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

Victor Aguirregabiria
University of Toronto and CEPR

Jiaying Gu
University of Toronto

Pedro Mira
CEMFI

March 3, 2021

Abstract

In the structural estimation of dynamic discrete choice games, the mis-specification of unobserved heterogeneity can generate important biases in two types of structural parameters: those that capture dynamic state dependence – e.g., costs of adjustment or switching – and those that represent strategic interactions between players – e.g., competition and peer effects. We study the identification of these parameters in models where market unobserved heterogeneity has a fixed effect structure – a nonparametric distribution conditional on the initial values of the state variables. We extend the sufficient statistics - conditional likelihood approach (Chamberlain, 1985) to dynamic games with multiple equilibibria with partial identification. We present identification results for different types of games according to complete vs. incomplete information, Stackelberg vs. full strategic interactions, and myopic vs. forward-looking players.

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

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

Victor Aguirregabiria. Address: 150 St. George Street. Toronto, ON, M5S 3G7, Canada.
E-mail: `victor.aguirregabiria@utoronto.ca`

Jiaying Gu. Address: 150 St. George Street. Toronto, ON, M5S 3G7, Canada.
E-mail: `jiaying.gu@utoronto.ca`

Pedro Mira. Address: Casado del Alisa, 5. Madrid, Spain.
E-mail: `mira@cemfi.es`

*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.

1 Introduction

Dynamic games are useful tools for the analysis of economic and social phenomena characterized by intertemporal interactions between agents. The structural estimation of dynamic games has received notorious attention in the study of dynamics of oligopoly competition (Ericson and Pakes, 1995) with empirical applications to different industries.[1] These econometric models have been recently applied also to study dynamic interactions within households (Eckstein and Lifshitz, 2015), long-term care decisions (Sovinsky and Stern, 2016), electoral competition (Sieg and Yoon, 2017), or the ratification of international treaties (Wagner, 2016), among other topics. There is also a substantial literature on dynamic discrete choice models with social interactions where agents are not forward looking (Brock and Durlauf, 2007; Blume, et al, 2011).

Two types of structural parameters play a fundamental role in the predictions of dynamic games: the parameters that capture dynamic state dependence, such as costs of switching, adjustment, investment, or entry and exit, (i.e., the *dynamic* part of the model); and the parameters that represent the effects of other players' actions on a player's payoff due to competition, spillovers, peer effects, or social interactions (i.e., the *game* part of the model). The identification of these two types of parameters rely crucially on the model assumptions about the stochastic properties of variables that are known to the players but unobservable to the researcher, i.e., what we can denote as the specification of *unobserved heterogeneity*. In dynamic models, it is well known that ignoring or misspecifying persistent unobserved heterogeneity can imply substantial biases in the estimation of structural parameters that capture true dynamics (Heckman, 1981). *Spurious dynamics* due to unobserved heterogeneity can be confounded with *true dynamics* due to state dependence. In the literature on the estimation of games, it is well-known that ignoring unobserved heterogeneity that is common or correlated across players can generate important biases in the estimation of the structural parameters that capture strategic (competition) or social (peer) interactions between players (Bresnahan and Reiss, 1991; Blume, et al, 2011). The common unobserved heterogeneity can be confounded with positive strategic, social, or peer effects.

In this paper, we study the identification and estimation of dynamic games where N players are observed playing the game over M markets and T periods of time, where the time dimension T

[1] Recent applications include industries such as automobiles (Hashmi and Van Biesebroeck, 2016), airlines (Aguirregabiria and Ho, 2012), pharmaceuticals (Ching, 2010, Gallant et al., 2017), procurement auctions (Jofre-Bonet and Pesendorfer, 2003, Groeger, 2014), construction materials (Ryan, 2012, Collard-Wexler, 2013), hotels (Suzuki, 2013), micro processors (Goetler and Gordon, 2011), hard drives (Igami, 2017), comercial radio (Sweeting, 2013, Jeziorski, 2014), movies (Einav, 2010, Takahashi, 2015), medical services (Dunne et al., 2013), ship building (Kalouptsidi, 2014), fishing (Huang and Smith, 2014), or retail stores (Aguirregabiria and Mira, 2007, Toivanen and Waterson, 2011, Igami and Yang, 2016), among others.

is small and the number of markets M is large. Most of our results apply to either a small or large number of players N. However, for concreteness and notational simplicity, we focus on the case of small N. We consider the identification of these models when there is time-invariant unobserved heterogeneity either at the market or player-market level, and the probability distribution of this heterogeneity and the initial conditions of the endogenous state variables is nonparametrically specified, i.e., a fixed effects panel data model.

The identification of this model should deal with the *incidental parameters problem*, and the *initial conditions problem*. The *incidental parameters problem* establishes that a simple dummy-variables estimator, that treats each market (or player-market) unobservable as a parameter to be estimated jointly with the parameters of interest, is inconsistent in nonlinear dynamic panel data models when T is fixed (Neyman and Scott, 1948, Lancaster, 2000). The *initial conditions problem* establishes that the joint distribution of the unobserved heterogeneity and the initial values of the endogenous state variables is not nonparametrically identified, but the misspecification of this joint distribution can generate important biases in the estimation of the parameters of interest (Heckman, 1981, Chamberlain, 1985, among others). That is, any approach that tries to jointly identify structural parameters and distribution of unobserved heterogeneity and initial conditions needs to impose restrictions on this distribution. In contrast, a fixed effects approach, as the one considered in this paper, is concerned with the identification of structural parameters but not with the identification of the distribution of the unobserved heterogeneity.

In this paper, we extend to dynamic discrete choice games the fixed effect conditional likelihood method pioneered by Cox (1958), Rasch (1961), Andersen (1970), and Chamberlain (1980). This approach is based on the derivation of sufficient statistics for the incidental parameters (for the fixed effects) and the maximization of a likelihood function of the data conditional on these sufficient statistics. The main advantage of this approach is that the estimation of the structural parameters is robust to any misspecification of the distribution of the unobserved heterogeneity. For those

This paper is related to the literature on identification and estimation of panel data dynamic discrete choice model with a fixed effects structure, and the seminal work by Chamberlain (1985) and Honoré and Kyriazidou (2000). More recently, Aguirregabiria, Gu, and Luo (2019) have applied this approach to the identification single-agent dynamic discrete choice structural models. Honoré and Kyriazidou (2017) study fixed effects panel data discrete choice VAR models. We build on the results in Honoré and Kyriazidou (2017) and extend their results in two directions. First, we study the identification of games where players are forward-looking. In these models, players decisions depend on continuation values which are nonlinear functions o unobserved heterogeneity and state variables. Controlling for these continuation values can complicate obtaining sufficient statistics for the the incidental parameters in the unobserved heterogeneity. Second, we study the identification of games with multiple equilibria. The model does not provide a unique prediction for the probability of a choice history but only bounds. We show how to obtain sufficient statistics for the effects of incidental parameters in these bounds, and how this implies the partial identification of the structural parameters. As far as we know, this is the first paper that combines the fixed effects - sufficient statistics approach with bounds and partial identification.

Furthermore, for those versions of the model where the conditional likelihood method cannot identify all the structural parameters, we consider a functional differencing method proposed by Bonhomme (2012). More specifically, we consider a version of functional differencing recently proposed by Dobroyni, Gu and Kim (2021) that is also similar to the one Honoré and Weidner (2020). This approach is based on the derivation of a general class of moment conditions and moment inequalities implied by the fixed effects dynamic model. We show that this alternative approach identifies some important parameters which are not identified using a conditional likelihood method.

Our paper also contributes to the literature of identification and estimation of dynamic games with unobserved heterogeneity. All the papers in this literature have considered a random effects approach with a finite mixture specification of the unobserved heterogeneity. Following an approach proposed by Heckman (1981), Aguirregabiria and Mira (2007) deal with the initial conditions problem by assuming that the initial conditions come from the market-type-specific ergodic distribution of the endogenous variables. Arcidiacono and Miller (2011) propose an Expectation-Maximization (EM) algorithm to deal with the computational burdens of estimating this type of models. Kasahara and Shimotsu (2009) show the nonparametric identification of choice probabilities and the distribution of unobserved heterogeneity in finite mixture models. Igami and Yang (2016) combine the results and methods in Arcidiacono and Miller (2011) and Kasahara and Shimotsu (2009) and apply them to the estimation of dynamic game of market entry in the Canadian fast food restaurant industry. All these previous results are based on the assumption that unobserved heterogeneity has a finite mixture distribution, and there are some restrictions on the initial conditions. Our model does not impose these restrictions.

In a recent paper, Berry and Compiani (2020) propose a two-step IV approach for the estimation of discrete choice dynamic structural models with serially correlated unobservables. There are complementarities between their model and method and our approach. Their method allows for time variant serially correlated unobservables while our approach imposes the restriction that the peristent component of unobservables is time invariant. However, their approach is based on a parametric specification for the stochastic process of the unobservables. They study two types of parametric models for the unobservables: a linear AR(1) model, and a Markov chain with discrete support. In contrast, we are interested in identification results under a fixed effects model that does not impose any restriction on the joint distribution of the time-invariant unobservables and the initial values of the state variables. Furthermore, for the case of dynamic games, they impose the restriction of equilibrium uniqueness.

Table 1 provides a summary – or road map – of our identification results. We study the identification of eight different models that result from the combination of three criteria: (i) players' discounting of future payoffs – myopic vs. forward-looking model; (ii) players' information – complete vs. incomplete information; and (iii) strategic interactions between players – Stackelberg model vs. full strategic interactions. In terms of the identification results, we distinguish two sets of parameters: switching cost (β) and strategic interactions (γ). We also distinguish between point and partial (set) identification.

Table 1
Summary of Identification Results: Eight Different Models(1)

	Complete		**Incomplete**	
	Stackelberg	**Full**	**Stackelberg**	**Full**
Myopic	Point identification of β_{11}, β_{22}, γ_2 (Proposition 3)	Partial identification of β_{11}, β_{22}, γ_1, γ_2 (Proposition 8)	Point identification of β_{11}, β_{22} Partial iden. γ_2 (Propositions 5,6)	Partial identification of β_{11}, β_{22}, γ_1, γ_2 (Proposition 9)
Forward Looking	Point identification of β_{11}, β_{22} (Proposition 10)	Partial identification of β_{11}, β_{22} (Proposition 11)	(Proposition 12)	(Proposition 13)

Note (1): Parameters $\beta_{11}$and β_{22} represent switching costs. Parameters γ_1 and γ_2 represent strategic / competition effects.

The rest of the paper is organized as follows. In section 2, we study the identification of *myopic* dynamic discrete choice games. These models are interesting by themselves, but they are also useful because they provide results that we use in the identification of dynamic games when players are forward-looking. Section 3 presents identification results for dynamic games when players are forward-looking. In sections 2 and 3 we distinguish different type of models according to the information structure (complete vs. incomplete information) and the nature of the strategic interactions (contemporaneous or lagged). Section 4 deals with estimation and inference. In section 5, we illustrate our identification results in the context of an empirical application. We summarize and conclude in section 6.

2 Identification of dynamic games with myopic players

2.1 Framework

We index players by $i, j \in \{1, 2\}$. Subindex m represents a realization of the *market* where a pair of agents play the game.[2] Time is discrete and indexed by t that belongs to $\{1, 2, ..., T\}$. Every period t, players simultaneously make a binary decision, that we represent using variables $y_{1mt} \in \{0, 1\}$ and $y_{2mt} \in \{0, 1\}$. Players maximize their respective expected utility in the market, $\mathbb{E}[U_{imt}(y_{imt}, y_{jmt}) \mid \mathcal{I}_{imt}]$, where $\mathcal{I}_{it}$ is the information set of player i in market m at period t, and $U_{imt}(y_i, y_j)$ is her utility if she chooses action y_i and the other player chooses y_j. This utility function has the following standard structure:

$$U_{imt} = \widetilde{\alpha}_{im}\left(y_{imt}\right) + \widetilde{\gamma}_i\left(y_{imt}, y_{jmt}\right) + \widetilde{\beta}_{ii}\left(y_{imt}, y_{im,t-1}\right) + \widetilde{\beta}_{ij}\left(y_{imt}, y_{jm,t-1}\right) + \widetilde{\varepsilon}_{imt}(y_{it}). \tag{1}$$

$\widetilde{\alpha}_{im}(0)$ and $\widetilde{\alpha}_{im}(1)$ are market and player characteristics that are unobservable to the researcher. The vector $\alpha_m \equiv \{\widetilde{\alpha}_{1m}(j),\ \widetilde{\alpha}_{2m}(j) : j \in \{0, 1\}\}$ represents the fixed effects for market m. The vector of fixed effects $\alpha \equiv \{\alpha_m : m = 1, 2, ..., M\}$ represents the *incidental parameters* of the model. Function $\widetilde{\gamma}_i\left(y_{imt}, y_{jmt}\right)$ represents the contemporaneous competition (strategic) effects. Function $\widetilde{\beta}_{ii}$ captures state dependence with respect the lagged value of the player's own action, e.g., adjustment costs, switching costs. Function $\widetilde{\beta}_{ij}$ represents state dependence with respect to the lagged value of the other player's action, i.e., dynamic strategic interactions between the two players. The unobservable variables $\{\widetilde{\varepsilon}_{imt}(a) : a \in \{0, 1\}\}$ are *i.i.d.* over (i, m, t, a) with a extreme value type I distribution.

The model can be extended to incorporate exogenous state variables, $\mathbf{x}_{mt}$. For instance, the functions $\widetilde{\gamma}_i$, $\widetilde{\beta}_{ii}$, and $\widetilde{\beta}_{ij}$ may depend on $\mathbf{x}_t$, or in a more parsimonious specification, we can add a term $\delta_i(y_{imt})'\mathbf{x}_{mt}$ to the utility function in equation (1). We also present identification results for the parameters δ_i.

We distinguish two types of models according to players' information structure: *complete information* and *incomplete information* games. The following assumptions about information structure

[2] Depending on the empirical application, the particular definition of market can be a geographic location (e.g., a city, a neighborhood), a school, a family, an industry, an election, etc.

are common to the two models: (a) lagged choices (y_{imt-1}, y_{jmt-1}), current exogenous state variables $\mathbf{x}_{mt}$, parameters $\{\widetilde{\alpha}_{im}, \widetilde{\alpha}_{jm}; \widetilde{\gamma}_i, \widetilde{\gamma}_j; \widetilde{\beta}_{ii}, \widetilde{\beta}_{jj}; \widetilde{\beta}_{ij}, \widetilde{\beta}_{ji}\}$, the probability distribution of $\widetilde{\varepsilon}$, and the transition probability function for $\mathbf{x}_{mt}$ are known to all the players at period t; and (2) future values of the exogenous state variables, $\mathbf{x}_{m,t+s}$ and $(\widetilde{\varepsilon}_{im,t+s}, \widetilde{\varepsilon}_{jm,t+s})$ for $s \geq 1$, are unknown to all the players. The difference between the two models is in the treatment of the current values of the $\widetilde{\varepsilon}$ variables. For *games of complete information* we assume that the $\widetilde{\varepsilon}$ variables are common knowledge to all the players. For *games of incomplete information* we assume that variables $\widetilde{\varepsilon}_{imt}(0)$ and $\widetilde{\varepsilon}_{imt}(1)$ are private information of firm i but unknown to the other players in the game.

Let ΔU_{imt} represent the utility difference $U_{imt}(1, y_{jmt}) - U_{imt}(0, y_{jmt})$. Since we can only identify parameters affecting the utility difference, it is convenient to represent the model in terms of ΔU_{imt}.

$$\Delta U_{imt} = \alpha_{im} + \gamma_i \ y_{jmt} + \beta_{ii} \ y_{im,t-1} + \beta_{ij} \ y_{jm,t-1} - \varepsilon_{imt} \tag{2}$$

where: $\alpha_{im} \equiv \widetilde{\alpha}_{im}(1) - \widetilde{\alpha}_{im}(0) + \widetilde{\gamma}_i(1,0) - \widetilde{\gamma}_i(0,0) + \widetilde{\beta}_{ii}(1,0) - \widetilde{\beta}_{ii}(0,0) + \widetilde{\beta}_{ij}(1,0) - \widetilde{\beta}_{ij}(0,0)$; $\gamma_i \equiv \widetilde{\gamma}_i(1,1) - \widetilde{\gamma}_i(0,1) - \widetilde{\gamma}_i(1,0) + \widetilde{\gamma}_i(0,0)$; $\beta_{ii} \equiv \widetilde{\beta}_{ii}(1,1) - \widetilde{\beta}_{ii}(0,1) - \widetilde{\beta}_{ii}(1,0) + \widetilde{\beta}_{ii}(0,0)$; $\beta_{ij} \equiv \widetilde{\beta}_{ij}(1,1) - \widetilde{\beta}_{ij}(0,1) - \widetilde{\beta}_{ij}(1,0) + \widetilde{\beta}_{ij}(0,0)$; and $\varepsilon_{imt} \equiv \widetilde{\varepsilon}_{imt}(0) - \widetilde{\varepsilon}_{imt}(1)$. Since $\widetilde{\varepsilon}_{imt}(j)'s$ are i.i.d. extreme value type 1 distributed, we have that ε_{imt} and ε_{jmt} are *i.i.d.* Logistic distributed.

A *Nash equilibrium* in the game of complete information is a solution in (y_1, y_2) to the system of best response conditions:

$$\left\{ \begin{array}{lcl} y_{1mt} & = & 1\{\alpha_{1m} + \gamma_1 \ y_{2mt} + \beta_{11} \ y_{1m,t-1} + \beta_{12} \ y_{2m,t-1} - \varepsilon_{1mt} \geq 0\} \\ \\ y_{2mt} & = & 1\{\alpha_{2m} + \gamma_2 \ y_{1mt} + \beta_{21} \ y_{1m,t-1} + \beta_{22} \ y_{2m,t-1} - \varepsilon_{2mt} \geq 0\} \end{array} \right. \tag{3}$$

where $1\{.\}$ is the indicator function.

A *Bayesian Nash equilibrium* in the incomplete information game is a solution in the choice probabilities (P_{1mt}, P_{2mt}) to the system of equations:

$$\left\{ \begin{array}{lcl} P_{1mt} & = & \Lambda\left(\alpha_{1m} + \gamma_1 \ P_{2mt} + \beta_{11} \ y_{1mt-1} + \beta_{12} \ y_{2mt-1}\right) \\ \\ P_{2mt} & = & \Lambda\left(\alpha_{2m} + \gamma_2 \ P_{1mt} + \beta_{21} \ y_{1mt-1} + \beta_{22} \ y_{2mt-1}\right) \end{array} \right. \tag{4}$$

where, for $i \in \{1, 2\}$, P_{imt} is the choice probability of player i, $\mathbb{P}\left[y_{imt} = 1 \mid \alpha_m, \ y_{1mt-1}, \ y_{2mt-1}\right]$, and $\Lambda(u)$ is the the Logistic function $\exp\{u\}/[1 + \exp\{u\}]$.

The researcher observes $\{y_{1mt}, y_{2mt}\}$ over M markets and T periods, where M is large and T is small. The model is a fixed effects model in the sense that the joint probability distribution of the incidental parameters $(\alpha_{1m}, \alpha_{2m})$ and the initial conditions $(y_{1m0}, , y_{2m0})$ is nonparametrically specified. We are interested in the identification of the vector of structural parameters $\theta = (\gamma_i, \beta_{ij} :$

$i, j \in \{1, 2\})'$. For the rest of this section, we present identification (and under-identification) results for different versions of this model. For notational simplicity, we omit the market subindex m and use α to represent the pair of fixed effects $(\alpha_{1m}, \alpha_{2m})$ for one market.

2.2 Model with no contemporaneous strategic interactions

Consider the model described above under the condition that $\gamma_1 = \gamma_2 = 0$. The best response equations for this model are:[3]

$$\begin{cases} y_{1t} & = & 1\{\alpha_1 + \beta_{11}\ y_{1t-1} + \beta_{12}\ y_{2t-1} - \varepsilon_{1t} \geq 0\} \\ \\ y_{2t} & = & 1\{\alpha_2 + \beta_{21}\ y_{1t-1} + \beta_{22}\ y_{2t-1} - \varepsilon_{2t} \geq 0\} \end{cases} \tag{5}$$

Note that this model can be interpreted either as complete or incomplete information because the value of the opponent's ε does not have any effect on the best response of a player. Though this model is dynamic (i.e., past choices affect current choices) and incorporates dynamic interactions between the agents (i.e., the past choice of player j has an effect on the current choice of player i), it seems that it is not a game because there are not explicit strategic interactions in the best responses of the agents. However, it can interpreted as a game under some conditions about the formation of beliefs.

EXAMPLE 1a (Fictitious play learning). Consider a dynamic binary choice game of incomplete information where the differential utility of a player is $\Delta U_{it} = \alpha_i + \gamma_i\ y_{jt} + \beta_{ii}\ y_{it-1} - \varepsilon_{it}$. The best response of this player is $y_{it} = 1\{\alpha_i + \gamma_i\ B_{it} + \beta_{ii}\ y_{it-1} - \varepsilon_{it} \geq 0\}$, where B_{it} is a probability that represents the belief that player i has about the probability that player j chooses alternative j. Under the assumption of Bayesian Nash Equilibrium, this belief correspond to the actual probability $\mathbb{P}(y_{jt} = 1 \mid \alpha, y_{1t-1}, y_{2t-1})$ in the equilibrium of the game. However, other solution concepts consider alternative assumptions on players' beliefs. In particular, under the assumption of *fictitious play learning* (Brown, 1951), a player's belief B_{it} consists of the frequency of the other player's choices during the last R times that the game has been played: $B_{it} = (1/R) \sum_{r=1}^{R} y_{j,t-r}$. For $R = 1$, the best response functions under fictitious play are the ones in equation (5). Examples of empirical applications include Holt (1999) and Lee and Pakes (2009), among others. ■

The terms α_1 and α_2 are fixed effects. The structural parameters are β_{11}, β_{12}, β_{21}, and β_{22}. The data for one market consist of the history of choices between periods 1 and T and the initial conditions (y_{10}, y_{20}). We represent these data using the vector $\widetilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, .., T)$ and we

[3] This model corresponds to the panel data binary choice VAR model studied by Honoré and Kyriazidou (2017).

refer to this vector as a *market history*. The model implies the following probability:

$$\mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) = \prod_{i=1}^{2}\prod_{t=1}^{T} \frac{\exp\left\{\ y_{it}\left[\alpha_i + \beta_{i1}\ y_{1t-1} + \beta_{i2}\ y_{2t-1}\right]\ \right\}}{1+\exp\left\{\alpha_i + \beta_{i1}\ y_{1t-1} + \beta_{i2}\ y_{2t-1}\right\}}\ p_\alpha\left(y_{10}, y_{20}\right) \tag{6}$$

where $p_\alpha\left(y_{10}, y_{20}\right)$ represents the probability of the initial condition given α.

A key property of the logit model is that it facilitates a representation of the log-likelihood which is additively separable in incidental parameters α and the structural parameters β. We derive now an expression that illustrates this separability and that plays a key role in the identification results.

Define, for $i \in \{1, 2\}$, function $\sigma_{\alpha i}\left(y_1, y_2\right) \equiv -\ln[1+\exp\left\{\alpha_i + \beta_{i1}\ y_1 + \beta_{i2}\ y_2\right\}]$, and let $\sigma_\alpha\left(y_1, y_2\right) \equiv \sigma_{\alpha 1}\left(y_1, y_2\right)+\sigma_{\alpha 2}\left(y_1, y_2\right)$. Given a choice history $\widetilde{\mathbf{y}}$, define the following statistics: $T_i^{(1)}$ is the number of times that player i chooses alternative 1, that is, $\sum_{t=1}^{T} y_{it}$; for $(y_1, y_2) \in \{0,1\}^2$, $T^{(y_1,y_2)}$ is the number of times that the two players choose the pair (y_1, y_2), that is, $\sum_{t=1}^{T} 1\{(y_{1t}, y_{2t}) = (y_1, y_2)\}$; C_{12} is the number of times that player 1 chooses alternative 1 given that player 2 choose alternative 1 at previous period, that is, $\sum_{t=1}^{T} y_{1t}\ y_{2t-1}$; and similarly, we define C_{21} as $\sum_{t=1}^{T} y_{2t} y_{1t-1}$, C_{11} as $\sum_{t=1}^{T} y_{1t} y_{1t-1}$, and C_{22} as $\sum_{t=1}^{T} y_{2t}\ y_{2t-1}$. Then, the logarithm of the probability of a market history can be written as:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) &= \ln p_\alpha\left(y_{10}, y_{20}\right) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} \\
&+ \sum_{y_1=0}^{1}\sum_{y_2=0}^{1} \sigma_\alpha\left(y_1, y_2\right)\ \left[T^{(y_1,y_2)} + 1\{(y_{10}, y_{20}) = (y_1, y_2)\} - 1\{(y_{1T}, y_{2T}) = (y_1, y_2)\}\right. \\
&+ \beta_{11}\ C_{11} + \beta_{12}\ C_{12} + \beta_{21}\ C_{21} + \beta_{22}\ C_{22}
\end{aligned} \tag{7}$$

Or using a more compact representation:

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) = \mathbf{s}\left(\widetilde{\mathbf{y}}\right)'\ \mathbf{g}_\alpha + \mathbf{c}\left(\widetilde{\mathbf{y}}\right)'\ \beta \tag{8}$$

where $\mathbf{s}\left(\widetilde{\mathbf{y}}\right)$ and $\mathbf{c}\left(\widetilde{\mathbf{y}}\right)$ are vectors of statistics, $\mathbf{g}_\alpha$ is a vector of functions of the incidental parameters, and β is the vector of structural parameters$(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22})'$. More specifically,

$$\left\{\begin{array}{rcllll}
\mathbf{s}\left(\widetilde{\mathbf{y}}\right)' &=& [1,\ y_{10},\ y_{20},\ y_{10}y_{20} & ;\ 1, y_{1T}, y_{2T}, y_{1T}y_{2T} & ;\ T,\ T_1^{(1)},\ T_2^{(1)},\ T^{(1,1)}] \\
\mathbf{g}_\alpha' &=& [\ln \mathbf{p}_\alpha^* + \sigma_\alpha^* & ;\ -\sigma_\alpha^* & ;\ \sigma_\alpha^* + (0, \alpha_1, \alpha_2, 0)] \\
\mathbf{c}\left(\widetilde{\mathbf{y}}\right)' &=& [C_{11},\ C_{12},\ C_{21},\ C_{22}] \\
\beta' &=& [\beta_{11},\ \beta_{12},\ \beta_{21},\ \beta_{22}]
\end{array}\right. \tag{9}$$

with $\ln \mathbf{p}_\alpha^* \equiv [\ln p_\alpha(0,0),\ \ln p_\alpha(1,0) - \ln p_\alpha(0,0),\ \ln p_\alpha(0,1) - \ln p_\alpha(0,0),\ \ln p_\alpha(1,1) - \ln p_\alpha(0,1) - \ln p_\alpha(1,0) + \ln p_\alpha(0,0)]$, and $\sigma_\alpha^* \equiv [\sigma_\alpha(0,0),\ \sigma_\alpha(1,0) - \sigma_\alpha(0,0),\ \sigma_\alpha(0,1) - \sigma_\alpha(0,0),\ \sigma_\alpha(1,1) - \sigma_\alpha(0,1) - \sigma_\alpha(1,0) + \sigma_\alpha(0,0)]$.

Given equation (8) we can establish the following identification result.

PROPOSITION 1. For the myopic dynamic game without contemporaneous effects described by equation (5): (A) The vector $\mathbf{s}(\widetilde{\mathbf{y}}) = [1,\ y_{10},\ y_{20},\ y_{10}y_{20},\ y_{1T},\ y_{2T},\ y_{1T}y_{2T},\ T,\ T_1^{(1)},\ T_2^{(1)},\ T^{(1,1)}]$ *is a minimal sufficient statistic for* α *such that* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \alpha, \beta)$ *does not depend on* α. *(B) Since* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \alpha, \beta) = \ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) - \ln \mathbb{P}(\mathbf{s}(\widetilde{\mathbf{y}}) \mid \alpha, \beta)$ *, we have that*

$$\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \beta) \quad = \quad \mathbf{c}(\widetilde{\mathbf{y}})' \beta - \ln \left(\sum_{\widetilde{\mathbf{y}}' : \mathbf{s}(\widetilde{\mathbf{y}}') = \mathbf{s}(\widetilde{\mathbf{y}})} \exp\{\mathbf{c}(\widetilde{\mathbf{y}}')' \beta)\} \right) \tag{10}$$

with $\mathbf{c}(\widetilde{\mathbf{y}}) = [C_{11},\ C_{12},\ C_{21},\ C_{22}]'$ *and* $\beta = [\beta_{11},\ \beta_{12},\ \beta_{21},\ \beta_{22}]'$. *(C) For* $T \geq 3$, *there are pairs of market histories, say* A *and* B, *with* $\mathbf{s}(A) = \mathbf{s}(B)$ *and* $\mathbf{c}(A) - \mathbf{c}(B) = [1,\ 0,\ 0,\ 0]'$ *such that the parameter* β_{11} *is identified as* $\beta_{11} = \ln \mathbb{P}(A) - \ln \mathbb{P}(B)$. *The same result applies to the identification of the the other structural parameters* β_{12}, β_{21}, *and* β_{22}. ■

Table 2 and Example 1b provide specific histories that identify the parameters of interest when $T = 3$.

Table 2
Myopic Game of Complete Information
with no 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_{11}
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_{22}
Case 3:	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	β_{12}
Case 4:	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	β_{21}

EXAMPLE 1b. 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,

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

Using the expressions in (11), we present in Table 2 examples of histories A and B and the corresponding parameter that is identified by $\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$. In cases 1 and 2, we identify the parameter β_{ii} 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 β parameters. In general, with $T = 3$, we have that the vector $(\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3)$ can take $4^4 = 256$ values. Let $\mathbf{C}_A$ and $\mathbf{C}_B$ be the matrices with dimension 256×4 such that each row contains the vector of statistics $\mathbf{c}(A)'$ and $\mathbf{c}(B)'$, respectively, for a particular value of $(\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3)$. Similarly, let $\ln \mathbf{P}_A$ and $\ln \mathbf{P}_B$ be the vectors with dimension 256×1 that contain the log-probabilities $\ln P(A)$ and $\ln P(B)$, respectively, for the different values of $(\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3)$. Equation (9) together with the fact that $\mathbf{s}(A) = \mathbf{s}(B)$ imply that $\ln \mathbf{P}_A - \ln \mathbf{P}_B = (\mathbf{C}_A - \mathbf{C}_B)\ \beta$. Furthermore, matrix $\mathbf{C}_A - \mathbf{C}_B$ is full column rank and we can identify β as:

$$\beta = \left[(\mathbf{C}_A - \mathbf{C}_B)' (\mathbf{C}_A - \mathbf{C}_B) \right]^{-1} \left[(\mathbf{C}_A - \mathbf{C}_B)' (\ln \mathbf{P}_A - \ln \mathbf{P}_B) \right] \tag{12}$$

The sample counterpart of this expression provides a root-M consistent and asymptotically normal estimator of β. However, this estimator is not as efficient as the Conditional Maximum Likelihood estimator that we present in section 4 below. ■

2.3 Complete information and triangular 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 $\beta_{21} = 0$. That is, the model is defined by the following best response functions:

$$\begin{cases} y_{1t} & = & 1\{\alpha_1 + \beta_{11}\ y_{1t-1} + \beta_{12}\ y_{2t-1} - \varepsilon_{1t} \geq 0\} \\ \\ y_{2t} & = & 1\{\alpha_2 + \gamma_2\ y_{1t} + \beta_{22}\ y_{2t-1} - \varepsilon_{2t} \geq 0\} \end{cases} \tag{13}$$

This system can be interpreted as a dynamic Stackelberg game where, at every period t, player 1 is the leader and decides first, and player 2 makes his choice after knowing the current choice of player 1.

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

$$\begin{array}{rcl} \ln \mathbb{P}(\widetilde{\mathbf{y}}|\alpha,\beta) & = & \ln p_\alpha(y_{10}, y_{20}) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} + \sum_{t=1}^{T} \sigma_{\alpha 1}(y_{1t-1}, y_{2t-1}) + \sigma_{\alpha 2}(y_{1t}, y_{2t-1}) \\ \\ & + & \beta_{11}\ C_{11} + \beta_{12}\ C_{12} + \beta_{22}\ C_{22} + \gamma_2\ T^{(1,1)} \end{array} \tag{14}$$

with $\sigma_{\alpha 1}(y_1, y_2) \equiv -\ln[1+\exp\{\alpha_1+\beta_{11}y_1+\beta_{12}y_2\}]$ and $\sigma_{\alpha 2}(y_1, y_2) \equiv -\ln[1+\exp\{\alpha_2+\gamma_2 y_1+\beta_{22}y_2\}]$. Comparing equations (14) and (7), we can see two important differences: the term $\gamma_2\ T^{(1,1)}$; and the term $\sum_{t=1}^{T}\sigma_{\alpha 2}(y_{1t}, y_{2t-1})$ instead of $\sum_{t=1}^{T}\sigma_{\alpha 2}(y_{1t-1}, y_{2t-1})$. These differences have implications on the identification of the parameters. The log-probability depends on γ_2 *only* through the statistic $T^{(1,1)}$, and this statistic is also associated with the incidental parameters – with $\sigma_{\alpha 1}$. Similarly, the log-probability depends on β_{12} *only* through the statistic C_{12}, but this statistic is also associated with the incidental parameters through the term $C_{12}\ [\sigma_{\alpha 2}(1,1) - \sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,1) + \sigma_{\alpha 2}(0,0)]$.[4] This implies that the parameters γ_2 and β_{12} cannot be identified.

Similarly as for the previous model, we can rewrite the right hand side of equation (14) as $\mathbf{s}(\widetilde{\mathbf{y}})'\ \mathbf{g}_\alpha + \mathbf{c}(\widetilde{\mathbf{y}})'\ \beta^*$, but now the vectors of statistics $\mathbf{s}(\widetilde{\mathbf{y}})$ and $\mathbf{c}(\widetilde{\mathbf{y}})$, and the vector of identified parameters β^* are different. More specifically,

$$\begin{cases} \mathbf{s}(\widetilde{\mathbf{y}})' & = & [1, y_{10}, y_{20}, y_{10}y_{20} & ; & 1, y_{1T}, y_{2T}, y_{1T}y_{2T} & ; & T, T_1^{(1)}, T_2^{(1)}, T^{(1,1)} & ; & C_{12}] \\ \mathbf{g}'_\alpha & = & [\ln \mathbf{p}^*_\alpha + \sigma^*_\alpha & ; & -\sigma^*_\alpha\ ; & ; & \begin{array}{c} \sigma^*_\alpha + \\ (0, \alpha_1, \alpha_2, -\Delta\sigma_{\alpha 2} + \gamma_2) \end{array} & ; & \begin{array}{c} \Delta\sigma_{\alpha 2} \\ +\beta_{12} \end{array}] \\ \mathbf{c}(\widetilde{\mathbf{y}})' & = & [C_{11},\ C_{22}] \\ \beta^{*\prime} & = & [\beta_{11},\ \beta_{22}] \end{cases} \tag{15}$$

where $\ln \mathbf{p}^*_\alpha$ and σ^*_α have the same definition as in section 2.2 above. For $i \in \{1, 2\}$, $\Delta\sigma_{\alpha i}$ is the incidental parameter $\sigma_{\alpha i}(1,1) - \sigma_{\alpha i}(0,1) - \sigma_{\alpha i}(1,0) + \sigma_{\alpha i}(0,0)$.

[4] Note that $\sum_{t=1}^{T}\sigma_{\alpha 2}(y_{1t}, y_{2t-1})$ can be written as $\left[\sum_{t=1}^{T}(1-y_{1t})(1-y_{2t-1})\right]\ \sigma_{\alpha 2}(0,0) + \left[\sum_{t=1}^{T}y_{1t}(1-y_{2t-1})\right]$ $\sigma_{\alpha 2}(1,0) + \left[\sum_{t=1}^{T}(1-y_{1t})y_{2t-1}\right]\ \sigma_{\alpha 2}(0,1) + \left[\sum_{t=1}^{T}y_{1t}y_{2t-1}\right]\ \sigma_{\alpha 2}(1,1)$, and this expression is equal to $T\ \sigma_{\alpha 2}(0,0) + T_1^{(1)}$ $[\sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,0)] + \left[T_2^{(1)} + y_{20} - y_{2T}\right]\ [\sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0)] + C_{12}\ [\sigma_{\alpha 2}(1,1) - \sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,1) + \sigma_{\alpha 2}(0,0)]$.

There are some fundamental differences with respect to the model without contemporaneous strategic interactions. First, the statistic C_{12} and the structural parameter β_{12} appear in the log-probability of a choice history through the term C_{12} $(\Delta\sigma_{\alpha 2} + \beta_{12})$. Without further restrictions we have that the incidental parameter $\Delta\sigma_{\alpha 2}$ is not zero. This implies that the parameter β_{12} is not identified. Second, the statistic $T^{(1,1)}$ and the structural parameter γ_2 appear in $\ln \mathbb{P}(\widetilde{\mathbf{y}}|\alpha, \beta)$ through the term $T^{(1,1)}$ $(\Delta\sigma_{\alpha 1} + \gamma_2)$. Therefore, without further restrictions, the parameter γ_2 is not identified.

PROPOSITION 2. For the myopic, complete information, triangular ('Stackelberg') dynamic game described by equation (13): (A) The vector $\mathbf{s}(\widetilde{\mathbf{y}}) = [1, y_{10}, y_{20}, y_{10}y_{20}, y_{1T}, y_{2T}, y_{1T}y_{2T}, T, T_1^{(1)}, T_2^{(1)}, T^{(1,1)}, S_{12}]'$ *is a minimal sufficient statistic for* α *such that* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \alpha, \beta)$ *does not depend on* α. *(B)* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \beta) = \mathbf{c}(\widetilde{\mathbf{y}})' \beta^* - \ln(\sum_{\widetilde{\mathbf{y}}':\mathbf{s}(\widetilde{\mathbf{y}}')=\mathbf{s}(\widetilde{\mathbf{y}})} \exp\{\mathbf{c}(\widetilde{\mathbf{y}}')'\beta^*\})$ *with* $\mathbf{c}(\widetilde{\mathbf{y}}) = [C_{11},\ C_{22}]'$ *and* $\beta^* = [\beta_{11},\ \beta_{22}]'$. *(C) For* $T \geq 3$, *there are histories* $\widetilde{\mathbf{y}}$ *such that* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \beta)$ *identifies the vector of parameters of interest* β^*. ■

Example 2 presents pairs of histories that identify parameters β_{11} and β_{22} in this model.

EXAMPLE 2. Consider the same framework as in Example 1b: $T = 3$ 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 Example 1b, 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 example 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_{11} and β_{22}, respectively, in this 'Stackelberg' dynamic game. We present this result in table 3. In contrast, for case 3 we have that $C_{12}(A) = 1$ and $C_{12}(B) = 0$, such that this history does not identify β_{12}. For case 4, we have that $C_{12}(A) = C_{12}(B) = 0$, but for this case we also have that $\mathbf{c}(A) = \mathbf{c}(B)$ such that $\ln \mathbb{P}(A \mid \mathbf{s}, \beta) = \ln \mathbb{P}(B \mid \mathbf{s}, \beta)$ such that this pair of histories does not identify any parameter. ■

Table 3
Myopic Game of Complete Information
Triangular with contemporaneous effect $y_1 \to y_2$
Examples of histories and identified parameters with T=3
$A = \{\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\};\quad B = \{\boldsymbol{y}_0,\ \boldsymbol{b},\ \boldsymbol{a},\ \boldsymbol{y}_3\}$ 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_{11}
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_{22}

As explained above, the no identification of the parameters γ_2 and β_{12} 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} $(\Delta\sigma_{\alpha 2} + \beta_{12})$, 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 $-\ln[1 + \exp\{\alpha_1 + \beta_{11}\ y_1 + \beta_{12}\ y_2\}]$. Taking this into account we have that $\Delta\sigma_{\alpha 1} = 0$ if and only if $\beta_{11} = 0$ or/and $\beta_{12} = 0$. Following a similar argument, we have the parameter β_{12} 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 $\beta_{22} = 0$. We summarize these identification results in the following Proposition 3.

PROPOSITION 3. For the myopic, complete information, triangular ('Stackelberg') dynamic game described by equation (13): (A) a necessary and sufficient condition for the identification of parameter γ_2 is that $\beta_{11} = 0$ or/and $\beta_{12} = 0$; (B) a necessary and sufficient condition for the identification of parameter β_{12} is that $\gamma_2 = 0$ or/and $\beta_{22} = 0$. ■

Based on Proposition 3, there is a myopic dynamic game where all the structural parameters are identified. This is the model with $\beta_{12} = 0$. That is,

$$\begin{cases} y_{1t} = 1\{\alpha_1 + \beta_{11}\ y_{1t-1} - \varepsilon_{1t} \geq 0\} \\ \\ y_{2t} = 1\{\alpha_2 + \gamma_2\ y_{1t} + \beta_{22}\ y_{2t-1} - \varepsilon_{2t} \geq 0\} \end{cases} \tag{16}$$

In this dynamic game both players have switching costs as represented by the parameters β_{11} and β_{22}. Player 1 is a Stackelberg leader such that player 2's past and current decisions do not have any effect on the decision of player 1. The best response of player 2 is affected by the contemporaneous choice of player 1.

EXAMPLE 3. Suppose that $T = 3$ and consider the pair of histories A and B with:

$$\begin{array}{rcl} A & = & \left[\begin{pmatrix} 0 \\ y_{20} \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ y_{23} \end{pmatrix} \right] \\ B & = & \left[\begin{pmatrix} 0 \\ y_{20} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ y_{23} \end{pmatrix} \right] \end{array} \tag{17}$$

We first verify that the values taken by the sufficient statistics $\boldsymbol{y}_0$, $\boldsymbol{y}_3$, $T_1^{(1)}$, $T_2^{(1)}$, and C_{12} are the same for the two histories. It is clear the initial and the final choices at $t = 0$ and $t = 3$ are the same in the two histories. We also have that: $T_1^{(1)}(A) = 2 = T_1^{(1)}(B)$; $T_2^{(1)}(A) = 1 + y_{23} = T_2^{(1)}(B)$; and $C_{12}(A) = 1 = C_{12}(B)$. Now, for the identifying statistics C_{11}, C_{22}, and $T^{(1,1)}$ we have that: $C_{11}(A) = 1 = C_{11}(B)$; $C_{22}(A) = y_{23}$ and $C_{22}(B) = y_{20}$; and $T^{(1,1)}(A) = 1+y_{23}$ and $T^{(1,1)}(B) = y_{23}$. Taking into account these results, we have that:

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) = \gamma_2 + (y_{23} - y_{20})\, \beta_{22} \tag{18}$$

Therefore, when $y_{23} - y_{20} = 0$, the frequencies of this pair of histories identify γ_2. ■

Functional differencing approach. So far, we have considered identification results using a conditional likelihood approach. Recently, Honoré and Weidner (2020) and Dobroyni, Gu and Kim (2021) have used a functional differencing approach in the spirit of Bonhomme (2012) to prove the identification of parameters in dynamic logit models that are not identify using a conditional likelihood method. We follow the same approach here, and more specifically the method in Dobroyni, Gu and Kim (2021). We are particularly interested in the identification of parameters γ_2 and β_{12} without the restrictions in Proposition 3. Without further restrictions, the functional differenting approach does not provide (point) identification of the parameters γ_2 and β_{12}. However, it does provide identification of these parameters when we restrict the fixed effects α_{1m} and α_{2m} to be the same for the two players. Importantly, the conditional likelihood approach does not identify γ_2 and β_{12} under this additional restriction. Proposition 4 presents this result. The proof, that includes a description of Dobroyni, Gu and Kim (2021) functional differenting approach for our model, is in the Appendix.

PROPOSITION 4. Consider the myopic, complete information, triangular ('Stackelberg') dynamic game described by equation (13) 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_{11}, β_{22}, β_{12},and γ_2. ■

2.4 Incomplete information and triangular strategic interaction

Now, consider the incomplete information version of the Stackelberg dynamic game in equation (16). That is,

$$\left\{ \begin{array}{lll} y_{1t} & = & 1\left\{\alpha_1 + \beta_{11}\ y_{1t-1} - \varepsilon_{1t} \geq 0\right\} \\ \\ y_{2t} & = & 1\left\{\alpha_2 + \gamma_2\ \Lambda_{\alpha_1}(y_{1t-1}) + \beta_{22}\ y_{2t-1} - \varepsilon_{2t} \geq 0\right\} \end{array} \right. \tag{19}$$

where $\Lambda_{\alpha_1}(y_{1t-1}) \equiv \Lambda\left(\alpha_1 + \beta_{11}\ y_{1t-1}\right)$ and $\Lambda(.)$ is the Logistic function. The log-probability of a choice history is:

$$\begin{array}{rcl} \ln \mathbb{P}\left(\widetilde{\mathbf{y}}|\alpha,\beta\right) & = & \ln p_{\alpha}\left(y_{10}, y_{20}\right) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} \\ \\ & + & \sum\limits_{t=1}^{T} \left[\sigma_{\alpha 1}\left(y_{1t-1}\right) + \gamma_2\ y_{2t}\ \Lambda_{\alpha_1}(y_{1t-1}) + \sigma_{\alpha 2}\left(y_{1t-1}, y_{2t-1}\right)\right] \\ \\ & + & \beta_{11}\ C_{11} + \beta_{22}\ C_{22} \end{array} \tag{20}$$

with $\sigma_{\alpha 1}\left(y_1\right) \equiv -\ln[1 + \exp\{\alpha_1 + \beta_{11}\ y_1\}]$, and $\sigma_{\alpha 2}\left(y_1, y_2\right) \equiv -\ln[1 + \exp\{\alpha_2 + \gamma_2\ \Lambda_{\alpha_1}(y_1) + \beta_{22}\ y_2\}]$.

In terms of incidental parameters and sufficient statistics, this model has two main differences with its complete information counterpart. First, the term $\gamma_2\ T^{(1,1)}$ in the model with complete information is now replaced by $\gamma_2\ [\Lambda_{\alpha_1}(1) - \Lambda_{\alpha_1}(0)]\ C_{21}$. Since γ_2 appears multiplying $\Lambda_{\alpha_1}(1) - \Lambda_{\alpha_1}(0)$, this structural parameter is not additively separable from the incidental parameters and it cannot be identified. This difference with respect to the complete information model also implies that now C_{21} is part of the vector of sufficient statistics. Second, the term $\sum_{t=1}^{T}\sigma_{\alpha 2}(y_{1t}, y_{2t-1})$, that appears in the model of complete information, is now replaced by $\sum_{t=1}^{T}\sigma_{\alpha 2}(y_{1t-1}, y_{2t-1})$. This implies that, in contrast to the complete information model, now $T^{(1,1)}$ is part of the vector of sufficient statistics but C_{12} is not.

PROPOSITION 5. For the myopic, incomplete information, triangular ('Stackelberg') dynamic game described by equation (19): (A) The vector $\mathbf{s}\left(\widetilde{\mathbf{y}}\right) = [1,\ y_{10},\ y_{20},\ y_{10}y_{20},\ y_{1T},\ y_{2T},\ y_{1T}y_{2T},\ T,\ T_1^{(1)},\ T_2^{(1)},\ T^{(1,1)},\ C_{21}]'$ *is a minimal sufficient statistic for* α *such that* $\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{s}\left(\widetilde{\mathbf{y}}\right), \alpha, \beta\right)$ *does not depend on* α. *(B)* $\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{s}\left(\widetilde{\mathbf{y}}\right), \beta\right) = \mathbf{c}\left(\widetilde{\mathbf{y}}\right)' \beta^{*} - \ln(\sum_{\widetilde{\mathbf{y}}':\mathbf{s}(\widetilde{\mathbf{y}}')=\mathbf{s}(\widetilde{\mathbf{y}})} \exp\left\{\mathbf{c}\left(\widetilde{\mathbf{y}}'\right)'\beta^{*}\right\})$ *with* $\mathbf{c}\left(\widetilde{\mathbf{y}}\right) = [C_{11},\ C_{22}]'$ *and* $\beta^{*} = [\beta_{11},\ \beta_{22}]'$. *(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* β_{11} *and* β_{22}. ■

In Example 5, we present pairs of histories that identify parameters β_{11} and β_{22}.

EXAMPLE 5. Consider the histories in Table 3. It is straightforward to verify that these two pairs of histories still identify the parameters β_{11} and β_{22}, respectively, in the model with incomplete information. Note that for these histories we also have that $C_{21}(A) = C_{21}(B)$ such that they also have the same value for the sufficient statistic $\mathbf{s}$ in Proposition 5. Therefore, for these particular histories, the identification of the parameters β_{11} and β_{22} is robust to the assumption of either private or common knowledge of the ε variables. ■

Partial identification of γ_2. Since γ_2 is not point identified, we explore here the identification of bounds for this parameter, i.e., partial identification. Let $P_1(y_{1,t-1}, \alpha_1)$ and $P_2(y_{t-1}, \alpha)$ represent the probabilities $\mathbb{P}(y_{1t} = 1|y_{1,t-1}, \alpha_1)$ and $\mathbb{P}(y_{2t} = 1|y_{t-1}, \alpha_1, \alpha_2)$, respectively. Since the first player is the leader and its behavior does not depend on the second player, we can treat it as a single agent dynamic logit model. For that model we know that, as soon as $T \geq 3$, the average effect $\Delta^{(1)} = \int [P_1(1, \alpha_1) - P_1(0, \alpha_1)] dF(\alpha_1)$ is point identified, as shown in Aguirregabiria and Carro (2020). For $(y_1, y_2), \in \{0, 1\}^2$, define the following average effect for player 2: $\Delta^{(2)}_{y_1,y_2} = \int [P_2(y_1, y_2, \alpha) - P_2(0, 0, \alpha)] dF(\alpha)$. Under the conditions $\beta_{11} \geq 0$, $\beta_{22} \geq 0$, and $\gamma_2 \leq 0$, it is simple to show that: $\Delta^{(2)}_{0,1} \geq 0$, $-\Delta^{(2)}_{1,0} \geq 0$, $\Delta^{(2)}_{1,1} - \Delta^{(2)}_{1,0} \geq 0$, and $\Delta^{(2)}_{0,1} - \Delta^{(2)}_{1,1} \geq 0$. Now making use of the fact that $\Lambda(x) - \Lambda(y) \leq \frac{1}{4}(x - y)$ if $x > y$, we get the following inequalities:

$$
\begin{aligned}
&0 \leq \Delta^{(2)}_{0,1} \leq \frac{1}{4}\beta_{22} \\
&0 \leq -\Delta^{(2)}_{1,0} \leq -\frac{\gamma_2}{4}\Delta^{(1)} \\
&0 \leq \Delta^{(2)}_{1,1} - \Delta^{(2)}_{1,0} \leq \frac{1}{4}\beta_{22} \\
&0 \leq \Delta^{(2)}_{0,1} - \Delta^{(2)}_{1,1} \leq -\frac{\gamma_2}{4}\Delta^{(1)}
\end{aligned}
$$

Now consider choice histories $C = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ and $D = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. We have that:

$$
\begin{aligned}
&P(D) - P(C) \\
&= \int p^*_\alpha(0,0)(1 - P_1(0,0,\alpha))(1 - P_1(1,0,\alpha))P_1(0,0,\alpha)(1 - P_2(0,0,\alpha))(P_2(1,0,\alpha) - P_2(0,0,\alpha))dF(\alpha) \\
&\geq \frac{1}{4}\Delta^{(2)}_{1,0} \geq \frac{\gamma_2\Delta^{(1)}}{16}
\end{aligned}
$$

Since we observe the left hand size and $\Delta^{(1)}$ is is point identified, the last inequality provides informative bounds on γ_2. There are other histories that provide informative bounds on this parameter.

PROPOSITION 6. In the myopic, incomplete information, triangular ('Stackelberg') dynamic game described by equation (19), under conditions $\beta_{11} \geq 0$, $\beta_{22} \geq 0$, and $\gamma_2 \leq 0$, three are choice histories that provide informative bounds for the parameter γ_2 such that this parameters is partially identified.. ■

2.5 Complete information and full contemporaneous strategic interactions

Consider the game with contemporaneous effects of y_2 on y_1 (i.e., $\gamma_1 \neq 0$) and of y_1 on y_2 (i.e., $\gamma_2 \neq 0$). As shown above, the identification results for the triangular model using the conditional likelihood approach establish that β_{12} and β_{21} are not point identified (Proposition 2), and the identification of γ_2 requires either $\beta_{12} = 0$ or $\beta_{11} = 0$ (Proposition 3). Now, in this model where γ_1 and γ_2 are unrestricted, we eliminate the lagged strategic interactions between players such that $\beta_{12} = \beta_{21} = 0$.

$$\begin{cases} y_{1t} & = & 1\{\alpha_1 + \gamma_1\ y_{2t} + \beta_{11}\ y_{1t-1} - \varepsilon_{1t} \geq 0\} \\ \\ y_{2t} & = & 1\{\alpha_2 + \gamma_2\ y_{1t} + \beta_{22}\ y_{2t-1} - \varepsilon_{2t} \geq 0\} \end{cases} \tag{21}$$

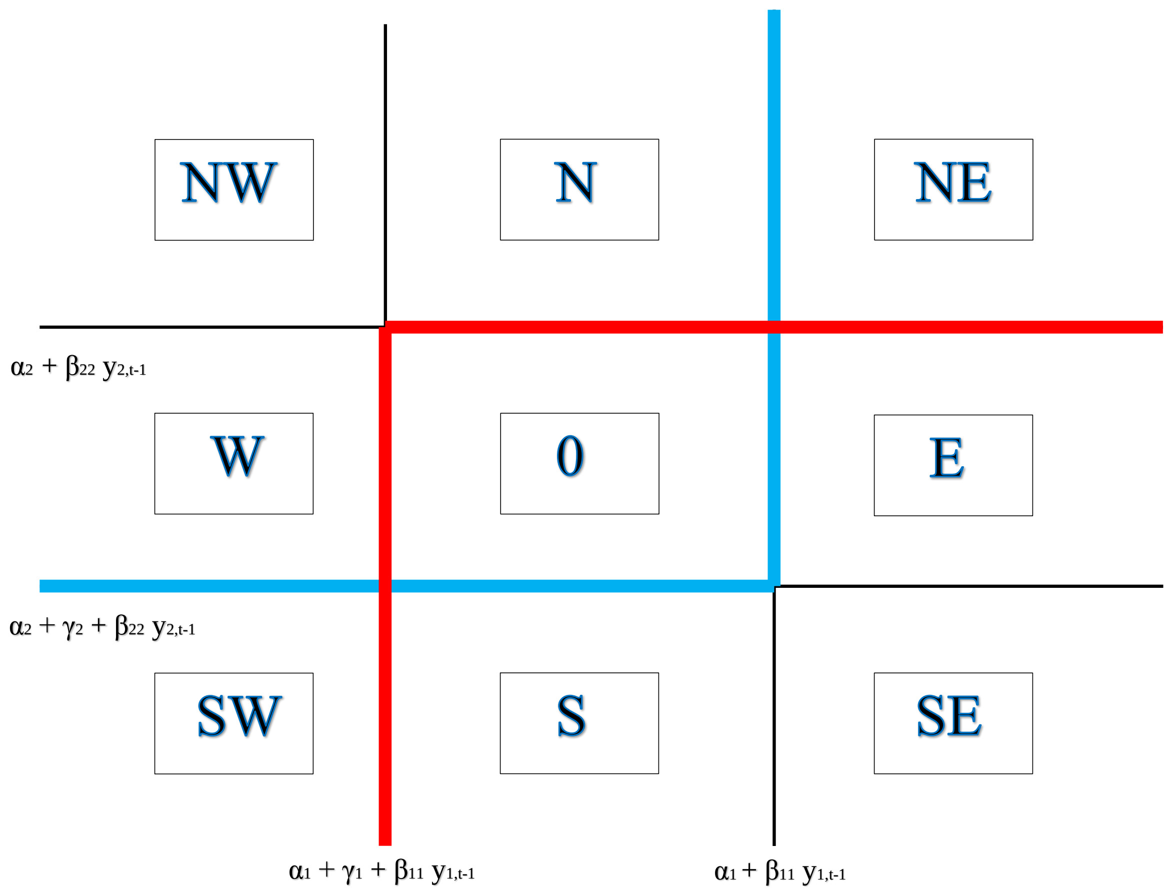
Here we concentrate on the (point) identification of the switching cost parameters – β_{11} and β_{22} – and on the partial identification of all the parameters.

This model is a dynamic panel data version of the two-player binary choice games in Bresnahan and Reiss (1991) and Tamer (2003), among others. Relative to the models considered in previous sections, this model has the additional complication of having multiple equilibria.

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}), and also to describe lower and upper bounds for the probabilities of these outcomes.

There are two vertical lines associated with the values $\alpha_1 + \gamma_1 + \beta_{11}\ y_{1t-1}$ and $\alpha_1 + \beta_{11}\ y_{1t-1}$, respectively, for ε_{1t}. Similarly, there are two horizontal lines associated with the values $\alpha_2 + \gamma_2 + \beta_{22}\ y_{2t-1}$ and $\alpha_2 + \beta_{22}\ y_{2t-1}$, respectively, for ε_{2t}. 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.

Figure 1: Regions for the Complete Information Game with Full Contemporaneous Strategic Interactions



The outcome of the game is $(y_{1t}, y_{2t}) = (1,1)$ if and only if $\varepsilon_{1t} \leq \alpha_1 + \gamma_1 + \beta_{11}\ y_{1t-1}$ and $\varepsilon_{2t} \leq \alpha_2 + \gamma_2 + \beta_{22}\ y_{2t-1}$ which corresponds to the Southwest (SW) quadrangle. Similarly, the outcome of the game is $(y_{1t}, y_{2t}) = (0,0)$ if and only if $\varepsilon_{1t} > \alpha_1 + \beta_{11}\ y_{1t-1}$ and $\varepsilon_{2t} > \alpha_2 + \beta_{22}$ y_{2t-1} which corresponds to the Northeast (NE) quadrangle. Therefore, the model provides unique predictions for the probabilities $\mathbb{P}((y_{1t}, y_{2t}) = (1,1) \mid \mathbf{y}_{t-1}; \alpha)$ and $\mathbb{P}((y_{1t}, y_{2t}) = (0,0) \mid \mathbf{y}_{t-1}; \alpha)$. That is,

$$\begin{cases} P(0,0 \mid \mathbf{y}_{t-1}; \alpha) & = \dfrac{1}{1+\exp\{\alpha_1 + \beta_{11}\ y_{1t-1}\}} \dfrac{1}{1+\exp\{\alpha_2 + \beta_{22}\ y_{2t-1}\}} \\ P(1,1 \mid \mathbf{y}_{t-1}; \alpha) & = \dfrac{\exp\{\alpha_1 + \beta_{11}\ y_{1t-1} + \gamma_1\}}{1+\exp\{\alpha_1 + \beta_{11}\ y_{1t-1} + \gamma_1\}} \dfrac{\exp\{\alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2\}}{1+\exp\{\alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2\}} \end{cases} \tag{22}$$

The quadrangle in the center of figure 1 – labeled as 0 – 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 $P(0,1 \mid \mathbf{y}_{t-1};\alpha)$ and $P(1,0 \mid \mathbf{y}_{t-1};\alpha)$. However, the model establishes bounds on the values of these probabilities.

The upper bound to the probability of outcome $(1,0)$ is given by 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 to 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:

$$\begin{cases} U(0,1|\mathbf{y}_{t-1};\alpha) & \equiv & \dfrac{1}{1+\exp\{\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\}}\ \dfrac{\exp\{\alpha_2+\beta_{22}\ y_{2t-1}\}}{1+\exp\{\alpha_2+\beta_{22}\ y_{2t-1}\}} \\ \\ U(1,0|\mathbf{y}_{t-1};\alpha) & \equiv & \dfrac{\exp\{\alpha_1+\beta_{11}\ y_{1t-1}\}}{1+\exp\{\alpha_1+\beta_{11}\ y_{1t-1}\}}\ \dfrac{1}{1+\exp\{\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\}} \end{cases} \tag{23}$$

This property plays an important role in our identification results.

The lower bounds to the probabilities of outcome $(0,1)$ and $(1,0)$ are equal to the upper bounds minus the probability associated to quadrangle 0 where multiple equilibria exists. Unfortunately, these (sharp) lower bounds do not have a logit structure. Without the logit structure it is not possible to derive sufficient statistics for α (Chamberlain, 2010). For this reason, we use non-sharp lower bounds which have a logit structure. We consider two different lower bounds. For outcome $(0,1)$, the non-sharp lower bounds are defined by the regions in the quadrangles $\{E, SE\}$ and $\{S, SE\}$. These regions imply the following (logit) lower bounds for the probability of outcome $(0,1)$:

$$\begin{cases} L^{\{E,SE\}}(0,1|\mathbf{y}_{t-1};\alpha) & \equiv & \dfrac{1}{1+\exp\{\alpha_1+\beta_{11}\ y_{1t-1}\}}\ \dfrac{\exp\{\alpha_2+\beta_{22}\ y_{2t-1}\}}{1+\exp\{\alpha_2+\beta_{22}\ y_{2t-1}\}} \\ \\ L^{\{S,SE\}}(0,1|\mathbf{y}_{t-1};\alpha) & \equiv & \dfrac{1}{1+\exp\{\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\}}\ \dfrac{\exp\{\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\}}{1+\exp\{\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\}} \end{cases} \tag{24}$$

For outcome $(1,0)$, the non-sharp lower bounds are defined by the regions in quadrangles $\{W, NW\}$ and $\{N, NW\}$, respectively. They imply the following (logit) lower bounds for the probability of outcome $(1,0)$:

$$\begin{cases} L^{\{W,NW\}}(1,0|\mathbf{y}_{t-1};\alpha) & \equiv & \dfrac{\exp\{\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\}}{1+\exp\{\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\}}\ \dfrac{1}{1+\exp\{\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\}} \\ \\ L^{\{N,NW\}}(1,0|\mathbf{y}_{t-1};\alpha) & \equiv & \dfrac{\exp\{\alpha_1+\beta_{11}\ y_{1t-1}\}}{1+\exp\{\alpha_1+\beta_{11}\ y_{1t-1}\}}\ \dfrac{1}{1+\exp\{\alpha_2+\beta_{22}\ y_{2t-1}\}} \end{cases} \tag{25}$$

2.5.1 Approach using only point predictions

First – before we present our bounds approach to derive identification results – we consider here a simpler but more restrictive approach that tries to avoid the problem of multiple equilibria. We use only market histories with $y_{1t} = y_{2t}$ at every period t. For these histories, the model provides point predictions for the probability of a choice history. Unfortunately, we show here that this approach cannot provide point identification of the structural parameters.

Consider a market history $\widetilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, .., T)$ and suppose – for the moment – that the model provides unique predictions about the possible outcomes such that: $\mathbb{P}(y_{1t}, y_{2t} \mid \mathbf{y}_{t-1}; \alpha) = \Lambda([2y_{1t} - 1]\ [\alpha_1 + \gamma_1\ y_{2t} + \beta_{11}\ y_{1t-1}])\ \Lambda([2y_{2t} - 1]\ [\alpha_2 + \gamma_2\ y_{1t} + \beta_{22}\ y_{2t-1}])$. Under this condition, the model implies the following log-probability:

$$
\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) \ &= \ \ln p_{\alpha}\left(y_{10}, y_{20}\right) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} + \textstyle\sum_{t=1}^{T} \sigma_{\alpha 1}\left(y_{1t-1}, y_{2t}\right) + \sigma_{\alpha 2}\left(y_{1t}, y_{2t-1}\right) \\
&+ \ \beta_{11}\ C_{11} + \beta_{22}\ C_{22} + (\gamma_1 + \gamma_2)\ T^{(1,1)}
\end{aligned} \tag{26}
$$

with $\sigma_{\alpha 1}\left(y_{1t-1}, y_{2t}\right) \equiv -\ln[1+\exp\{\alpha_1+\beta_{11}y_{1t-1}+\gamma_1 y_{2t}\}]$ and $\sigma_{\alpha 2}\left(y_{1t}, y_{2t-1}\right) \equiv -\ln[1+\exp\{\alpha_2+\gamma_2\ y_{1t} + \beta_{22}\ y_{2t-1}\}]$. Representing $\sum_{t=1}^{T} \sigma_{\alpha 1}\left(y_{1t-1}, y_{2t}\right) + \sigma_{\alpha 2}\left(y_{1t}, y_{2t-1}\right)$ in terms of the statistics T and C, we can write this log-probability as:

$$
\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) \ &= \ \ln p_{\alpha}\left(y_{10}, y_{20}\right) + T\left[\sigma_{\alpha 1}(0,0) + \sigma_{\alpha 2}(0,0)\right] \\
&+ \ T_1^{(1)}\ \left[\alpha_1 + \Delta\sigma_{\alpha 1}(1,0) + \Delta\sigma_{\alpha 2}(1,0)\right] + T_2^{(1)}\ \left[\alpha_2 + \Delta\sigma_{\alpha 1}(0,1) + \Delta\sigma_{\alpha 2}(0,1)\right] \\
&+ \ \left(y_{10} - y_{1T}\right)\ \Delta\sigma_{\alpha 1}(1,0) + \left(y_{20} - y_{2T}\right)\ \Delta\sigma_{\alpha 2}(0,1) + C_{21}\ \Delta^2\sigma_{\alpha 1} + C_{12}\ \Delta^2\sigma_{\alpha 2} \\
&+ \ \beta_{11}\ C_{11} + \beta_{22}\ C_{22} + (\gamma_1 + \gamma_2)\ T^{(1,1)}
\end{aligned} \tag{27}
$$

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

Now, we take into account that the model provides unique predictions only for market histories with $y_{1t} = y_{2t} = y_t$ for every period $t \geq 1$. This implies the following restrictions on the statistics: (a) $T_1^{(1)} = T_2^{(1)} = T^{(1,1)}$; (b) $C_{11} = C_{21} = \sum_{t=2}^{T} y_t y_{t-1} + y_1\ y_{10}$; and (c) $C_{22} = C_{12} = \sum_{t=2}^{T} y_t y_{t-1} + y_1\ y_{20}$. Plugging these restrictions into the equation for the log-probability, we obtain this expression:

$$
\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) \ &= \ \ln p_{\alpha}\left(y_{10}, y_{20}\right) + T\left[\sigma_{\alpha 1}(0,0) + \sigma_{\alpha 2}(0,0)\right] \\
&+ \ T^{(1,1)}\ \left[\alpha_1 + \alpha_2 + \Delta\sigma_{\alpha 1}(1,0) + \Delta\sigma_{\alpha 2}(1,0) + \Delta\sigma_{\alpha 1}(0,1) + \Delta\sigma_{\alpha 2}(0,1) + \gamma_1 + \gamma_2\right] \\
&+ \ \left(y_{10} - y_{1T}\right)\ \Delta\sigma_{\alpha 1}(1,0) + \left(y_{20} - y_{2T}\right)\ \Delta\sigma_{\alpha 2}(0,1) \\
&+ \ C_{11}\ \left[\beta_{11} + \Delta^2\sigma_{\alpha 1}\right] + C_{22}\ \left[\beta_{22} + \Delta^2\sigma_{\alpha 2}\right]
\end{aligned} \tag{28}
$$

Equation (28) shows that, without further restrictions, neither the parameters γ_1 and γ_2 nor the switching costs β_{11} and β_{22} can be identified. The parameters γ_1 and γ_2 are only related to the statistic $T^{(1,1)}$, and this statistic is also associated with the incidental parameters. Similarly, the parameters β_{11} and β_{22} are only related to the statistics C_{11} and C_{22}, respectively, but these statistics are also associated with the incidental parameters.

Equation (28) shows also that a necessary and sufficient condition for the identification of the switching cost β_{ii}: $\Delta^2\sigma_{\alpha i}$ should be equal to zero for any possible value of the incidental parameter α_i. Using its definition, we have that $\Delta^2\sigma_{\alpha i}$ is equal to:

$$-\ln[1+\exp\{\alpha_i+\beta_{ii}+\gamma_i\}]+\ln[1+\exp\{\alpha_i+\beta_{ii}\}]+\ln[1+\exp\{\alpha_i+\gamma_i\}]-\ln[1+\exp\{\alpha_i\}] \tag{29}$$

It is clear that $\Delta^2\sigma_{\alpha i}=0$ for every value of α_i if and only if $\gamma_i=0$ or $\beta_{ii}=0$. In this section we consider a model without restrictions on γ_1 and γ_2. Therefore, we are ruling out the restriction $\gamma_i=0$. Of course, the restriction $\beta_{ii}=0$ trivially identifies β_{ii}.

PROPOSITION 7. Consider the myopic dynamic game with contemporaneous effects – with $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$ – described by equation (21). If we use data only from market histories with $y_{1t}=y_{2t}$ for every $t \geq 1$ (such that the model has unique predictions), the structural parameters of the model cannot be point identified using a conditional likelihood approach. ■

2.5.2 Conditional likelihood - Bounds approach

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 $\ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \alpha,\beta)=\mathbf{s}_L(\widetilde{\mathbf{y}})'\,\mathbf{g}_\alpha+\mathbf{c}_L(\widetilde{\mathbf{y}})'\,\beta$ and the upper bound is $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \alpha,\beta)=\mathbf{s}_U(\widetilde{\mathbf{y}})'\,\mathbf{g}_\alpha+\mathbf{c}_U(\widetilde{\mathbf{y}})'\,\beta$, 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 $\mathbf{g}_\alpha$ is a vector of incidental parameters. Given this structure, the logarithm of the probability of a market history $\widetilde{\mathbf{y}}$ (unconditional on α) has the following bounds:

$$h(\mathbf{s}_L(\widetilde{\mathbf{y}}))+\mathbf{c}_L(\widetilde{\mathbf{y}})'\,\beta \;\leq\; \ln \mathbb{P}(\widetilde{\mathbf{y}}) \;\leq\; h(\mathbf{s}_U(\widetilde{\mathbf{y}}))+\mathbf{c}_U(\widetilde{\mathbf{y}})'\,\beta \tag{30}$$

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 α. 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:

$$[\mathbf{c}_L(A)-\mathbf{c}_U(B)]'\,\beta \;\leq\; \ln \mathbb{P}(A)-\ln \mathbb{P}(B) \tag{31}$$

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

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \quad \leq \quad [\mathbf{c}_U(A) - \mathbf{c}_L(B)]' \beta \tag{32}$$

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 $\mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$. 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, $\ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \alpha, \beta) = \mathbf{s}_L(\widetilde{\mathbf{y}})' \mathbf{g}_\alpha + \mathbf{c}_L(\widetilde{\mathbf{y}})' \beta$ and $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \alpha, \beta) = \mathbf{s}_U(\widetilde{\mathbf{y}})' \mathbf{g}_\alpha + \mathbf{c}_U(\widetilde{\mathbf{y}})' \beta,$ 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 $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$. These bounds are:

$$\begin{cases} \ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \equiv & \ln p_\alpha(y_{10}, y_{20}) + \sum_{t=1}^T \ln L(\mathbf{y}_t | \mathbf{y}_{t-1}; \alpha, \beta) \\ \ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \equiv & \ln p_\alpha(y_{10}, y_{20}) + \sum_{t=1}^T \ln U(\mathbf{y}_t | \mathbf{y}_{t-1}; \alpha, \beta) \end{cases} \tag{33}$$

For outcomes $(0,0)$ and $(1,1)$, the upper bounds and the lower bounds are the same and they are the probabilities in equation (22). For outcomes $(0,1)$ and $(1,0)$, the upper bounds $U(\mathbf{y}_t | \mathbf{y}_{t-1}; \alpha, \beta)$ are the ones in equation (23), and the lower bounds $L(\mathbf{y}_t | \mathbf{y}_{t-1}; \alpha, \beta)$ come from equations (24) and (25).

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 (21), the log-probability of a market history has lower bounds $\ln \mathbb{P}_{L\{E,W\}}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$ *and*

$\ln \mathbb{P}_{L\{S,N\}}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$ *and upper bound* $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \alpha, \beta)$ *which have the following expressions:*

$$\begin{aligned}
\ln \mathbb{P}_{L\{E,W\}}(\widetilde{\mathbf{y}} \mid \alpha, \beta) &= \mathbf{s}^1(\widetilde{\mathbf{y}})' \mathbf{g}_\alpha^1 + \left[T_1^{(1)},\, T_1^{(1)},\, C_{11},\, C_{12}\right] \mathbf{g}_\alpha^2 \\
&+ C_{11}\,\beta_{11} + C_{22}\,\beta_{22} + T_1^{(1)}\,\gamma_1 + T^{(1,1)}\,\gamma_2 \\
\ln \mathbb{P}_{L\{S,N\}}(\widetilde{\mathbf{y}} \mid \alpha, \beta) &= \mathbf{s}^1(\widetilde{\mathbf{y}})' \mathbf{g}_\alpha^1 + \left[T_2^{(1)},\, T_2^{(1)},\, C_{21},\, C_{22}\right] \mathbf{g}_\alpha^2 \\
&+ C_{11}\,\beta_{11} + C_{22}\,\beta_{22} + T^{(1,1)}\,\gamma_1 + T_2^{(1)}\,\gamma_2 \\
\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \alpha, \beta) &= \mathbf{s}^1(\widetilde{\mathbf{y}})' \mathbf{g}_\alpha^1 + \left[T_2^{(1)},\, T_1^{(1)},\, C_{21},\, C_{12}\right] \mathbf{g}_\alpha^2 \\
&+ C_{11}\,\beta_{11} + C_{22}\,\beta_{22} + T^{(1,1)}\,[\gamma_1 + \gamma_2]
\end{aligned} \tag{34}$$

where $\mathbf{g}_\alpha^1$ *and* $\mathbf{g}_\alpha^2$ *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 8.

PROPOSITION 8. Consider the myopic complete information dynamic game with contemporaneous effects in equation (21). 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* $\Delta(A, B, \beta_{11}, \beta_{22})$ *be* $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) - [C_{11}(A) - C_{11}(B)]\,\beta_{11} - [C_{22}(A) - C_{22}(B)]\,\beta_{22}$.

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

$$\Delta(A, B, \beta_{11}, \beta_{22}) \geq \left[T_1^{(1)}(A) - T^{(1,1)}(B)\right]\,\gamma_1 + \left[T^{(1,1)}(A) - T^{(1,1)}(B)\right]\,\gamma_2 \tag{35}$$

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

$$\Delta(A, B, \beta_{11}, \beta_{22}) \leq \left[T^{(1,1)}(A) - T_1^{(1)}(B)\right]\,\gamma_1 + \left[T^{(1,1)}(A) - T^{(1,1)}(B)\right]\,\gamma_2 \tag{36}$$

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

$$\Delta(A, B, \beta_{11}, \beta_{22}) \geq \left[T^{(1,1)}(A) - T^{(1,1)}(B)\right]\,\gamma_1 + \left[T_2^{(1)}(A) - T^{(1,1)}(B)\right]\,\gamma_2 \tag{37}$$

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

$$\Delta(A, B, \beta_{11}, \beta_{22}) \leq \left[T^{(1,1)}(A) - T^{(1,1)}(B)\right]\,\gamma_1 + \left[T^{(1,1)}(A) - T_2^{(1)}(B)\right]\,\gamma_2 \tag{38}$$

Based on these inequalities, we can find pairs of market histories – A and B – that set identify the parameters β_{11}, β_{22}, 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 $\Delta(A, B, \beta_{11}, \beta_{22}) = \ln \mathbb{P}(A) - \ln \mathbb{P}(B) - \beta_{22}$. We now check conditions (i) to (iv) in Proposition 8.

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

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \geq \beta_{22} + \gamma_1 + \gamma_2 \tag{39}$$

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

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \leq \beta_{22} + \gamma_1 \tag{40}$$

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

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \geq \beta_{22} + \gamma_1 + \gamma_2 \tag{41}$$

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 $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \geq \beta_{11} + \gamma_1 + \gamma_2$ and $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \leq \beta_{11} + \gamma_2$

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.

2.6 Incomplete information and full contemporaneous effects

We now consider the myopic incomplete information game where with $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$. The Bayesian Nash equilibrium (BNE) at period t can be described as a pair of probabilities, P_1 and P_2, that solve the system of equations:

$$\begin{cases} P_1 & = & \Lambda\left(\alpha_1 + \gamma_1\ P_2 + \beta_{11}\ y_{1t-1}\right) \\ \\ P_2 & = & \Lambda\left(\alpha_2 + \gamma_2\ P_1 + \beta_{22}\ y_{2t-1}\right) \end{cases} \tag{42}$$

Then, given a solution $(P_1(\mathbf{y}_{t-1};\alpha), P_2(\mathbf{y}_{t-1};\alpha))$, the model implies the outcome probabilities $\mathbb{P}((y_{1t}, y_{2t}) = (j,k) \mid \mathbf{y}_{t-1};\alpha) = P_1(\mathbf{y}_{t-1};\alpha)^j\ P_2(\mathbf{y}_{t-1};\alpha)^k$, for any $j,k \in \{0,1\}$.

There are two issues with these outcome probabilities. First, in general, the model does not have a unique BNE. There may be multiple pairs of probabilities that solve the system of equations in (42) such that the model does not have a unique prediction for its outcome. Second, even under uniqueness, the equilibrium probabilities do not have a logit structure.

We deal with these issues using a bounds approach. However, the way in which we derive bounds for this incomplete information game is different than with complete information. We use the followings Lemmas 3 to 5.

LEMMA 3. Let $\Lambda(x)$ be the Logistic function. Then, then $\Lambda(x) - \Lambda(y) \leq \frac{1}{4}(x-y)$ if $x > y$ and $\Lambda(x) - \Lambda(y) \geq \frac{1}{4}(x-y)$ if $x < y$. ■

Proof. By the mean value theorem, $\Lambda(x) - \Lambda(y) = (x-y)\ \Lambda'(\xi)$, for $\xi \in [y,x]$. For the Logistic function, $\Lambda'(x) = \Lambda(x)(1-\Lambda(x))$. Since $\Lambda(x) \in [0,1]$, we have $\Lambda'(x) \leq 1/4$ for all $x \in \mathbb{R}$. Therefore, if $x > y$, we have the first inequality, and if if $x < y$ we have the second inequality. ■

LEMMA 4. Suppose that the following conditions hold: (i) non-negative switching costs, $\beta_{11} \geq 0$ and $\beta_{22} \geq 0$; (ii) negative strategic interactions, $\gamma_1 \leq 0$ and $\gamma_2 \leq 0$; and (iii) upper bound on strategic effects, $\gamma_1\gamma_2 \leq 16$. Under these conditions, the following inequalities hold for any for any $(\alpha_1, \alpha_2) \in \mathbb{R}^2$:

$$P_1(1,0,\alpha) - P_1(0,0,\alpha) \geq 0;\ P_1(0,1,\alpha) - P_1(0,0,\alpha) \leq 0$$

$$P_2(1,0,\alpha) - P_2(0,0,\alpha) \leq 0;\ P_2(0,1,\alpha) - P_2(0,0,\alpha) \geq 0$$

Proof. See the Appendix. ■

LEMMA 5. For player i, let $\Delta^{(i)}_{y_1,y_2}$ be the average effect $\int [P_i(y_1,y_2,\alpha) - P_i(0,0,\alpha)]dF(\alpha)$. Under *conditions (i) to (iii) in Lemma 4, the following inequality restrictions hold:*

$$\Delta^{(2)}_{1,0} \geq \frac{\gamma_2}{4}\Delta^{(1)}_{1,0}$$
$$\Delta^{(1)}_{1,0} \leq \frac{\gamma_1}{4}\Delta^{(2)}_{1,0} + \frac{\beta_{11}}{4}$$
$$\Delta^{(1)}_{0,1} \geq \frac{\gamma_1}{4}\Delta^{(2)}_{0,1}$$
$$\Delta^{(2)}_{0,1} \leq \frac{\gamma_2}{4}\Delta^{(1)}_{0,1} + \frac{\beta_{22}}{4}$$

In turn, these inequalities imply the following:

$$0 \geq \Delta^{(2)}_{1,0} \geq \frac{\gamma_2\beta_{11}/16}{1-\gamma_1\gamma_2/16}$$
$$0 \leq \Delta^{(1)}_{1,0} \leq \frac{\beta_{11}/4}{1-\gamma_1\gamma_2/16}$$
$$0 \geq \Delta^{(1)}_{0,1} \geq \frac{\gamma_1\beta_{22}/16}{1-\gamma_1\gamma_2/16}$$
$$0 \leq \Delta^{(2)}_{0,1} \leq \frac{\beta_{22}/4}{1-\gamma_1\gamma_2/16}$$

Proof of Lemma 5. See the Appendix. ■

PROPOSITION 9. In the myopic, incomplete information game with full interactions, under conditions $\beta_{11} \geq 0$, $\beta_{22} \geq 0$, $\gamma_1 \leq 0$, $\gamma_2 \leq 0$, and $\gamma_1\gamma_2 \leq 16$, there are market histories that provide informative bounds on the structural parameters. All the parameters are partially identified. ■

Proof of Proposition 9. To get bounds for the parameters $(\gamma_1, \gamma_2, \beta_{11}, \beta_{22})$, we now link the probability of the market histories to the average effects $\Delta^{(1)}_{1,0}$, $\Delta^{(1)}_{0,1}$, $\Delta^{(2)}_{1,0}$, and $\Delta^{(2)}_{0,1}$. Consider the histories $A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, and $D = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. We can see that

$$\mathbb{P}(A) = \int p_\alpha(0,0)P_1((0,0))^2(1-P_1((1,0))(1-P_2((0,0))^2(1-P_2((1,0)))dF(\alpha)$$

and

$$\mathbb{P}(B) = \int p_\alpha(0,0)P_1((0,0))(1-P_1((0,0)))P_1((1,0))(1-P_2((0,0)))^2(1-P_2((1,0)))dF(\alpha)$$

and therefore, under conditions (i) to (iii) in Lemma 4, we have:

$$\begin{aligned}
\mathbb{P}(B)-\mathbb{P}(A) &= \int p_\alpha(0,0)(1-P_2((0,0)))^2(1-P_2((1,0)))P_1((0,0))\Big[P_1((1,0))-P_1((0,0))\Big]dF(\alpha) \\
&\leq \int \Big[P_1((1,0))-P_1((0,0))\Big]dF(\alpha) \\
&= \Delta^{(1)}_{1,0} \leq \frac{\beta_{11}/4}{1-\gamma_1\gamma_2/16}
\end{aligned}$$

where the first inequality holds since $\{p_\alpha(0,0), P_2((0,0)), P_2((1,0)), P_1((0,0))\} \in [0,1]^4$.

We also have:

$$\begin{aligned}
\mathbb{P}(C) - \mathbb{P}(D) &= \int p_\alpha(0,0)P_1((0,0))(1-P_1((0,0)))(1-P_1((1,0)))(1-P_2((0,0)))\Big[P_2((1,0)) - P_2((0,0))\Big]dF(\alpha) \\
&\geq \frac{1}{4}\int \Big[P_2(1,0) - P_2(0,0)\Big]dF(\alpha) \\
&= \frac{1}{4}\Delta^{(2)}_{1,0} \geq \frac{1}{4}\frac{\gamma_2\beta_{11}/16}{1-\gamma_1\gamma_2/16}
\end{aligned}$$

where the first inequality holds because we know $P_1((0,0))(1-P_1((0,0))) \leq 1/4$ and all other probabilities are in $[0,1]$.

Similarly, consider the histories $E = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$, $F = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$, $G = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, and $H = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$. We then have

$$\begin{aligned}
\mathbb{P}(F) - \mathbb{P}(E) &= \int p_\alpha(0,0)(1-P_1((0,0)))^2(1-P_1((0,1)))P_2((0,0))\Big[P_2((0,1)) - P_2((0,0))\Big]dF(\alpha) \\
&\leq \int \Big[P_2((0,1)) - P_2((0,0))\Big]dF(\alpha) \\
&= \Delta^{(2)}_{0,1} \leq \frac{\beta_{22}/4}{1-\gamma_1\gamma_2/16}
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{P}(G) - \mathbb{P}(H) &= \int p_\alpha(0,0)(1-P_1((0,0)))(1-P_2((0,0)))P_2((0,0))(1-P_2((0,1)))\Big[P_1((0,1)) - P_1((0,0))\Big]dF(\alpha) \\
&\geq \frac{1}{4}\int \Big[P_1((0,1)) - P_1((0,0))\Big]dF(\alpha) \\
&= \frac{1}{4}\Delta^{(1)}_{0,1}\pi^a_{01} \geq \frac{1}{4}\frac{\gamma_1\beta_{22}/16}{1-\gamma_1\gamma_2/16}
\end{aligned}$$

3 Identification of dynamic games with forward-looking players

3.1 Framework

Every period t, the two players choose simultaneously their binary actions in market m – $y_{1mt} \in \{0,1\}$ and $y_{2mt} \in \{0,1\}$ – to maximize their respective expected and discounted intertemporal payoffs: $\mathbb{E}_t \left[\sum_{s=0}^{\infty} \delta_{im}^{s} \; U_{im,t+s}(y_{im,t+s}, y_{jm,t+s})\right]$, where $\delta_{im} \in (0,1)$ is the discount factor of player i in market m, and $U_{imt}(y_i, y_j)$ is the one-period payoff. The one-period payoff has the same structure as in equation (1), such that the utility difference $\Delta U_{imt} \equiv U_{imt}(1, y_{jmt}) - U_{imt}(0, y_{jmt})$ can be represented as in (2). The unobservable variables – $\varepsilon_{imt}(0)$ and $\varepsilon_{imt}(1)$ – have the same interpretation and the same properties as in the myopic model.

We distinguish two types of models: *complete information games* where all the variables including ε's are common knowledge to all the firms; and *incomplete information games* where the ε_i's are private information of player i but unknown to the other players in the game. Note that the fixed effects are assumed common knowledge in both cases.

Following most of the empirical literature on dynamic discrete games, we assume that players' decisions come from a *Markov Perfect Equilibrium* (MPE). This implies that players' strategies depend only on payoff relevant state variables. At any period t, the action of player i is a function of the variables known by this player that affects her payoff, or the payoff of other players, at period t. For the incomplete information game, this means that $y_{imt} = \sigma_{im}(\varepsilon_{imt}, \mathbf{y}_{mt-1})$, where $\sigma_i(.)$ represents the strategy function of player i. For the complete information game, the strategy function of a player depends also on the current ε variable of the other player, such that $y_{imt} = \sigma_{im}(\varepsilon_{imt}, \varepsilon_{jmt}, \mathbf{y}_{mt-1})$.

A player takes the strategies of the other players as given and chooses her strategy to maximize her own intertemporal value. This best response function is the solution to a single-agent dynamic programming problem.

A *MPE* in the game of *complete information* is a solution in (y_1, y_2) to the system of best response conditions:

$$\left\{ \begin{array}{lcl} y_{1mt} & = & 1\left\{\alpha_{1m} + \gamma_1 \; y_{2mt} + \beta_{11} \; y_{1mt-1} + \beta_{12} \; y_{2mt-1} + \widetilde{v}_{1m}(y_{2mt}) - \varepsilon_{1mt} \geq 0\right\} \\ & & \\ y_{2mt} & = & 1\left\{\alpha_{2m} + \gamma_2 \; y_{1mt} + \beta_{21} \; y_{1mt-1} + \beta_{22} \; y_{2mt-1} + \widetilde{v}_{2m}(y_{1mt}) - \varepsilon_{2mt} \geq 0\right\} \end{array} \right. \tag{43}$$

where $\widetilde{v}_{im}\left(y_{jmt}\right) \equiv v_{im}\left(1, y_{jmt}\right) - v_{im}\left(0, y_{jmt}\right)$ is the difference between the continuation value of choosing alternative 1 and the continuation value of choosing alternative 0. In the game of complete information, these continuation values depend on the current decision of the other player – y_{jmt} –which is known to player i at period t. Importantly – for our identification results –

these continuation values do not depend on the state variables at period t. Given players' choices át period t – which are common knowledge – their choices at period $t-1$ do not contain any additional information that is relevant to predict future payoffs at periods $t+1$ and beyond.

In the game of incomplete information, variable ε_{imt} becomes private information of player i. This implies that a player does not know with centainty the current decision of the other player – she knows only a probability distribution of that decision. A *MPE* in the game of *incomplete information* is a solution in the choice probabilities functions $(P_{1m}(\mathbf{y}_{mt-1}),\ P_{2m}(\mathbf{y}_{mt-1}))$ to the system of equations:

$$\left\{ \begin{array}{lcl} P_{1mt} & = & \Lambda\left(\alpha_{1m} + \beta_{11}\ y_{1mt-1} + \beta_{12}\ y_{2mt-1} + [1 - P_{2mt}]\ \widetilde{v}_{1m}(0) + P_{2mt}\ \widetilde{v}_{1m}(1)\right) \\ \\ P_{2mt} & = & \Lambda\left(\alpha_{2m} + \beta_{21}\ y_{1mt-1} + \beta_{22}\ y_{2mt-1} + [1 - P_{1mt}]\ \widetilde{v}_{2m}(0) + P_{1mt}\ \widetilde{v}_{2m}(1)\right) \end{array} \right. \tag{44}$$

where, for $i \in \{1, 2\}$, $P_{imt} \equiv P_{imt}(\mathbf{y}_{mt-1}) = \mathbb{P}\left[y_{imt} = 1 \mid \alpha_m,\ y_{1mt-1},\ y_{2mt-1}\right]$; and $\widetilde{v}_{im}(y_{jmt})$ is still the difference between the continuation values of alternatives 1 and 0.

It is important to distinguish two manifestations of the problem of multiple equilibria. In general, given the primitives, the model may have multiple strategy functions – $\sigma_{1m}(.)$ and $\sigma_{2m}(.)$ – that satisfy the system of best response restrictions characterizing the equilibrium of the model. This is a general description of phenomenom of multiple equilibria, and it may happen with complete or with incomplete information. There is a more specific manifestation of multiplicity of equilibria that appears only in games of completete information. Even if we fix the continuation value functions – $v_{1m}(.)$ and $v_{2m}(.)$ – and the value of the state variables – $\mathbf{y}_{mt-1}$, ε_{1mt}, and ε_{2mt} – there are regions in the space of $(\varepsilon_{1mt}, \varepsilon_{2mt})$ with multiple predictions about the best response values of (y_{1mt}, y_{2mt}). This is the same type problem with multiple equilibria that appears in static games of complete information – Bresnahan and Reiss (1990) or Tamer (2003).

The sampling framework is the same as for the myopic model. For notational simplicity, we omit the market subindex m for the rest of this section. We use the vector $\widetilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, .., T)$ to represent a *market history*.

3.2 Forward-looking, complete information, triangular dynamic game

Consider the complete information game in equation (43). It is convenient to represent this model as follows:

$$y_{it} \quad = \quad 1\left\{\widetilde{\alpha}_i + \beta_{ii}\ y_{i,t-1} + \widetilde{\gamma}_{i\alpha}\ y_{jt} - \varepsilon_{it} \geq 0\right\} \tag{45}$$

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_{11} and β_{22}; (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 $\beta_{12} = \gamma_1 = 0$. 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:

$$y_{1t} \quad = \quad 1\left\{ \varepsilon_{1t} \leq \quad \alpha_1 + \beta_{11}\ y_{1t-1} + \widetilde{v}_{1\alpha} \ \right\} \tag{46}$$

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_{11} is identified from the maximization of the condtional likelihood function.

We now establish the point identification of parameters β_{11} and β_{22}. The best response of player 2 in this triangular model is:

$$y_{2t} \quad = \quad 1\left\{ \varepsilon_{2t} \leq \quad \widetilde{\alpha}_2 + \beta_{21}\ y_{1,t-1} + \beta_{22}\ y_{2,t-1} + \widetilde{\gamma}_{2\alpha}\ y_{1t} \ \right\} \tag{47}$$

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 (46) and (47), the log-probability of the market history $\widetilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, .., T)$ has the following structure:

$$\begin{aligned} \ln \mathbb{P}(\widetilde{\mathbf{y}}|\alpha,\beta) \quad = \quad & \ln p_\alpha(y_{10}, y_{20}) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} + \widetilde{\gamma}_{2\alpha}\ T^{(1,1)} + \sum_{t=1}^{T} \sigma_{\alpha 1}(y_{1t-1}) + \sigma_{\alpha 2}(y_{1t}, y_{2t-1}) \\ + \quad & \beta_{11}\ C_{11} + \beta_{22}\ C_{22} \end{aligned} \tag{48}$$

where $\sigma_{\alpha 1}(y_1) \equiv -\ln[1 + \exp\{\alpha_1 + \beta_{11}\ y_1\}]$ and $\sigma_{\alpha 2}(y_1, y_2) \equiv -\ln[1 + \exp\{\alpha_2 + \widetilde{\gamma}_{2\alpha}\ y_1 + \beta_{22}\ y_2\}]$. We can rewrite this equation for the log-probability of a market history as $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$ as $\mathbf{s}(\widetilde{\mathbf{y}})'$ $\mathbf{g}_\alpha + \mathbf{c}(\widetilde{\mathbf{y}})'\ \beta^*$, with

$$\left\{ \begin{array}{rcl} \mathbf{s}(\widetilde{\mathbf{y}})' & = & [1,\ y_{10},\ y_{20},\ y_{10}y_{20} \quad ; \quad 1,\ y_{1T},\ y_{2T},\ y_{1T}y_{2T} \quad ; \quad T,\ T_1^{(1)},\ T_2^{(1)},\ T^{(1,1)} \quad ; \quad C_{12}] \\ \mathbf{c}(\widetilde{\mathbf{y}})' & = & [C_{11},\ C_{22}] \\ \beta^{*\prime} & = & [\beta_{11},\ \beta_{22}] \end{array} \right. \tag{49}$$

PROPOSITION 10. For the forward-looking, complete information, triangular ('Stackelberg') dynamic game described by equations (46) and (47): (A) The vector $\mathbf{s}(\widetilde{\mathbf{y}}) = [1,\ y_{10},\ y_{20},\ y_{10}y_{20},\ y_{1T},\ y_{2T},\ y_{1T}y_{2T},\ T,\ T_1^{(1)},\ T_2^{(1)},\ T^{(1,1)},\ C_{12}]'$ *is a minimal sufficient statistic for* α *such that* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{u}(\widetilde{\mathbf{y}}), \alpha, \beta)$ *does not depend on* α*. (B)* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{c}(\widetilde{\mathbf{y}}), \beta) = \mathbf{s}(\widetilde{\mathbf{y}})' \beta^* - \ln(\sum_{\widetilde{\mathbf{y}}':\mathbf{s}(\widetilde{\mathbf{y}}')=\mathbf{s}(\widetilde{\mathbf{y}})} \exp\{\mathbf{c}(\widetilde{\mathbf{y}}')'\beta^*\})$ *with* $\mathbf{c}(\widetilde{\mathbf{y}}) = [C_{11},\ C_{22}]'$ *and* $\beta^* = [\beta_{11},\ \beta_{22}]'$*. (C) For* $T \geq 3$*, there are histories* $\widetilde{\mathbf{y}}$ *such that* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \beta)$ *identifies the vector of parameters of interest* β^**.* ■

EXAMPLE 10. The same histories in Example 3 that – in the myopic, complete information, triangular model – identify parameters β_{11} and β_{22}, 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_{11}; 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_{22}. ■

3.3 Forward-looking, complete information: bounds approach

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

$$\begin{cases} y_{1t} = 1\{\ \varepsilon_{1t} \leq \ \widetilde{\alpha}_1 + \beta_{11}\ y_{1,t-1} + \widetilde{\gamma}_{1\alpha}\ y_{2t}\ \} \\ \\ y_{2t} = 1\{\ \varepsilon_{2t} \leq \ \widetilde{\alpha}_2 + \beta_{22}\ y_{2,t-1} + \widetilde{\gamma}_{2\alpha}\ y_{1t}\ \} \end{cases} \tag{50}$$

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 11 establsihes that these parameters are partially identified.

PROPOSITION 11. Consider the forward-looking complete information dynamic game with contemporaneous effects in equation (50). Under conditions $\beta_{11} \geq 0$, $\beta_{22} \geq 0$, $\tilde{\gamma}_{1\alpha} \leq 0$, *and* $\tilde{\gamma}_{2\alpha} \leq 0$, *there are market histories that provide informative bounds on the parameters* β_{11} *and* β_{22}*. These parameters are partially identified.* ■

Proof of Proposition 11. Denote the following terms:

$$\begin{aligned}
\sigma_{\alpha_1}(y_{1t-1}, y_{2t}) &= -\ln\{1+\exp(\tilde{\alpha}_1 + \beta_{11}y_{1t-1} + \tilde{\gamma}_{1\alpha}y_{2t})\} \\
\sigma_{\alpha_2}(y_{1t}, y_{2t-1}) &= -\ln\{1+\exp(\tilde{\alpha}_2 + \beta_{22}y_{2t-1} + \tilde{\gamma}_{2\alpha}y_{1t}\} \\
\Delta\sigma_{\alpha_1}(1,0) &= \sigma_{\alpha_1}(1,0) - \sigma_{\alpha_1}(0,0) \\
\Delta\sigma_{\alpha_1}(0,1) &= \sigma_{\alpha_1}(0,1) - \sigma_{\alpha_1}(0,0) \\
\Delta\sigma_{\alpha_2}(1,0) &= \sigma_{\alpha_2}(1,0) - \sigma_{\alpha_2}(0,0) \\
\Delta\sigma_{\alpha_2}(0,1) &= \sigma_{\alpha_2}(0,1) - \sigma_{\alpha_2}(0,0) \\
\Delta^2\sigma_{\alpha_1} &= \sigma_{\alpha_1}(1,1) - \sigma_{\alpha_1}(1,0) - \sigma_{\alpha_1}(0,1) + \sigma_{\alpha_1}(0,0) \\
\Delta^2\sigma_{\alpha_2} &= \sigma_{\alpha_2}(1,1) - \sigma_{\alpha_2}(1,0) - \sigma_{\alpha_2}(0,1) + \sigma_{\alpha_2}(0,0)
\end{aligned}$$

Under the conditions of Proposition 11, 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:

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

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

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

Define the statistics $s^1(\tilde{y}) = [T, T_1^{(1)}, T_2^{(1)}]$

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

$$\Delta(A, B, \beta_{11}, \beta_{22}) \leq 0$$

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, $[T_2^{(1)}(A), T_1^{(1)}(A), S_{21}(A), S_{12}(A)] - [T_1^{(1)}(B), T_1^{(1)}(B), S_{11}(B), S_{12}(B)] \leq 0$.

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

$$\Delta(A, B, \beta_{11}, \beta_{22}) \leq 0$$

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 $L\{E, N\}$:

$$\Delta(A, B, \beta_{11}, \beta_{22}) \leq 0$$

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, $[T_2^{(1)}(A), T_1^{(1)}(A), C_{21}(A), C_{12}(A)] - [T^{(1,1)}(B), T^{(1,1)}(B), R_1^{(1,1)}($ 0.

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

$$\Delta(A, B, \beta_{11}, \beta_{22}) \leq 0$$

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 11 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 11such 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 $\ln P(A) - \ln P(B) \leq \beta_{22}$

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 $\ln P(A) - \ln P(B) \leq \beta_{11}$. ■

3.4 Identification of dynamic games of incomplete information

Consider the two-player binary choice model of incomplete information. We can represent the best response decisions of the players as follows:

$$\begin{cases} y_{1t} &=& 1\{\alpha_1 + \beta_{11}\ y_{1,t-1} + \beta_{12}\ y_{2,t-1} + \widetilde{v}_{1\alpha}(y_{1,t-1}, y_{2,t-1}) - \varepsilon_{1t} \geq 0\} \\ y_{2t} &=& 1\{\alpha_2 + \beta_{21}\ y_{1,t-1} + \beta_{22}\ y_{2,t-1} + \widetilde{v}_{2\alpha}(y_{1,t-1}, y_{2,t-1}) - \varepsilon_{2t} \geq 0\} \end{cases} \tag{51}$$

where $\widetilde{v}_{i\alpha}(y_{1,t-1}, y_{2,t-1}) \equiv v_{i\alpha}(1; y_{1,t-1}, y_{2,t-1}) - v_{i\alpha}(0; y_{1,t-1}, y_{2,t-1})$. The continuation value function $\widetilde{v}_{i\alpha}(y_{1,t-1}, y_{2,t-1})$ depends on $(y_{1,t-1}, y_{2,t-1})$ because player i does not know the current choice of the other player j. She has to predict the choice of the other player, but this prediction (the equilibrium choice probability) depends on $(y_{1,t-1}, y_{2,t-1})$ and on the incidental parameters.

Equation (51) already shows that the structural parameters β cannot be identified in this model. In the log-probability of a market history, these statistics are associated with the statistics

$S_{ij} \equiv \sum_{t=1}^{T} y_{it}\ y_{jt-1}$ for $i, j \in \{1, 2\}$. However, in the log-probability of any market history, these statistics S_{ij} also appear associated with the incidental parameters through the terms $\sum_{t=1}^{T} y_{it}$ $\widetilde{v}_{i\alpha}\left(y_{1,t-1}, y_{2,t-1}\right)$. Therefore, without further restrictions, the structural parameters β are not identified in this model.

We consider a 'Stackelberg' version of this model where every period t, player 1 decides first and then player 2 makes her decision given that she knows the decision of player 1.

4 Estimation and inference

TBW

5 Empirical application

TBW

6 Conclusions

TBW

APPENDIX

Proof of Lemma 1. Given the structure for the lower bound – $\ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \alpha, \beta) = \mathbf{s}_L' \mathbf{g}_\alpha + \mathbf{c}_L' \beta$ – and for the upper bound – $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \alpha, \beta) = \mathbf{s}_U' \mathbf{g}_\alpha + \mathbf{c}_U' \beta$ – we have that:

$$\exp\left\{\mathbf{s}_L' \mathbf{g}_\alpha + \mathbf{c}_L' \beta\right\} \leq \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha) \leq \exp\left\{\mathbf{s}_U' \mathbf{g}_\alpha + \mathbf{c}_U' \beta\right\} \tag{A.1.1}$$

Integrating the inequalities in (A.1.1) over the distribution of α we have that the inequalities still hold and they take the following form:

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

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

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

Proof of Lemma 2 (Complete Information). For the derivations below, we use the following definitions: $\sigma_{\alpha 1}(y_{1t-1}, y_{2t}) \equiv -\ln[1 + \exp\{\alpha_1 + \beta_{11}\ y_{1t-1} + \gamma_1 y_{2t}\}]$ and $\sigma_{\alpha 2}(y_{1t}, y_{2t-1}) \equiv -\ln[1 + \exp\{\alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2 y_{1t}\}]$, and

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

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:

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

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 $\ln \mathbb{P}_{L\{E,W\}}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$. To obtain this lower bound, we use the bounds $L^{\{E,SE\}}(0,1|\mathbf{y}_{t-1};\alpha) \equiv [1 - \Lambda(\alpha_1 + \beta_{11}\ y_{1t-1})]\ \Lambda(\alpha_2 + \beta_{22}\ y_{2t-1})$ and $L^{\{W,NW\}}(1,0|\mathbf{y}_{t-1};\alpha) \equiv \Lambda(\alpha_1 + \beta_{11}\ y_{1t-1} + \gamma_1)$

$[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2)]$ for the choice probabilities. Then,

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha,\beta) \geq\; & \ln p_\alpha(y_{10},y_{20}) \\
+ & \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \;\; (\ln[1-\Lambda(\alpha_1+\beta_{11}\ y_{1t-1})] + \ln[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})]) \\
+ & \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \;\; (\ln[1-\Lambda(\alpha_1+\beta_{11}\ y_{1t-1})] + \ln\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})) \\
+ & \sum_{t=1}^{T} y_{1t}(1-y_{2t}) \;\; (\ln\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1) + \ln[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2)]) \\
+ & \sum_{t=1}^{T} y_{1t}y_{2t} \;\; (\ln\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1) + \ln\Lambda(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2))
\end{aligned} \tag{A.3}$$

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

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha,\beta) \geq\; & \ln p_\alpha(y_{10},y_{20}) \\
+ & \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \;\; [\sigma_{\alpha 1}(y_{1t-1},0)+\sigma_{\alpha 2}(0,y_{2t-1})] \\
+ & \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \;\; [\sigma_{\alpha 1}(y_{1t-1},0)+\sigma_{\alpha 2}(0,y_{2t-1})+\alpha_2+\beta_{22}\ y_{2t-1}] \\
+ & \sum_{t=1}^{T} y_{1t}(1-y_{2t}) \;\; [\sigma_{\alpha 1}(y_{1t-1},1)+\sigma_{\alpha 2}(1,y_{2t-1})+\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1] \\
+ & \sum_{t=1}^{T} y_{1t}y_{2t} \;\; [\sigma_{\alpha 1}(y_{1t-1},1)+\sigma_{\alpha 2}(1,y_{2t-1})+\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1+\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2]
\end{aligned} \tag{A.4}$$

Grouping terms, we have:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha,\beta) \geq\; & \ln p_\alpha(y_{10},y_{20}) \\
+ & \sum_{t=1}^{T} (1-y_{1t}) \;\; [\sigma_{\alpha 1}(y_{1t-1},0)+\sigma_{\alpha 2}(0,y_{2t-1})] \\
+ & \sum_{t=1}^{T} y_{2t} \;\; [\alpha_2+\beta_{22}\ y_{2t-1}] \\
+ & \sum_{t=1}^{T} y_{1t} \;\; [\sigma_{\alpha 1}(y_{1t-1},1)+\sigma_{\alpha 2}(1,y_{2t-1})+\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1] \\
+ & \sum_{t=1}^{T} y_{1t}y_{2t} \;\; [\gamma_2]
\end{aligned} \tag{A.5}$$

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 $\ln \mathbb{P}_{L\{E,W\}}(\widetilde{\mathbf{y}} \mid \alpha,\beta)$:

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

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

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

(b) Lower Bound $\ln \mathbb{P}_{L\{S,N\}}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$. To obtain this lower bound, we use the bounds $L^{\{S,SE\}}(0,1|\mathbf{y}_{t-1};\alpha) \equiv [1-\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1)]\ \Lambda(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2)$ and $L^{\{N,NW\}}(1,0|\mathbf{y}_{t-1};\alpha) \equiv \Lambda(\alpha_1+\beta_{11}\ y_{1t-1})$ $[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})]$. Then,

$$
\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) &\geq \ln p_\alpha(y_{10}, y_{20}) \\
&+ \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \ \ (\ln[1-\Lambda(\alpha_1+\beta_{11}\ y_{1t-1})] + \ln[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})]) \\
&+ \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \ (\ln[1-\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1)] + \ln\Lambda(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2)) \\
&+ \sum_{t=1}^{T} y_{1t}\,(1-y_{2t}) \ \ (\ln\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}) + \ln[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})]) \\
&+ \sum_{t=1}^{T} y_{1t} y_{2t} \ \ (\ln\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1) + \ln\Lambda(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2))
\end{aligned} \tag{A.8}
$$

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

$$
\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) &\geq \ln p_\alpha(y_{10}, y_{20}) \\
&+ \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1})] \\
&+ \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2] \\
&+ \sum_{t=1}^{T} y_{1t}\,(1-y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1}) + \alpha_1 + \beta_{11}\ y_{1t-1}] \\
&+ \sum_{t=1}^{T} y_{1t} y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_1 + \beta_{11}\ y_{1t-1} + \gamma_1 + \alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2]
\end{aligned} \tag{A.9}
$$

Grouping terms, we have:

$$
\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) &\geq \ln p_\alpha(y_{10}, y_{20}) \\
&+ \sum_{t=1}^{T} (1-y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1})] \\
&+ \sum_{t=1}^{T} y_{1t} \ \ [\alpha_1 + \beta_{11}\ y_{1t-1}] \\
&+ \sum_{t=1}^{T} y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2] \\
&+ \sum_{t=1}^{T} y_{1t} y_{2t} \ \ [\gamma_1]
\end{aligned} \tag{A.10}
$$

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 $\ln \mathbb{P}_{L\{S,N\}}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right)$:

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

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

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

(c) Lower Bound $\ln \mathbb{P}_{L\{E,N\}}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right)$. To obtain this lower bound, we use the bounds $L^{\{E,SE\}}(0,1|\mathbf{y}_{t-1};\alpha) \equiv [1-\Lambda(\alpha_1+\beta_{11}\ y_{1t-1})]\ \Lambda(\alpha_2+\beta_{22}\ y_{2t-1})$ and $L^{\{N,NW\}}(1,0|\mathbf{y}_{t-1};\alpha) \equiv \Lambda(\alpha_1+\beta_{11}\ y_{1t-1})\ [1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})]$ for the choice probabilities. Then,

$$\begin{array}{rcl} \ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right) & \geq & \ln p_{\alpha}\left(y_{10},y_{20}\right) \\ & + & \sum\limits_{t=1}^{T}\left(1-y_{1t}\right)\left(1-y_{2t}\right)\ \left(\ln\left[1-\Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}\right)\right]+\ln\left[1-\Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}\right)\right]\right) \\ & + & \sum\limits_{t=1}^{T}\left(1-y_{1t}\right)y_{2t}\ \left(\ln\left[1-\Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}\right)\right]+\ln\Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}\right)\right) \\ & + & \sum\limits_{t=1}^{T}y_{1t}\left(1-y_{2t}\right)\ \left(\ln\Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}\right)+\ln\left[1-\Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}\right)\right]\right) \\ & + & \sum\limits_{t=1}^{T}y_{1t}y_{2t}\ \left(\ln\Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\right)+\ln\Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\right)\right) \end{array} \tag{A.13}$$

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

$$\begin{array}{rcl} \ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right) & \geq & \ln p_{\alpha}\left(y_{10},y_{20}\right) \\ & + & \sum\limits_{t=1}^{T}\left(1-y_{1t}\right)\left(1-y_{2t}\right)\ \left[\sigma_{\alpha 1}(y_{1t-1},0)+\sigma_{\alpha 2}(0,y_{2t-1})\right] \\ & + & \sum\limits_{t=1}^{T}\left(1-y_{1t}\right)y_{2t}\ \left[\sigma_{\alpha 1}(y_{1t-1},0)+\sigma_{\alpha 2}(0,y_{2t-1})+\alpha_2+\beta_{22}\ y_{2t-1}\right] \\ & + & \sum\limits_{t=1}^{T}y_{1t}\left(1-y_{2t}\right)\ \left[\sigma_{\alpha 1}(y_{1t-1},0)+\sigma_{\alpha 2}(0,y_{2t-1})+\alpha_1+\beta_{11}\ y_{1t-1}\right] \\ & + & \sum\limits_{t=1}^{T}y_{1t}y_{2t}\ \left[\sigma_{\alpha 1}(y_{1t-1},1)+\sigma_{\alpha 2}(1,y_{2t-1})+\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1+\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\right] \end{array} \tag{A.14}$$

Grouping terms, we have:

$$\begin{array}{rcl}\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \geq & \ln p_{\alpha}(y_{10}, y_{20}) \\ & + & \sum_{t=1}^{T}(1 - y_{1t}y_{2t}) \;\; [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\ & + & \sum_{t=1}^{T} y_{2t} \;\; [\alpha_2 + \beta_{22} \; y_{2t-1}] \\ & + & \sum_{t=1}^{T} y_{1t} \;\; [\alpha_1 + \beta_{11} \; y_{1t-1}] \\ & + & \sum_{t=1}^{T} y_{1t}y_{2t} \;\; [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \gamma_1 + \gamma_2] \end{array} \tag{A.15}$$

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 $\ln \mathbb{P}_{L\{E,N\}}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$:

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

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

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

(d) Lower Bound $\ln \mathbb{P}_{L\{S,W\}}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$. To obtain this lower bound, we use the bounds $L^{\{S,SE\}}(0,1|\mathbf{y}_{t-1};\alpha) \equiv [1 - \Lambda(\alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1)] \; \Lambda(\alpha_2 + \beta_{22} \; y_{2t-1} + \gamma_2)$ and $L^{\{W,NW\}}(1,0|\mathbf{y}_{t-1};\alpha) \equiv \Lambda(\alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1)$ $[1 - \Lambda(\alpha_2 + \beta_{22} \; y_{2t-1} + \gamma_2)]$ for the choice probabilities. Then,

$$\begin{array}{rcl}\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \geq & \ln p_{\alpha}(y_{10}, y_{20}) \\ & + & \sum_{t=1}^{T}(1 - y_{1t})(1 - y_{2t}) \;\; (\ln[1 - \Lambda(\alpha_1 + \beta_{11} \; y_{1t-1})] + \ln[1 - \Lambda(\alpha_2 + \beta_{22} \; y_{2t-1})]) \\ & + & \sum_{t=1}^{T}(1 - y_{1t}) \, y_{2t} \;\; (\ln[1 - \Lambda(\alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1)] + \ln \Lambda(\alpha_2 + \beta_{22} \; y_{2t-1} + \gamma_2)) \\ & + & \sum_{t=1}^{T} y_{1t}(1 - y_{2t}) \;\; (\ln \Lambda(\alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1) + \ln[1 - \Lambda(\alpha_2 + \beta_{22} \; y_{2t-1} + \gamma_2)]) \\ & + & \sum_{t=1}^{T} y_{1t}y_{2t} \;\; (\ln \Lambda(\alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1) + \ln \Lambda(\alpha_2 + \beta_{22} \; y_{2t-1} + \gamma_2)) \end{array} \tag{A.18}$$

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

$$\begin{array}{rcl}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \geq & \ln p_\alpha (y_{10}, y_{20}) \\
& + & \sum_{t=1}^{T} (1 - y_{1t}) (1 - y_{2t}) \;\; [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\
& + & \sum_{t=1}^{T} (1 - y_{1t}) y_{2t} \;\; [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_2 + \beta_{22} \; y_{2t-1} + \gamma_2] \\
& + & \sum_{t=1}^{T} y_{1t} (1 - y_{2t}) \;\; [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1] \\
& + & \sum_{t=1}^{T} y_{1t} y_{2t} \;\; [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1 + \alpha_2 + \beta_{22} \; y_{2t-1} + \gamma_2]
\end{array} \tag{A.19}$$

Grouping terms, we have:

$$\begin{array}{rcl}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \geq & \ln p_\alpha (y_{10}, y_{20}) \\
& + & \sum_{t=1}^{T} [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\
& + & \sum_{t=1}^{T} y_{2t} \;\; [\alpha_2 + \beta_{22} \; y_{2t-1} + \gamma_2] \\
& + & \sum_{t=1}^{T} y_{1t} \;\; [\alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1] \\
& + & \sum_{t=1}^{T} [y_{1t} + y_{2t} - y_{1t} y_{2t}] \;\; [\sigma_{\alpha 1}(y_{1t-1}, 1) - \sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(1, y_{2t-1}) - \sigma_{\alpha 2}(0, y_{2t-1})]
\end{array} \tag{A.20}$$

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 $\ln \mathbb{P}_{L\{S,W\}} (\widetilde{\mathbf{y}} \mid \alpha, \beta)$:

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

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

$$\begin{array}{rcl}
\ln \mathbb{P}_{L\{S,W\}} (\widetilde{\mathbf{y}} \mid \alpha, \beta) & = & \mathbf{s}^1 (\widetilde{\mathbf{y}})' \mathbf{g}_\alpha^1 \\
& + & \left[T_1^{(1)} + T_2^{(1)} - T^{(1,1)}, \; T_1^{(1)} + T_2^{(1)} - T^{(1,1)}, \; C_{11} + C_{21} - R_1^{(1,1)}, \; C_{12} + C_{22} - R_2^{(1,1)}\right] \; \mathbf{g}_\alpha^2 \\
& + & C_{11} \;\; \beta_{11} + C_{22} \;\; \beta_{22} + T_1^{(1)} \;\; \gamma_1 + T_2^{(1)} \;\; \gamma_2
\end{array} \tag{A.22}$$

(e) Upper Bound $\ln \mathbb{P}_U (\widetilde{\mathbf{y}} \mid \alpha, \beta)$. For the upper bounds, we use the bounds for the choice probabilities $U(0, 1|\mathbf{y}_{t-1}; \alpha) \equiv [1 - \Lambda (\alpha_1 + \beta_{11} \; y_{1t-1} + \gamma_1)] \; \Lambda (\alpha_2 + \beta_{22} \; y_{2t-1})$ and $U(1, 0|\mathbf{y}_{t-1}; \alpha) \equiv$

$\Lambda(\alpha_1 + \beta_{11}\ y_{1t-1})\ [1 - \Lambda(\alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2)]$. Then,

$$
\begin{array}{rcl}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \leq & \ln p_\alpha(y_{10}, y_{20}) \\
& + & \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \ \ (\ln[1-\Lambda(\alpha_1+\beta_{11}\ y_{1t-1})] + \ln[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})]) \\
& + & \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \ (\ln[1-\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1)] + \ln\Lambda(\alpha_2+\beta_{22}\ y_{2t-1})) \\
& + & \sum_{t=1}^{T} y_{1t}\,(1-y_{2t}) \ \ (\ln\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}) + \ln[1-\Lambda(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2)]) \\
& + & \sum_{t=1}^{T} y_{1t}y_{2t} \ \ (\ln\Lambda(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1) + \ln\Lambda(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2))
\end{array}
\tag{A.23}
$$

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

$$
\begin{array}{rcl}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \leq & \ln p_\alpha(y_{10}, y_{20}) \\
& + & \sum_{t=1}^{T} (1-y_{1t})(1-y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0,y_{2t-1})] \\
& + & \sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(0,y_{2t-1}) + \alpha_2 + \beta_{22}\ y_{2t-1}] \\
& + & \sum_{t=1}^{T} y_{1t}\,(1-y_{2t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_1 + \beta_{11}\ y_{1t-1}] \\
& + & \sum_{t=1}^{T} y_{1t}y_{2t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1,y_{2t-1}) + \alpha_1 + \beta_{11}\ y_{1t-1} + \gamma_1 + \alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2]
\end{array}
\tag{A.24}
$$

Grouping terms, we have:

$$
\begin{array}{rcl}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) & \leq & \ln p_\alpha(y_{10}, y_{20}) \\
& + & \sum_{t=1}^{T} (1-y_{2t})\ \ \sigma_{\alpha 1}(y_{1t-1},0) + (1-y_{1t})\ \ \sigma_{\alpha 2}(0,y_{2t-1}) \\
& + & \sum_{t=1}^{T} y_{2t}\ \sigma_{\alpha 1}(y_{1t-1},1) + y_{2t}\ \ [\alpha_2 + \beta_{22}\ y_{2t-1}] \\
& + & \sum_{t=1}^{T} y_{1t}\ \sigma_{\alpha 2}(1,y_{2t-1}) + y_{1t}\ \ [\alpha_1 + \beta_{11}\ y_{1t-1}] \\
& + & \sum_{t=1}^{T} y_{1t}y_{2t}\ \ [\gamma_1 + \gamma_2]
\end{array}
\tag{A.25}
$$

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 $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \alpha, \beta)$:

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

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

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

Proof of Proposition 4. Consider the myopic, complete information, Stackelberg model 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 choice histories. The choice probabilities conditional on the market level unobserved heterogeneity can be represented as

	$\{y_{11}, y_{21}\}$	$\{y_{12}, y_{22}\}$	$P(y)$
1	$\{0,0\}$	$\{0,0\}$	$(\frac{1}{1+A})^4$
2	$\{0,1\}$	$\{0,0\}$	$\frac{1}{1+A}\frac{A}{1+A}\frac{1}{1+AB_{12}}\frac{1}{1+AB_{22}}$
3	$\{1,0\}$	$\{0,0\}$	$\frac{A}{1+A}\frac{1}{1+A}\frac{1}{1+AB_{11}}\frac{1}{1+A}$
4	$\{1,1\}$	$\{0,0\}$	$\frac{A}{1+A}\frac{AC}{1+AC}\frac{1}{AB_{11}B_{12}}\frac{1}{AB_{22}}$
5	$\{0,0\}$	$\{0,1\}$	$(\frac{1}{1+A})^3\frac{A}{1+A}$
6	$\{0,1\}$	$\{0,1\}$	$\frac{1}{1+A}\frac{A}{1+A}\frac{1}{1+AB_{12}}\frac{AB_{22}}{1+AB_{22}}$
7	$\{1,0\}$	$\{0,1\}$	$\frac{A}{1+A}\frac{1}{1+AC}\frac{1}{1+AB_{11}}\frac{A}{1+A}$
8	$\{1,1\}$	$\{0,1\}$	$\frac{A}{1+A}\frac{AC}{1+AC}\frac{1}{1+AB_{11}B_{12}}\frac{AB_{22}}{1+AB_{22}}$
9	$\{0,0\}$	$\{1,0\}$	$\frac{1}{1+A}\frac{1}{1+A}\frac{A}{1+A}\frac{1}{1+AC}$
10	$\{0,1\}$	$\{1,0\}$	$\frac{1}{1+A}\frac{A}{1+A}\frac{AB_{12}}{1+AB_{12}}\frac{1}{1+ACB_{22}}$
11	$\{1,0\}$	$\{1,0\}$	$\frac{A}{1+A}\frac{1}{1+AC}\frac{AB_{11}}{1+AB_{11}}\frac{1}{1+A}$
12	$\{1,1\}$	$\{1,0\}$	$\frac{A}{1+A}\frac{AC}{1+AC}\frac{AB_{11}B_{12}}{1+AB_{11}B_{12}}\frac{1}{ACB_{22}}$
13	$\{0,0\}$	$\{1,1\}$	$\frac{1}{1+A}\frac{1}{1+A}\frac{A}{1+A}\frac{AC}{1+AC}$
14	$\{0,1\}$	$\{1,1\}$	$\frac{1}{1+A}\frac{A}{1+A}\frac{AB_{12}}{1+AB_{12}}\frac{ACB_{22}}{1+ACB_{22}}$
15	$\{1,0\}$	$\{1,1\}$	$\frac{A}{1+A}\frac{1}{1+AC}\frac{AB_{11}}{1+AB_{11}}\frac{AC}{1+AC}$
16	$\{1,1\}$	$\{1,1\}$	$\frac{A}{1+A}\frac{AC}{1+AC}\frac{AB_{11}B_{12}}{1+AB_{11}B_{12}}\frac{ACB_{22}}{1+ACB_{22}}$

where $A = \exp(\alpha)$, $B_{11} = \exp(\beta_{11})$, $B_{12} = \exp(\beta_{12})$, $B_{22} = \exp(\beta_{22})$ and $C = \exp(\gamma)$.

Denote $g(\alpha,\theta) = (1+A)^4(1+AC)^2(1+AB_{11})(1+AB_{12})(1+AB_{22})(1+AB_{11}B_{12})(1+ACB_{22})$. We can verify that the function $g(\alpha,\theta)$ is such that $P(y|\alpha,\theta)g(\alpha,\theta)$ becomes a polynomial function of A with its coefficient being polynomials of $(B_{11}, B_{12}, B_{22}, C)$. Also for any $\alpha \in \mathbb{R}$, $1/g(\alpha,\theta) \in (0,1]$, this implies that

$$P(y|\theta, y_0) = \int P(y|\alpha,\theta,y_0)dQ(\alpha|y_0) = \int P(y|\alpha,\theta,y_0)g(\alpha|\theta)d\bar{Q}(\alpha|\theta,y_0)$$

where Q is the distribution of the market level fixed effect and $d\bar{Q}(\alpha|\theta,y_0) = \frac{1}{g(\alpha,\theta)}dQ(\alpha|y_0)$. $\bar{Q}(\alpha|\theta,y_0)$ is a positive Borel measure on the support $[0,\infty)$. $\bar{Q}$ is not a probability measure, but it can be made into a probability measure by dividing $\int d\bar{Q}(\alpha|\theta,y_0)$ since $1/g(\alpha,\theta)$ is finite everywhere on the support of α, this integral exists and is finite. Some calculation shows that in particular, we can write

$$p(y|\theta) = G(\theta)m_A$$

where $G(\theta)$ is a 16×12 matrix with its elements only involving $\{B_{11}, B_{12}, B_{22}, C\}$ and m_A is a vector of length 12 and $m_A = \int \begin{pmatrix}1 & A & A^2 & \dots & A^{11}\end{pmatrix}' d\bar{Q}(\alpha|\theta)$, that is, the power moments of the measure $\bar{Q}$. Moment conditions for θ by finding a vector $v \in \mathbb{R}^{16}$, allowed to depends on θ such that

$$v'G(\theta) = 0$$

These collection of v is nothing but elements in the left null space of the matrix $G(\theta)$, hence we can just take all elements in a basis that spans the left null space of $G(\theta)$. In our specific case here with $T = 2$, the rank of $G(\theta)$ is 4, hence the dimension of the left null space of $G(\theta)$ is 4, therefore we expect to find 4 linearly independent moment conditions for θ. In particular, two of them take the form:

$$\begin{aligned}
&-B_{11}P_7 + P_{11} = 0\\
&-CP_3 - B_{11}CP_7 + CP_9 + P_{13} = 0
\end{aligned}$$

Clearly these two moment conditions identified γ_2 as well as β_{11}. We have two more moment conditions for the identification of β_{12} and β_{22}, which takes the form

$$\begin{aligned}
&\frac{B_{22}(C-1)}{B_{22}-C}(P_3 - P_2) - B_{22}P_4\\
&+ \frac{B_{12}C - B_{22}C + B_{22}^2 - B_{12}B_{22}C}{B_{22}(B_{22}-C)}P_6\\
&+ \frac{B_{11}B_{22}(C-1)}{B_{22}-C}P_7 - B_{11}B_{12}P_8 + P_{10} + P_{14} = 0
\end{aligned}$$

and

$$\frac{B_{11}B_{12}(C-1)^2}{C(B_{22}-C)(B_{11}B_{12}-B_{22}C)(B_{22}-1)}(P_3-P_2)-\frac{B_{11}B_{12}(C-1)}{B_{22}C^2-B_{11}B_{12}C}P_4$$

$$-\frac{B_{11}B_{12}(B_{12}C-B_{22}C-B_{22}^3C+B_{22}^2+B_{22}^2C^2-2B_{12}B_{22}C+B_{12}B_{22}^2C)}{B_{22}^2C(B_{22}-C)(B_{11}B_{12}-B_{22}C)(B_{22}-1)}P_6$$

$$+\frac{B_{11}^2B_{12}(C-1)^2}{C(B_{22}-C)(B_{11}B_{12}-B_{22}C)(B_{22}-1)}P_7$$

$$-\frac{B_{11}B_{12}}{B_{22}C}P_8+\frac{B_{11}(B_{22}C-1)(B_{12}-B_{22}C)}{B_{22}C(B_{11}B_{12}-B_{22}C)(B_{22}-1)}P_{10}+P_{12}=0$$

Proof of Lemma 4. Consider the equilibrium model defined by the system of equations:

$$\left\{\begin{array}{lcl} P_1 & = & \Lambda\left(a_1+\gamma_1\ P_2\right) \\ \\ P_2 & = & \Lambda\left(a_2+\gamma_2\ P_1\right) \end{array}\right. \tag{52}$$

An equilibrium in our model for a value of $(y_{1,t-1}, y_{2,t-1}) \in \{0,1\}^2$ is equivalent to an equilibrium of model (52) for a value of the intercepts (a_1, a_2) with $a_1 \in \{\alpha_1, \alpha_1+\beta_1\}$ and $a_2 \in \{\alpha_2, \alpha_2+\beta_2\}$. Let me use $P_1^*(a_1, a_2)$ and $P_2^*(a_1, a_2)$ to represent equilibrium probabilities in model (52) given value (a_1, a_2) for the intercept parameters. There is an obvious relationship between the equilibrium probabilities using the representation $\{P_1(y_{1,t-1}, y_{2,t-1}; \alpha_1, \alpha_2),\ P_2(y_{1,t-1}, y_{2,t-1}; \alpha_1, \alpha_2)\}$ and using $\{P_1^*(a_1, a_2), P_2^*(a_1, a_2)\}$. That is:

$$\left\{\begin{array}{lcl} \{P_1(0,0;\alpha_1,\alpha_2), P_2(0,0;\alpha_1,\alpha_2)\} & = & \{P_1^*(\alpha_1,\alpha_2), P_2^*(\alpha_1,\alpha_2)\} \\ \{P_1(1,0;\alpha_1,\alpha_2), P_2(1,0;\alpha_1,\alpha_2)\} & = & \{P_1^*(\alpha_1+\beta_1,\alpha_2), P_2^*(\alpha_1+\beta_1,\alpha_2)\} \\ \{P_1(0,1;\alpha_1,\alpha_2), P_2(0,1;\alpha_1,\alpha_2)\} & = & \{P_1^*(\alpha_1,\alpha_2+\beta_2), P_2^*(\alpha_1,\alpha_2+\beta_2)\} \end{array}\right. \tag{53}$$

Accordingly, we have that the following statement:

$$sign\{P_1(1,0;\alpha_1,\alpha_2)-P_1(0,0;\alpha_1,\alpha_2)\} \quad = \quad sign\{P_1^*(\alpha_1+\beta_1,\alpha_2)-P_1^*(\alpha_1,\alpha_2)\} \tag{54}$$

Similarly, we also have the following statements:

$$\begin{array}{lcl} sign\{P_1(0,1;\alpha_1,\alpha_2)-P_1(0,0;\alpha_1,\alpha_2)\} & = & sign\{P_1^*(\alpha_1,\alpha_2+\beta_2)-P_1^*(\alpha_1,\alpha_2)\} \\ sign\{P_2(1,0;\alpha_1,\alpha_2)-P_2(0,0;\alpha_1,\alpha_2)\} & = & sign\{P_2^*(\alpha_1+\beta_1,\alpha_2)-P_2^*(\alpha_1,\alpha_2)\} \\ sign\{P_2(0,1;\alpha_1,\alpha_2)-P_2(0,0;\alpha_1,\alpha_2)\} & = & sign\{P_2^*(\alpha_1,\alpha_2+\beta_2)-P_2^*(\alpha_1,\alpha_2)\} \end{array} \tag{55}$$

In model (52), we can consider that (a_1, a_2) can take any value in the Euclidean space $\mathbb{R}^2$. Given the continuous differentiability of function Λ, the following equivalence statements should be clear:

$$\begin{array}{ccc}
\left\{ \begin{array}{c} P_1^*(\alpha_1+\beta_1,\alpha_2) - P_1^*(\alpha_1,\alpha_2) \geq 0 \\ \text{for any } (\alpha_1,\alpha_2) \in \mathbb{R}^2 \end{array} \right\} & \Leftrightarrow & \left\{ \begin{array}{c} \dfrac{\partial P_1^*(a_1,a_2)}{\partial a_1} \; sign\{\beta_1\} \geq 0 \\ \text{for any } (a_1,a_2) \in \mathbb{R}^2 \end{array} \right\} \\
\left\{ \begin{array}{c} P_1^*(\alpha_1,\alpha_2+\beta_2) - P_1^*(\alpha_1,\alpha_2) \leq 0 \\ \text{for any } (\alpha_1,\alpha_2) \in \mathbb{R}^2 \end{array} \right\} & \Leftrightarrow & \left\{ \begin{array}{c} \dfrac{\partial P_1^*(a_1,a_2)}{\partial a_2} \; sign\{\beta_2\} \leq 0 \\ \text{for any } (a_1,a_2) \in \mathbb{R}^2 \end{array} \right\} \\
\left\{ \begin{array}{c} P_2^*(\alpha_1+\beta_1,\alpha_2) - P_2^*(\alpha_1,\alpha_2) \leq 0 \\ \text{for any } (\alpha_1,\alpha_2) \in \mathbb{R}^2 \end{array} \right\} & \Leftrightarrow & \left\{ \begin{array}{c} \dfrac{\partial P_2^*(a_1,a_2)}{\partial a_1} \; sign\{\beta_1\} \leq 0 \\ \text{for any } (a_1,a_2) \in \mathbb{R}^2 \end{array} \right\} \\
\left\{ \begin{array}{c} P_2^*(\alpha_1,\alpha_2+\beta_2) - P_2^*(\alpha_1,\alpha_2) \geq 0 \\ \text{for any } (\alpha_1,\alpha_2) \in \mathbb{R}^2 \end{array} \right\} & \Leftrightarrow & \left\{ \begin{array}{c} \dfrac{\partial P_2^*(a_1,a_2)}{\partial a_2} \; sign\{\beta_2\} \geq 0 \\ \text{for any } (a_1,a_2) \in \mathbb{R}^2 \end{array} \right\}
\end{array} \tag{56}$$

To obtain an expression for the derivative $\dfrac{\partial P_1^*(a_1,a_2)}{\partial a_1}$ in terms of primitives of the model, we differentiate the system of equations (52) with respect to a_1. We get:

$$\left\{ \begin{array}{lcl} \dfrac{\partial P_1^*}{\partial a_1} & = & \Lambda_1' \left[1 + \gamma_1 \; \dfrac{\partial P_2^*}{\partial a_1} \right] \\ \\ \dfrac{\partial P_2^*}{\partial a_1} & = & \Lambda_2' \left[\gamma_2 \; \dfrac{\partial P_1^*}{\partial a_1} \right] \end{array} \right. \tag{57}$$

where Λ_1' and Λ_2' are the derivatives of the logistic function evaluated at $(a_1 + \gamma_1 \; P_2^*)$ and $(a_2 + \gamma_2 \; P_2^*)$, respectively. Solving this system of linear equations in $(\partial P_1^*/\partial a_1, \, \partial P_2^*/\partial a_1)$ we obtain the following solution:

$$\left\{ \begin{array}{lcl} \dfrac{\partial P_1^*}{\partial a_1} & = & \dfrac{\Lambda_1'}{1 - \Lambda_1' \; \Lambda_2' \; \gamma_1 \gamma_2} \\ \\ \dfrac{\partial P_2^*}{\partial a_1} & = & \dfrac{\Lambda_1' \; \Lambda_2' \; \gamma_2}{1 - \Lambda_1' \; \Lambda_2' \; \gamma_1 \gamma_2} \end{array} \right. \tag{58}$$

Given that $\Lambda_1' > 0$ and $\Lambda_2' > 0$, we have that:

$$\left\{ \begin{array}{lcl} sign\left\{ \dfrac{\partial P_1^*}{\partial a_1} \right\} & = & sign\left\{ 1 - \Lambda_1' \; \Lambda_2' \; \gamma_1 \gamma_2 \right\} \\ \\ sign\left\{ \dfrac{\partial P_2^*}{\partial a_1} \right\} & = & sign\left\{ 1 - \Lambda_1' \; \Lambda_2' \; \gamma_1 \gamma_2 \right\} \; sign\left\{ \gamma_2 \right\} \end{array} \right. \tag{59}$$

Similarly, we can differentiate the system of equations (52) with respect to a_2, solve the system of linear equations in $(\partial P_1^*/\partial a_2, \, \partial P_2^*/\partial a_2)$, and obtain the following conditions.

$$\left\{ \begin{array}{lcl} sign\left\{ \dfrac{\partial P_2^*}{\partial a_2} \right\} & = & sign\left\{ 1 - \Lambda_1' \; \Lambda_2' \; \gamma_1 \gamma_2 \right\} \\ \\ sign\left\{ \dfrac{\partial P_1^*}{\partial a_2} \right\} & = & sign\left\{ 1 - \Lambda_1' \; \Lambda_2' \; \gamma_1 \gamma_2 \right\} \; sign\left\{ \gamma_1 \right\} \end{array} \right. \tag{60}$$

Plugging the results in (59) and (60) into the conditions in (56), we have that:

$$\begin{array}{ccc}
\left\{ \begin{array}{c} P_1^*(\alpha_1+\beta_1,\alpha_2) - P_1^*(\alpha_1,\alpha_2) \geq 0 \\ \text{for any } (\alpha_1,\alpha_2) \in \mathbb{R}^2 \end{array} \right\} & \Leftrightarrow & \left\{ \begin{array}{c} sign\left\{1-\Lambda_1' \ \Lambda_2' \ \gamma_1\gamma_2\right\} sign\{\beta_1\} \geq 0 \\ \text{for any } (a_1,a_2) \in \mathbb{R}^2 \end{array} \right\} \\
\left\{ \begin{array}{c} P_1^*(\alpha_1,\alpha_2+\beta_2) - P_1^*(\alpha_1,\alpha_2) \leq 0 \\ \text{for any } (\alpha_1,\alpha_2) \in \mathbb{R}^2 \end{array} \right\} & \Leftrightarrow & \left\{ \begin{array}{c} sign\left\{1-\Lambda_1' \ \Lambda_2' \ \gamma_1\gamma_2\right\} \ sign\left\{\gamma_1\right\} sign\{\beta_2\} \leq 0 \\ \text{for any } (a_1,a_2) \in \mathbb{R}^2 \end{array} \right\} \\
\left\{ \begin{array}{c} P_2^*(\alpha_1+\beta_1,\alpha_2) - P_2^*(\alpha_1,\alpha_2) \leq 0 \\ \text{for any } (\alpha_1,\alpha_2) \in \mathbb{R}^2 \end{array} \right\} & \Leftrightarrow & \left\{ \begin{array}{c} sign\left\{1-\Lambda_1' \ \Lambda_2' \ \gamma_1\gamma_2\right\} \ sign\left\{\gamma_2\right\} sign\{\beta_1\} \leq 0 \\ \text{for any } (a_1,a_2) \in \mathbb{R}^2 \end{array} \right\} \\
\left\{ \begin{array}{c} P_2^*(\alpha_1,\alpha_2+\beta_2) - P_2^*(\alpha_1,\alpha_2) \geq 0 \\ \text{for any } (\alpha_1,\alpha_2) \in \mathbb{R}^2 \end{array} \right\} & \Leftrightarrow & \left\{ \begin{array}{c} sign\left\{1-\Lambda_1' \ \Lambda_2' \ \gamma_1\gamma_2\right\} sign\{\beta_2\} \geq 0 \\ \text{for any } (a_1,a_2) \in \mathbb{R}^2 \end{array} \right\}
\end{array} \tag{61}$$

Suppose that $\beta_1 \geq 0$, $\beta_2 \geq 0$, $\gamma_1 \leq 0$, and $\gamma_2 \leq 0$. Then, a necessary and sufficient condition to obtain the inequalities in (61) is:

$$\Lambda_1' \ \Lambda_2' \ \gamma_1\gamma_2 \leq 1 \quad \text{for any } (a_1,a_2) \in \mathbb{R}^2 \tag{62}$$

Or equivalently,

$$\sup_{(a_1,a_2)\in\mathbb{R}^2} \left\{\Lambda_1'(a_1,a_2) \ \Lambda_2'(a_1,a_2)\right\} \ \gamma_1\gamma_2 \leq 1 \tag{63}$$

For the logistic function, we know that $\Lambda_1'(a_1,a_2) \leq 1/4$ and $\Lambda_2'(a_1,a_2) \leq 1/4$ for any $(a_1,a_2) \in \mathbb{R}^2$. We also know that there are values $(a_1,a_2) \in \mathbb{R}^2$ such that the upper bound $1/4$ is reach. Therefore, we have that $\sup_{(a_1,a_2)\in\mathbb{R}^2} \{\Lambda_1'(a_1,a_2) \ \Lambda_2'(a_1,a_2)\} = 1/4 * 1/4 = 1/16$. Therefore, if $\beta_1 \geq 0$, $\beta_2 \geq 0$, $\gamma_1 \leq 0$, and $\gamma_2 \leq 0$, a necessary and sufficient condition to obtain the inequalities in (61) is:

$$\gamma_1 \ \gamma_2 \leq 16 \tag{64}$$

This condition also implies equilibrium uniqueness. ■

Proof of Lemma 5. Under the conditions in Lemma 4, we can sign the average effects $\Delta_{y_1,y_2}^{(i)}$. More specifically,

$$\Delta_{1,0}^{(1)} \geq 0; \ \Delta_{1,0}^{(2)} \leq 0; \ \Delta_{0,1}^{(1)} \leq 0; \ \Delta_{0,1}^{(2)} \geq 0$$

We show the first two inequalities, and the last two can be shown similarly. Giben that $P_2(1,0,\alpha) - P_2(0,0,\alpha) \leq 0$ for all $\alpha \in \mathbb{R}^2$ and $P_2(1,0) - P_2(0,0) = \Lambda(\alpha_2 + \gamma_2 P_1(1,0)) - \Lambda(\alpha_2 + \gamma_2 P_1(0,0))$, we know that $\alpha_2 + \gamma_2 P_1(1,0) \leq \alpha_2 + \gamma_2 P_1(0,0)$. Applying Lemma 3, we get

$$\Delta_{1,0}^{(2)} \geq \frac{1}{4}\gamma_2 \Delta_{1,0}^{(1)}$$

Similarly, since $P_1(1,0,\alpha) - P_1(0,0,\alpha) \geq 0$ for all $\alpha \in \mathbb{R}^2$, and $P_1(1,0) - P_1(0,0) = \Lambda(\alpha_1 + \gamma_1 P_2(1,0) + \beta_{11}) - \Lambda(\alpha_1 + \gamma_1 P_2(0,0))$, we know that $\alpha_1 + \gamma_1 P_2(1,0) + \beta_{11} \geq \alpha_1 + \gamma_1 P_2(0,0)$.

Applying Lemma 3, we have

$$\Delta_{1,0}^{(1)} \leq \frac{1}{4}\{\gamma_1 \Delta_{1,0}^{(2)} + \beta_{11}\}$$

We can draw two inequalities in the second quadrant of $\mathbb{R}^2$. The intersection of the areas, illustrated in Figure 1, provides bound for $\Delta_{1,0}^{(1)}$ and $\Delta_{1,0}^{(2)}$.

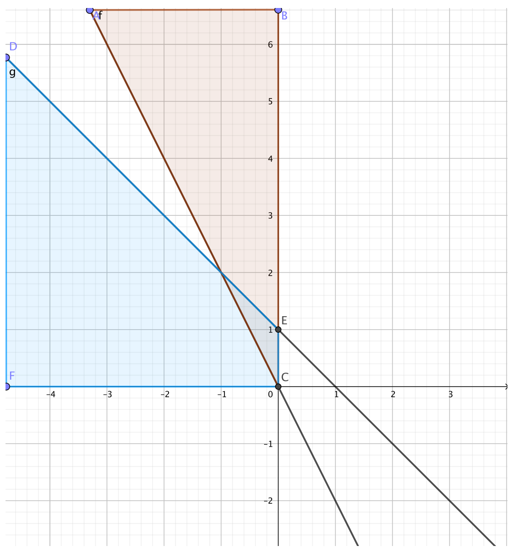


Figure 1: The x-axis is π_{10}^{b} and the y-axis is π_{10}^{a}. The red area portraits the first inequality and the blue area portraits the second inequality diaplayed in Lemma 5.

We then deduce that we have the following inequalities.

$$0 \geq \Delta_{1,0}^{(2)} \geq \frac{\gamma_2 \beta_{11}/16}{1-\gamma_1\gamma_2/16}$$
$$0 \leq \Delta_{1,0}^{(1)} \leq \frac{\beta_{11}/4}{1-\gamma_1\gamma_2/16}$$
$$0 \geq \Delta_{0,1}^{(1)} \geq \frac{\gamma_1 \beta_{22}/16}{1-\gamma_1\gamma_2/16}$$
$$0 \leq \Delta_{0,1}^{(2)} \leq \frac{\beta_{22}/4}{1-\gamma_1\gamma_2/16}$$

References

[1] Aguirregabiria, V., and J. Carro (2020): ”Identification of Average Marginal Effects in Fixed Effects Dynamic Discrete Choice Models,” manuscript, University of Toronto.

[2] Aguirregabiria, V., J. Gu, Y. Luo (2019): ”Sufficient Statistics for Unobserved Heterogeneity in Structural Dynamic Logit Models,” *Journal of Econometrics*, forthcoming.

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

[4] Aguirregabiria, V. and P. Mira (2007): “Sequential Estimation of Dynamic Discrete Games,” *Econometrica*, 75, 1-53.

[5] Andersen, E (1970): “Asymptotic Properties of Conditional Maximum Likelihood Estimators,” *Journal of the Royal Statistical Society, Series B*, 32, 283 301.

[6] Arcidiacono, P., and R. Miller (2011): “Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity,” *Econometrica*, 79(6), 1823-1867.

[7] Arellano, M., and S. Bonhomme (2011): ”Nonlinear Panel Data Analysis”, *Annual Review of Economics*, 3, 395-424.

[8] Arellano, M., and B. Honoré (2001): ”Panel Data Models: Some Recent Developments,” in J. J. Heckman and E. Leamer (eds.) *Handbook of Econometrics*, Volume 5, Chapter 53, North-Holland, 3229-3296.

[9] Bajari, P., L. Benkard and J. Levin (2006): ”Estimating dynamic models of imperfect competition,” *Econometrica*, 75, 1331-1370.

[10] Berry, S., and G. Compiani (2020): ”An Instrumental Variable Approach to Dynamic Models,” Becker Friedman Institute for Economics Working Paper No. 2020-106.

[11] Bonhomme, S. (2012): ”Functional Differencing”, *Econometrica*, 80, 1337 - 1385.

[12] 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.

[13] Bresnahan, T., and P. Reiss (1991): ”Empirical models of discrete games,” *Journal of Econometrics*, 48(1-2), 57-81.

[14] Brock, W., and S. Durlauf (2007): ”Identification of binary choice models with social interactions,” Journal of Econometrics, 140(1), 52-75.

[15] Brown, G. (1951) "Iterative Solutions of Games by Fictitious Play," in *Activity Analysis of Production and Allocation*, T. C. Koopmans (editor). New York: Wiley.

[16] Chamberlain, G. (1985): "Heterogeneity, Omitted Variable Bias, and Duration Dependence," in Longitudinal Analysis of Labor Market Data, edited by J. J. Heckman and B. Singer. Cambridge: Cambridge University Press.

[17] Chamberlain, G. (1993): "Feedback in Panel Data Models," unpublished manuscript, Department of Economics, Harvard University.

[18] Chamberlain, G. (2010): "Binary response models for panel data: Identification and information," *Econometrica*, 78(1), 159-168.

[19] Chen, L. Y. (2017): "Identification of Discrete Choice Dynamic Programming Models with Nonparametric Distribution of Unobservables," Econometric Theory, 33(3), 551-577.

[20] Ching, A. (2010): "A dynamic oligopoly structural model for the prescription drug market after ptent expiration," International Economic Review, 51(4), 1175-1207.

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

[22] Dobronyi, C., J. Gu, and K. Kim (2021): "Identification of Dynamic Panel Logit Models with Fixed Effects", manuscript.

[23] Dunne, T., S. Klimek, M. Roberts, and D. Xu (2013): "Entry, Exit and the Determinants of Market Structure," *RAND Journal of Economics*, 44(3), 462-487.

[24] Eckstein, Z., & Lifshitz, O. (2015): "Household interaction and the labor supply of married women," International Economic Review, 56(2), 427-455.

[25] Einav, L. (2010): "Not all rivals look alike: Estimating an equilibrium model of the release date timing game," Economic Inquiry, 48(2), 369-390.

[26] Ericson, R. and A. Pakes (1995): "Markov Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62, 53-82.

[27] Gallant, R., H. Hong, and A. Khwaja (2017): "The dynamic spillovers of entry: an application to the generic drug industry," Management Science 64 (3), 1189-1211.

[28] Gallant, 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.

[29] Goetler, R. and B. Gordon (2011): "Does AMD spur Intel to innovate more?" Journal of Political Economy, 119(6), 1141-1200.

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

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

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

[33] Heckman, J., and B. Singer (1984): "A method for minimizing the impact of distributional assumptions in economic models for duration data," *Econometrica*, 52, 271-320.

[34] Heckman, J. and S. Navarro (2007): "Dynamic Discrete Choice and Dynamic Treatment Effects," *Journal of Econometrics,* 136, 341-396.

[35] Holt, D. (1999): "An Empirical Model of Strategic Choice with an Application to Coordination Games," *Games and Economic Behavior,* 27(1), 86-105.

[36] Honoré, B. , and E. Kyriazidou (2000): "Panel data discrete choice models with lagged dependent variables," *Econometrica*, 68(4), 839-874.

[37] Honoré, B. , and E. Kyriazidou (2017): "Panel Vector Autoregressions with Binary Data," manuscript. Princeton University.

[38] Honoré, B. , and M. Weidner (2020): "Moment Conditions for Dynamic Panel Logit Models with Fixed Effects", manuscript, arXiv:2005.05942

[39] Hotz, J., and R.A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models," *Review of Economic Studies,* 60, 497-529.

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

[41] 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.

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

[43] Jeziorski, P. (2014): "Estimation of cost efficiencies from mergers: Application to US radio," The RAND Journal of Economics, 45(4), 816-846.

[44] Jofre-Bonet, M. and M. Pesendorfer (2003): "Estimation of a Dynamic Auction Game," *Econometrica,* 71, 1443–1489.

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

[46] Kasahara, H. and K. Shimotsu, 2007a, Nested Pseudo-likelihood Estimation and Bootstrap-based Inference for Structural Discrete Markov Decision Models. Manuscript. Department of Economics. The University of Western Ontario.

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

[48] Kano, K. (2013): "Menu Costs and Dynamic Duopoly," *International Journal of Industrial Organization*, 31(1), 102-118.

[49] Lancaster, T. (2000): "The incidental parameter problem since 1948," *Journal of Econometrics*, 95(2), 391-413.

[50] Lee, R., and A. Pakes (2009): "Multiple equilibria and selection by learning in an applied setting," *Economics Letters*, 104(1), 13-16.

[51] Neyman, J. and E. Scott (1948): "Consistent Estimates Based on Partially Consistent Observations," *Econometrica*, 16(1), 1-32.

[52] Pakes, A., M. Ostrovsky, and S. Berry (2007), "Simple Estimators for the Parameters of Discrete Dynamic Games, with Entry/Exit Examples", RAND Journal of Economics, 38(2), 373-399.

[53] Roberts, M. and J. Tybout (1997): "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 87(4), 545-564.

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

[55] Sieg, H., & Yoon, C. (2017): "Estimating dynamic games of electoral competition to evaluate term limits in us gubernatorial elections," *American Economic Review*, 107(7), 1824-57.

[56] Sovinsky, M., & Stern, S. (2016): "Dynamic modelling of long-term care decisions," Review of Economics of the Household, 14(2), 463-488.

[57] Suzuki, J. (2013): "Land use regulation as a barrier to entry: evidence from the Texas lodging industry," International Economic Review, 54(2), 495-523.

[58] 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.

[59] Takahashi, Y. (2015): "Estimating a war of attrition: The case of the us movie theater industry,". American Economic Review, 105(7), 2204-41.

[60] Toivanen, O. and M. Waterson (2011): "Retail Chain Expansion: The Early Years of McDonalds in Great Britain," CEPR Discussion Papers 8534, C.E.P.R. Discussion Papers.

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