L\{E,N\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$. To obtain this lower bound, we use the bounds $L^{\{E,SE\}}(0,1|\mathbf{y}_{t-1};\boldsymbol{\alpha}) \equiv [1-\Lambda(\alpha_1+\beta_1\; y_{1t-1})]\; \Lambda(\alpha_2+\beta_2\; y_{2t-1})$ and $L^{\{N,NW\}}(1,0|\mathbf{y}_{t-1};\boldsymbol{\alpha}) \equiv \Lambda(\alpha_1+\beta_1\; y_{1t-1})\; [1-\Lambda(\alpha_2+\beta_2\; y_{2t-1})]$ for the choice probabilities. Then,

$$\begin{array}{rcl} \ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) & \geq & \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) \\ & + & \sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \;\; \left(\ln\left[1 - \Lambda\left(\alpha_1 + \beta_1 \; y_{1t-1}\right)\right] + \ln\left[1 - \Lambda\left(\alpha_2 + \beta_{22} \; y_{2t-1}\right)\right]\right) \\ & + & \sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \;\; \left(\ln\left[1 - \Lambda\left(\alpha_1 + \beta_1 \; y_{1t-1}\right)\right] + \ln\Lambda\left(\alpha_2 + \beta_2 \; y_{2t-1}\right)\right) \\ & + & \sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \;\; \left(\ln\Lambda\left(\alpha_1 + \beta_1 \; y_{1t-1}\right) + \ln\left[1 - \Lambda\left(\alpha_2 + \beta_2 \; y_{2t-1}\right)\right]\right) \\ & + & \sum_{t=1}^{T} y_{1t} y_{2t} \;\; \left(\ln\Lambda\left(\alpha_1 + \beta_1 \; y_{1t-1} + \gamma_1\right) + \ln\Lambda\left(\alpha_2 + \beta_2 \; y_{2t-1} + \gamma_2\right)\right) \end{array} \tag{A.13}$$

Using the definitions of $\sigma_{\alpha 1}(y_{1t-1}, y_{2t})$ and $\sigma_{\alpha 2}(y_{1t}, y_{2t-1})$, we have:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \geq & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
+ & \sum_{t=1}^{T} (1 - y_{1t})(1 - y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\
+ & \sum_{t=1}^{T} (1 - y_{1t}) y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_2 + \beta_2 \ y_{2t-1}] \\
+ & \sum_{t=1}^{T} y_{1t}(1 - y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_1 + \beta_1 \ y_{1t-1}] \\
+ & \sum_{t=1}^{T} y_{1t} y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_1 \ y_{1t-1} + \gamma_1 + \alpha_2 + \beta_2 \ y_{2t-1} + \gamma_2]
\end{aligned} \tag{A.14}$$

Grouping terms, we have:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \geq & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
+ & \sum_{t=1}^{T} (1 - y_{1t} y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\
+ & \sum_{t=1}^{T} y_{2t} \ \ [\alpha_2 + \beta_2 \ y_{2t-1}] \\
+ & \sum_{t=1}^{T} y_{1t} \ \ [\alpha_1 + \beta_1 \ y_{1t-1}] \\
+ & \sum_{t=1}^{T} y_{1t} y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \gamma_1 + \gamma_2]
\end{aligned} \tag{A.15}$$

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$,$T^{(1,1)}$, C_{11}, and C_{12}, we have the following expression for the lower bound $\ln \mathbb{P}_{L\{E,N\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \geq & \ln \mathbb{P}_{L\{E,N\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) \\
\equiv & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) + T \ \ [\sigma_{\alpha 1}(0,0) + \sigma_{\alpha 2}(0,0)] \\
+ & (y_{10} - y_{1T}) \ \Delta\sigma_{\alpha 1}(1,0) + (y_{20} - y_{2T}) \ \Delta\sigma_{\alpha 2}(0,1) \\
+ & T_1^{(1)} \ \ [\alpha_1 + \Delta\sigma_{\alpha 1}(1,0)] + T_2^{(1)} \ \ [\alpha_2 + \Delta\sigma_{\alpha 2}(0,1)] \\
+ & T^{(1,1)} \ \ [\Delta\sigma_{\alpha 1}(0,1) + \Delta\sigma_{\alpha 2}(1,0)] \\
+ & R_1^{(1,1)} \ \Delta^2\sigma_{\alpha 1} + R_2^{(1,1)} \ \Delta^2\sigma_{\alpha 2} \\
+ & C_{11} \ \beta_1 + C_{22} \ \beta_2 + T^{(1,1)} \ \ [\gamma_1 + \gamma_2]
\end{aligned} \tag{A.16}$$

Finally, using the definitions of $\mathbf{s}^1(\widetilde{\mathbf{y}})' \, \mathbf{g}_{\boldsymbol{\alpha}}^1$ and $\mathbf{g}_{\boldsymbol{\alpha}}^2$, we get:

$$\begin{aligned}
\ln \mathbb{P}_{L\{E,N\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) = & \mathbf{s}^1(\widetilde{\mathbf{y}})' \, \mathbf{g}_{\boldsymbol{\alpha}}^1 + \left[T^{(1,1)}, \, T^{(1,1)}, \, R_1^{(1,1)}, \, R_2^{(1,1)}\right] \ \mathbf{g}_{\boldsymbol{\alpha}}^2 \\
+ & C_{11} \ \beta_1 + C_{22} \ \beta_2 + T^{(1,1)} \ \ [\gamma_1 + \gamma_2]
\end{aligned} \tag{A.17}$$

(d) Lower Bound $\ln \mathbb{P}_{L\{S,W\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$. To obtain this lower bound, we use the bounds $L^{\{S,SE\}}(0,1|\mathbf{y}_{t-1};\boldsymbol{\alpha}) \equiv [1-\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1)]\ \Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)$ and $L^{\{W,NW\}}(1,0|\mathbf{y}_{t-1};\boldsymbol{\alpha}) \equiv \Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1)$ $[1-\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)]$ for the choice probabilities. Then,

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \ \geq\ & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
+ & \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \ \left(\ln\left[1-\Lambda(\alpha_1+\beta_1\ y_{1t-1})\right] + \ln\left[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})\right]\right) \\
+ & \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \left(\ln\left[1-\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1)\right] + \ln\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)\right) \\
+ & \sum_{t=1}^{T} y_{1t}(1-y_{2t}) \ \left(\ln\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1) + \ln\left[1-\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)\right]\right) \\
+ & \sum_{t=1}^{T} y_{1t}y_{2t} \ \left(\ln\Lambda(\alpha_1+\beta_1\ y_{1t-1}+\gamma_1) + \ln\Lambda(\alpha_2+\beta_2\ y_{2t-1}+\gamma_2)\right)
\end{aligned} \tag{A.18}$$

Using the definitions of $\sigma_{\alpha 1}(y_{1t-1}, y_{2t})$ and $\sigma_{\alpha 2}(y_{1t}, y_{2t-1})$, we have:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \ \geq\ & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
+ & \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \ \left[\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1})\right] \\
+ & \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \left[\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_2+\beta_2\ y_{2t-1}+\gamma_2\right] \\
+ & \sum_{t=1}^{T} y_{1t}(1-y_{2t}) \ \left[\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_1+\beta_1\ y_{1t-1}+\gamma_1\right] \\
+ & \sum_{t=1}^{T} y_{1t}y_{2t} \ \left[\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_1+\beta_1\ y_{1t-1}+\gamma_1+\alpha_2+\beta_2\ y_{2t-1}+\gamma_2\right]
\end{aligned} \tag{A.19}$$

Grouping terms, we have:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) \ \geq\ & \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
+ & \sum_{t=1}^{T} \left[\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1})\right] \\
+ & \sum_{t=1}^{T} y_{2t} \ \left[\alpha_2+\beta_2\ y_{2t-1}+\gamma_2\right] \\
+ & \sum_{t=1}^{T} y_{1t} \ \left[\alpha_1+\beta_1\ y_{1t-1}+\gamma_1\right] \\
+ & \sum_{t=1}^{T} \left[y_{1t}+y_{2t}-y_{1t}y_{2t}\right] \ \left[\sigma_{\alpha 1}(y_{1t-1},1) - \sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(1,y_{2t-1}) - \sigma_{\alpha 2}(0,y_{2t-1})\right]
\end{aligned} \tag{A.20}$$

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$,$T^{(1,1)}$, C_{11}, and C_{12}, we have the following expression for the lower bound $\ln \mathbb{P}_{L\{S,W\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$:

$$
\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) &\geq \ln \mathbb{P}_{L\{S,W\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) \\
&\equiv \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) + T\ \left[\sigma_{\alpha 1}(0,0) + \sigma_{\alpha 2}(0,0)\right] \\
&+ (y_{10} - y_{1T})\ \left[\sigma_{\alpha 1}(1,0) - \sigma_{\alpha 1}(0,0)\right] + (y_{20} - y_{2T})\ \left[\sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0)\right] \\
&+ T_1^{(1)}\ \left[\alpha_1 + \Delta\sigma_{\alpha 1}(1,0)\right] + T_2^{(1)}\ \left[\alpha_2 + \Delta\sigma_{\alpha 2}(0,1)\right] \\
&+ \left[T_1^{(1)} + T_2^{(1)} - T^{(1,1)}\right]\ \left[\Delta\sigma_{\alpha 1}(0,1) + \Delta\sigma_{\alpha 2}(1,0)\right] \\
&+ \left[C_{11} + C_{21} - R_1^{(1,1)}\right]\ \Delta^2\sigma_{\alpha 1} + \left[C_{12} + C_{22} - R_2^{(1,1)}\right]\ \Delta^2\sigma_{\alpha 2} \\
&+ C_{11}\ \beta_1 + C_{22}\ \beta_2 + T_1^{(1)}\ \gamma_1 + T_2^{(1)}\ \gamma_2
\end{aligned}
\tag{A.21}
$$

Finally, using the definitions of $\mathbf{s}^1(\widetilde{\mathbf{y}})'\,\mathbf{g}_{\boldsymbol{\alpha}}^1$ and $\mathbf{g}_{\boldsymbol{\alpha}}^2$, we get:

$$
\begin{aligned}
\ln \mathbb{P}_{L\{S,W\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) &= \mathbf{s}^1(\widetilde{\mathbf{y}})'\,\mathbf{g}_{\boldsymbol{\alpha}}^1 \\
&+ \left[T_1^{(1)} + T_2^{(1)} - T^{(1,1)},\ T_1^{(1)} + T_2^{(1)} - T^{(1,1)},\ C_{11} + C_{21} - R_1^{(1,1)},\ C_{12} + C_{22} - R_2^{(1,1)}\right]\ \mathbf{g} \\
&+ C_{11}\ \beta_1 + C_{22}\ \beta_2 + T_1^{(1)}\ \gamma_1 + T_2^{(1)}\ \gamma_2
\end{aligned}
\tag{A.22}
$$

(e) Upper Bound $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$. For the upper bounds, we use the bounds for the choice probabilities $U(0,1|\mathbf{y}_{t-1}; \boldsymbol{\alpha}) \equiv \left[1 - \Lambda(\alpha_1 + \beta_1\ y_{1t-1} + \gamma_1)\right]\ \Lambda(\alpha_2 + \beta_2\ y_{2t-1})$ and $U(1,0|\mathbf{y}_{t-1}; \boldsymbol{\alpha}) \equiv \Lambda(\alpha_1 + \beta_1\ y_{1t-1})\ \left[1 - \Lambda(\alpha_2 + \beta_2\ y_{2t-1} + \gamma_2)\right]$. Then,

$$
\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}) &\leq \ln p_{\boldsymbol{\alpha}}(y_{10}, y_{20}) \\
&+ \sum_{t=1}^{T} (1 - y_{1t})(1 - y_{2t})\ \left(\ln\left[1 - \Lambda(\alpha_1 + \beta_1\ y_{1t-1})\right] + \ln\left[1 - \Lambda(\alpha_2 + \beta_{22}\ y_{2t-1})\right]\right) \\
&+ \sum_{t=1}^{T} (1 - y_{1t})\, y_{2t}\ \left(\ln\left[1 - \Lambda(\alpha_1 + \beta_1\ y_{1t-1} + \gamma_1)\right] + \ln \Lambda(\alpha_2 + \beta_2\ y_{2t-1})\right) \\
&+ \sum_{t=1}^{T} y_{1t}(1 - y_{2t})\ \left(\ln \Lambda(\alpha_1 + \beta_1\ y_{1t-1}) + \ln\left[1 - \Lambda(\alpha_2 + \beta_2\ y_{2t-1} + \gamma_2)\right]\right) \\
&+ \sum_{t=1}^{T} y_{1t} y_{2t}\ \left(\ln \Lambda(\alpha_1 + \beta_1\ y_{1t-1} + \gamma_1) + \ln \Lambda(\alpha_2 + \beta_2\ y_{2t-1} + \gamma_2)\right)
\end{aligned}
\tag{A.23}
$$

Using the definitions of $\sigma_{\alpha 1}(y_{1t-1}, y_{2t})$ and $\sigma_{\alpha 2}(y_{1t}, y_{2t-1})$, we have:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \leq\ & \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) \\
+\ & \sum_{t=1}^{T} \left(1-y_{1t}\right)\left(1-y_{2t}\right)\ \left[\sigma_{\alpha 1}(y_{1t-1}, 0)+\sigma_{\alpha 2}(0, y_{2t-1})\right] \\
+\ & \sum_{t=1}^{T} \left(1-y_{1t}\right) y_{2t}\ \left[\sigma_{\alpha 1}(y_{1t-1}, 1)+\sigma_{\alpha 2}(0, y_{2t-1})+\alpha_2+\beta_2\ y_{2t-1}\right] \\
+\ & \sum_{t=1}^{T} y_{1t}\left(1-y_{2t}\right)\ \left[\sigma_{\alpha 1}(y_{1t-1}, 0)+\sigma_{\alpha 2}(1, y_{2t-1})+\alpha_1+\beta_1\ y_{1t-1}\right] \\
+\ & \sum_{t=1}^{T} y_{1t} y_{2t}\ \left[\sigma_{\alpha 1}(y_{1t-1}, 1)+\sigma_{\alpha 2}(1, y_{2t-1})+\alpha_1+\beta_1\ y_{1t-1}+\gamma_1+\alpha_2+\beta_2\ y_{2t-1}+\gamma_2\right]
\end{aligned} \tag{A.24}$$

Grouping terms, we have:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \leq\ & \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) \\
+\ & \sum_{t=1}^{T} \left(1-y_{2t}\right)\ \sigma_{\alpha 1}(y_{1t-1}, 0)+\left(1-y_{1t}\right)\ \sigma_{\alpha 2}(0, y_{2t-1}) \\
+\ & \sum_{t=1}^{T} y_{2t}\ \sigma_{\alpha 1}(y_{1t-1}, 1)+y_{2t}\ \left[\alpha_2+\beta_2\ y_{2t-1}\right] \\
+\ & \sum_{t=1}^{T} y_{1t}\ \sigma_{\alpha 2}(1, y_{2t-1})+y_{1t}\ \left[\alpha_1+\beta_1\ y_{1t-1}\right] \\
+\ & \sum_{t=1}^{T} y_{1t} y_{2t}\ \left[\gamma_1+\gamma_2\right]
\end{aligned} \tag{A.25}$$

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$,$T^{(1,1)}$, C_{11}, and C_{12}, we have the following expression for the upper bound $\ln \mathbb{P}_U\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \leq\ & \ln \mathbb{P}_U\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right) \\
\equiv\ & \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right)+T\ \left[\sigma_{\alpha 1}(0,0)+\sigma_{\alpha 2}(0,0)\right] \\
+\ & \left(y_{10}-y_{1T}\right)\ \Delta\sigma_{\alpha 1}(1,0)+\left(y_{20}-y_{2T}\right)\ \Delta\sigma_{\alpha 2}(0,1) \\
+\ & T_1^{(1)}\ \left[\alpha_1+\Delta\sigma_{\alpha 1}(1,0)\right]+T_2^{(1)}\ \left[\alpha_2+\Delta\sigma_{\alpha 2}(0,1)\right] \\
+\ & T_2^{(1)}\ \Delta\sigma_{\alpha 1}(0,1)+T_1^{(1)}\ \Delta\sigma_{\alpha 2}(1,0) \\
+\ & C_{21}\ \Delta^2\sigma_{\alpha 1}+C_{12}\ \Delta^2\sigma_{\alpha 2} \\
+\ & C_{11}\ \beta_1+C_{22}\ \beta_2+T^{(1,1)}\ \left[\gamma_1+\gamma_2\right]
\end{aligned} \tag{A.26}$$

Finally, using the definitions of $\mathbf{s}^1\left(\widetilde{\mathbf{y}}\right)'\,\mathbf{g}^1_{\boldsymbol{\alpha}}$ and $\mathbf{g}^2_{\boldsymbol{\alpha}}$, we get:

$$\begin{aligned}
\ln \mathbb{P}_U\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right) =\ & \mathbf{s}^1\left(\widetilde{\mathbf{y}}\right)'\,\mathbf{g}^1_{\boldsymbol{\alpha}}+\left[T_2^{(1)},\, T_1^{(1)},\, C_{21},\, C_{12}\right]\ \mathbf{g}^2_{\boldsymbol{\alpha}} \\
+\ & C_{11}\ \beta_1+C_{22}\ \beta_2+T^{(1,1)}\ \left[\gamma_1+\gamma_2\right]
\end{aligned} \tag{A.27}$$

References

Aguirregabiria, V., J. Gu, and Y. Luo (2021): "Sufficient statistics for unobserved heterogeneity in dynamic structural logit models," *Journal of Econometrics*, 223(2), 280–311.

Aguirregabiria, V., and C.-Y. Ho (2012): "A dynamic oligopoly game of the US airline industry: Estimation and policy experiments," *Journal of Econometrics*, 168(1), 156–173.

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

Andersen, E. (1970): "Asymptotic properties of conditional maximum likelihood estimators," *Journal of the Royal Statistical Society, Series B*, 32, 283–301.

Arcidiacono, P., and R. A. Miller (2011): "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity," *Econometrica*, 79(6), 1823–1867.

Blume, L., W. Brock, S. Durlauf, and Y. Ioannides (2011): "Identification of social interactions," in *Handbook of Social Economics, Vol. 1*, pp. 853–964. North-Holland.

Bonhomme, S. (2012): "Functional differencing," *Econometrica*, 80(4), 1337–1385.

Bresnahan, T., and P. Reiss (1991): "Empirical models of discrete games," *Journal of Econometrics*, 48(1), 57–81.

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

Chamberlain, G. (1980): "Analysis of covariance with qualitative data," *Review of Economic Studies*, 47(1), 225–238.

——— (2010): "Binary response models for panel data: Identification and information," *Econometrica*, 78(1), 159–168.

Ching, A. T. (2010): "A dynamic oligopoly structural model for the prescription drug market after patent expiration," *International Economic Review*, 51(4), 1175–1207.

Collard-Wexler, A. (2013): "Demand fluctuations in the ready-mix concrete industry," *Econometrica*, 81(3), 1003–1037.

Cox, D. (1958): "The regression analysis of binary sequences," *Journal of the Royal Statistical Society. Series B*, 20(2), 215–242.

Dobronyi, C., J. Gu, and K. Kim (2021): "Identification of Dynamic Panel Logit Models with Fixed Effects," *arXiv:2104.04590*.

Dunne, T., S. D. Klimek, M. J. Roberts, and D. Y. Xu (2013): "Entry, exit, and the determinants of market structure," *The RAND Journal of Economics*, 4(3), 462–487.

Eckstein, Z., and O. Lifshitz (2015): "Household interaction and the labor supply of married women," *International Economic Review*, 56(2), 427–455.

Einav, L. (2010): "Not all rivals look alike: estimating an equilibrium model of the release date timing game," *Economic Inquiry*, 48(2), 369–390.

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

Gallant, A. R., H. Hong, and A. Khwaja (2018): "A Bayesian approach to estimation of dynamic models with small and large number of heterogeneous players and latent serially correlated states," *Journal of Econometrics*, 203(1), 19–32.

Goettler, R. L., and B. R. Gordon (2011): "Does AMD spur Intel to innovate more?," *Journal of Political Economy*, 119(6), 1141–1200.

Groeger, J. R. (2014): "A study of participation in dynamic auctions," *International Economic Review*, 55(4), 1129–1154.

Hashmi, A., and J. V. Biesebroeck (2016): "The relationship between market structure and innovation in industry equilibrium: a case study of the global automobile industry," *The Review of Economics and Statistics*, 98(1), 192–208.

Heckman, J. (1981): "The incidental parameters problem and the problem of initial conditions in estimating a discrete time - discrete data stochastic process," in *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press.

Honoré, B., and A. De Paula (2021): "Identification in simple binary outcome panel data models," *The Econometrics Journal*, 24(2), C78–C93.

Honoré, B., and E. Kyriazidou (2019): "Identification in Binary Response Panel Data Models: Is Point-Identification More Common Than We Thought?," *Annals of Economics and Statistics*, 134, 207–226.

Honoré, B., and M. Weidner (2020): "Dynamic Panel Logit Models with Fixed Effects," *arXiv preprint arXiv:2005.05942.*

Huang, L., and M. D. Smith (2014): "The dynamic efficiency costs of common-pool resource exploitation," *American Economic Review*, 104(12), 4071–4103.

Igami, M. (2017): "Estimating the innovator's dilemma: structural analysis of creative destruction in the hard disk drive industry, 1981–1998," *Journal of Political Economy*, 125(3), 798–847.

Igami, M., and N. Yang (2016): "Unobserved heterogeneity in dynamic games: Cannibalization and preemptive entry of hamburger chains in Canada," *Quantitative Economics*, 7(2), 483–521.

Jeziorski, P. (2014): "Estimation of cost efficiencies from mergers: application to US radio," *The RAND Journal of Economics*, 45(4), 816–846.

Jofre-Bonet, M., and M. Pesendorfer (2003): "Estimation of a dynamic auction game," *Econometrica*, 71(5), 1443–1489.

Kalouptsidi, M. (2014): "Time to build and fluctuations in bulk shipping," *American Economic Review*, 104(2), 564–608.

Kasahara, H., and K. Shimotsu (2009): "Nonparametric identification of finite mixture models of dynamic discrete choices," *Econometrica*, 77(1), 135–175.

Narendranthan, W., S. Nickell, and D. Metcalf (1985): "An investigation into the incidence and dynamic structure of sickness and unemployment in Britain, 1965–75," *Journal of the Royal Statistical Society: Series A*, 148(3), 254–267.

Rasch, G. (1961): "On General Laws and the Meaning of Measurement in Psychology," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 4, 321–333.

Ryan, S. P. (2012): "The costs of environmental regulation in a concentrated industry," *Econometrica*, 80(3), 1019–1061.

Sieg, H., and C. Yoon (2017): "Estimating dynamic games of electoral competition to evaluate term limits in us gubernatorial elections," *American Economic Review*, 107(7), 1824–1857.

Sovinsky, M., and S. Stern (2016): "Dynamic modelling of long-term care decisions," *Review of Economics of the Household*, 14(2), 463–488.

Suzuki, J. (2013): "Land use regulation as a barrier to entry: evidence from the texas lodging industry," *International Economic Review*, 54(2), 495–523.

Sweeting, A. (2013): "Dynamic product positioning in differentiated product markets: the effect of fees for musical performance rights on the commercial radio industry," *Econometrica*, 81(5), 1763–1803.

Takahashi, Y. (2015): "Estimating a war of attrition: the case of the US movie theater industry," *American Economic Review*, 105(7), 2204–2241.

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

Wagner, U. (2016): "Estimating strategic models of international treaty formation," *The Review of Economic Studies*, 83(4), 1741–1778.