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

February 18, 2019

(Preliminary and Incomplete)

Abstract

In the structural estimation of dynamic discrete choice games, the misspecification of unobserved heterogeneity can generate important biases in two types of important structural parameters: those that capture dynamic state dependence (e.g., costs of adjustment or switching), and those that capture strategic interactions between players (e.g., competition and peer effects). In this paper, we study the identification of these structural parameters in models with a large number of markets and a short time dimension, and where players' and market unobserved heterogeneity have a fixed effect structure, i.e., a fully nonparametric specification of the joint distribution of these unobservables and of the initial conditions of the endogenous state variables. We present identification results for both games of incomplete and of complete information, and with myopic and forward-looking players. Based on our identification results, we propose conditional maximum likelihood estimators of the structural parameters.

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`

1 Introduction

Dynamic games are useful tools for the analysis of economic and social phenomena characterized by intertemporal strategic interactions between a small number of 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.[2]

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 investment, entry/exit, adjustment, or switching (i.e., the *dynamic* part of the model); and the parameters that represent strategic interactions between players, such as competition, social, or peer effects (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 interactions between players (Bresnahan and Reiss, 1991). 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 is small and the number of markets M is large.[3] We consider the identification of these models when

[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.

[2] There is also a substantial literature on dynamic discrete choice models with social interactions where agents are not forward looking. See Brock and Durlauf (2007) and Blume, et al (2011).

[3] Most of our results apply to either a small or large number of players N. However, for concreteness and notational

there is time-invariant market, or player-market, unobserved heterogeneity 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 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. Its main limitation is that the distribution of the unobserved heterogeneity is not identified, and therefore there is not identification of marginal effects, or at least the identification of these effects requires additional restrictions.

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 (2018) 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 and extend the results in Honoré and Kyriazidou (2017). 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. Aguirregabiria and Mira (2007) deal with the initial conditions problem

simplicity, we focus on the case of small N.

by assuming that the initial conditions come from the market-type-specific ergodic distribution of the endogenous variables.[4] 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.

Robust estimation of some parameters of interest is the key advantage of the fixed effects approach. In the case of games, robustness extends to equilibrium selection - we can be completely agnostic about this in the sense that different markets may be selecting different equilibria. The equilibrium is part of the fixed effect - and like the fixed effect, we have to assume that it is time-invariant (within market) in our short panel.

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 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. In the paper we focus on binary choice games with two players. We show in the appendix how the results extend to multinomial choice games with N players.

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 game or *market* where a pair of agents play the game. 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, $\mathbb{E}\left[U_{imt}(y_{imt}, y_{jmt})\right]$, where $U_{imt}(y_i, y_j)$ is the utility of player i if she chooses action y_i and the

[4] This approach follows Heckman (1981).

other player chooses y_j in market m at period t. This utility function has the following standard structure:

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

$\widetilde{\alpha}_{im}(0)$ 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, and all the fixed effects $\{\alpha_m : m = 1, 2, ..., M\}$ are 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 $\{\widetilde{\varepsilon}_{imt}(j) : j \in \{0,1\}\}$ are *i.i.d.* over (i, m, t, j) with a extreme value type I distribution. We distinguish two types of models: *complete information games* where all the variables including $\widetilde{\varepsilon}$'s are common knowledge to all the firms; and *incomplete information games* where the $\widetilde{\varepsilon}_i$'s are private information of firm i but unknown to the other players in the game. The fixed effects α_m are assumed common knowledge in both cases.

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

$$\Delta U_{imt} = \alpha_{im} + \gamma_i\ y_{jmt} + \beta_{ii}\ y_{imt-1} + \beta_i\ y_{jmt-1} - \varepsilon_{imt} \quad (2)$$

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_{1mt} and ε_{2mt} are *i.i.d.* over (m, t) Logistic distributed.

A *Nash equilibrium* in the game of complete information is a solution to the system of best response conditions:

$$\begin{cases} y_{1mt} &= 1\left\{\alpha_{1m} + \gamma_1\ y_{2mt} + \beta_{11}\ y_{1mt-1} + \beta_{12}\ y_{2mt-1} - \varepsilon_{1mt} \geq 0\right\} \\ y_{2mt} &= 1\left\{\alpha_{2m} + \gamma_2\ y_{1mt} + \beta_{21}\ y_{1mt-1} + \beta_{22}\ y_{2mt-1} - \varepsilon_{2mt} \geq 0\right\} \end{cases} \quad (3)$$

where $1\{.\}$ is the indicator function. Similarly, a *Bayesian Nash equilibrium* in the incomplete information game is a solution to the system of equations:

$$\begin{cases} y_{1mt} &= 1\left\{\alpha_{1m} + \gamma_1\ P_{2m} + \beta_{11}\ y_{1mt-1} + \beta_{12}\ y_{2mt-1} - \varepsilon_{1mt} \geq 0\right\} \\ y_{2mt} &= 1\left\{\alpha_{2m} + \gamma_2\ P_{1m} + \beta_{21}\ y_{1mt-1} + \beta_{22}\ y_{2mt-1} - \varepsilon_{2mt} \geq 0\right\} \end{cases} \quad (4)$$

where $P_{jm} \equiv \mathbb{E}\left[y_{jmt} \mid \alpha_m,\, y_{1mt-1},\, y_{2mt-1}\right]$ in the equilibrium.

The researcher observe $\{y_{1mt}, y_{2mt}\}$ over M markets T periods where M is large and T is small. The model is a fixed effects model in the sense that 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 $\beta = (\gamma_i, \beta_{ij} : i, j \in \{1, 2\})'$. 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 Game with no contemporaneous effects

Consider the game described above under the condition that $\gamma_1 = \gamma_2 = 0$. Note that this game can bet interpreted either as complete or incomplete information. This model corresponds to the panel data binary choice VAR model studied by Honoré and Kyriazidou (2017).

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

The terms α_1 and α_2 are fixed effects. The parameters β_{11}, β_{12}, β_{21}, and β_{22} are the parameters of interest.

The data for one market consist of the history of choices between periods 1 and T, $\{y_{1t}, y_{2t} : t = 1, 2, .., T\}$, 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 refer to this vector as a *market history*. The model implies the following probability:

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

A key property of the logit model is that it facilitates a representation of the likelihood which is 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 $\sigma_{\alpha i}\left(y_1, y_2\right) \equiv -\ln[1 + \exp\left\{\alpha_i + \beta_{i1}\; y_1 + \beta_{i2}\; y_2\right\}]$ and $\sigma_\alpha\left(y_1, y_2\right) \equiv \sigma_{\alpha 1}\left(y_1, y_2\right) + \sigma_{\alpha 2}\left(y_1, y_2\right)$, and the statistics $T_i^{(1)} \equiv \sum_{t=1}^{T} y_{it}$, $T^{(j,k)} \equiv \sum_{t=1}^{T} 1\{(y_{1t}, y_{2t}) = (j, k)\}$, and $S_{ij} \equiv \sum_{t=1}^{T} y_{it}\; y_{jt-1}$ for

$i, j \in \{1, 2\}$. Then, the logarithm of the probability of a market history can be written as:

$$\begin{aligned} \ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) &= \ln p(y_{10}, y_{20} \mid \alpha) + \alpha_1 \, T_1^{(1)} + \alpha_2 \, T_2^{(1)} \\ &+ \textstyle\sum_{j=0}^{1} \sum_{k=0}^{1} \sigma_\alpha(j,k) \; [T^{(j,k)} + 1\{(y_{10}, y_{20}) = (j,k)\} - 1\{(y_{1T}, y_{2T}) = (j,k)\} \\ &+ \beta_{11} \, S_{11} + \beta_{12} \, S_{12} + \beta_{21} \, S_{21} + \beta_{22} \, S_{22} \end{aligned} \tag{7}$$

Or using a more compact representation:

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

where $\mathbf{u}(\widetilde{\mathbf{y}})$ is the vector of statistics $[T^{(j,k)}, \, 1\{(y_{10}, y_{20}) = (j,k)\}, \, 1\{(y_{1T}, y_{2T}) = (j,k)\} : (j,k) = \{0,1\}^2]'$; $\mathbf{g}_\alpha$ is the vector of functions of the incidental parameters $[\sigma_\alpha(j,k) + j \, \alpha_1 + k \, \alpha_2, \, \sigma_\alpha(j,k) + \ln p(j,k \mid \alpha), \, -\sigma_\alpha(j,k) : (j,k) = \{0,1\}^2]'$; $\mathbf{s}(\widetilde{\mathbf{y}})$ is the vector of statistics $[S_{11}, \, S_{12}, \, S_{21}, \, S_{22}]'$; and β is the vector parameters$(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22})'$.

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

PROPOSITION 1. For the myopic dynamic game with no contemporaneous effects described by equation (5): (A) The vector $\mathbf{u}(\widetilde{\mathbf{y}}) = [(y_{10}, y_{20}), \, (y_{1T}, y_{2T}), \, T^{(j,k)} : (j,k) \in \{0,1\}^2]'$ *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) Furthermore,* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{u}(\widetilde{\mathbf{y}}), \alpha, \beta) = \ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) - \ln \mathbb{P}(\mathbf{u}(\widetilde{\mathbf{y}}) \mid \alpha, \beta)$ *such that*

$$\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{u}(\widetilde{\mathbf{y}}), \beta) = \mathbf{s}(\widetilde{\mathbf{y}})' \beta - \ln \left(\textstyle\sum_{\widetilde{\mathbf{y}}' : \mathbf{u}(\widetilde{\mathbf{y}}') = \mathbf{u}(\widetilde{\mathbf{y}})} \exp\{\mathbf{s}(\widetilde{\mathbf{y}}')' \beta)\} \right) \tag{9}$$

with $\mathbf{s}(\widetilde{\mathbf{y}}) = [S_{11}, \, S_{12}, \, S_{21}, \, S_{22}]'$ *and* $\beta = [\beta_{11}, \, \beta_{12}, \, \beta_{21}, \, \beta_{22}]'$*. (C) For* $T \geq 3$*, there is a pair of market histories, say* A *and* B*, with* $\mathbf{u}(A) = \mathbf{u}(B)$ *and* $\mathbf{s}(A) - \mathbf{s}(B) = [1, \, 0, \, 0, \, 0]'$ *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}*.* ■

The following example provides specific histories that identify the parameters of interest.

EXAMPLE 1. 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 $\mathbf{u}$. 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 $T_A^{(j,k)} = T_B^{(j,k)}$. Therefore, $\mathbf{u}(A) = \mathbf{u}(B)$. Now, for $\boldsymbol{a} \neq \boldsymbol{b}$ we have that $\mathbf{s}(A) \neq \mathbf{s}(B)$ and the difference between the log-probabilities of these histories identifies parameters

of interest. Note that,

$$\begin{cases} S_{11}(A) - S_{11}(B) &= (a_1 - b_1)\ (y_{10} - y_{13}) \\ S_{12}(A) - S_{12}(B) &= (a_1 - b_1)y_{20} - (a_2 - b_2)y_{13} + a_2 b_1 - a_1 b_2 \\ S_{21}(A) - S_{21}(B) &= (a_2 - b_2)y_{10} - (a_1 - b_1)y_{23} + a_1 b_2 - a_2 b_1 \\ S_{22}(A) - S_{22}(B) &= (a_2 - b_2)\ (y_{20} - y_{23}) \end{cases} \tag{10}$$

Using the expressions in (10), we present in Table 1 examples of histories A and B and the corresponding parameter that is identified by $\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$.

Table 1
Myopic Game of Complete Information
with no contemporaneous effects
Examples of histories and identified parameters with T=3
$A = \{\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\}; \quad B = \{\boldsymbol{y}_0,\ \boldsymbol{b},\ \boldsymbol{a},\ \boldsymbol{y}_3\}$

	$\boldsymbol{y}_0$	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{y}_3$	$\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$
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}

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" (i.e., $(0,1,0,1)$) with the frequency of the history where she "stays" (i.e., $(0,0,1,1)$). In cases 3 and 4, we compare the probability of history $y_i = (0,0,0,1)$ for player i when the other player chooses alternative 1 at period $t = 2$ ($y_j = (0,0,1,0)$) and when this choice is $t = 1$ ($y_j = (0,1,0,0)$).

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 $\mathbf{S}_A$ and $\mathbf{S}_B$ be the matrices with dimension 256×4 such that each row contains the vector of statistics $\mathbf{s}(A)'$ and $\mathbf{s}(B)'$, 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 $(\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3)$. Matrix $\mathbf{S}_A - \mathbf{S}_B$ is full column rank and we can identify β as:

$$\beta = \left[(\mathbf{S}_A - \mathbf{S}_B)'(\mathbf{S}_A - \mathbf{S}_B)\right]^{-1}\left[(\mathbf{S}_A - \mathbf{S}_B)'(\ln \mathbf{P}_A - \ln \mathbf{P}_B)\right] \tag{11}$$

As we show in section 4, the sample counterpart of this expression provides a root-M consistent and asymptotically normal estimator of β. ■

2.3 Complete information and triangular contemporaneous effects

Now, consider that we relax the condition of no contemporaneous effects such that we restrict the parameter γ_1 to be zero (i.e., no contemporaneous effect of y_2 on y_1) but we allow γ_2 to be different to zero (i.e., there may be a contemporaneous effect of y_1 on y_2). Here we consider the restriction $\beta_{21} = 0$ (i.e., no effect of lagged y_1 on current y_2), but the model $\beta_{21} \neq 0$ has the same identification properties.[5]

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

This could be interpreted as a dynamic Stackelberg game where 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:

$$\begin{aligned} \ln \mathbb{P}\left(\widetilde{\mathbf{y}}|\alpha, \beta\right) &= \ln p\left(y_{10}, y_{20} \mid \alpha\right) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} + \sum_{t=1}^{T} \sigma_{\alpha 1}\left(y_{1t-1}, y_{2t-1}\right) + \sigma_{\alpha 2}\left(y_{1t}, y_{2t-1}\right) \\ &+ \beta_{11}\ S_{11} + \beta_{12}\ S_{12} + \beta_{22}\ S_{22} + \gamma_2\ T^{(1,1)} \end{aligned} \tag{13}$$

Comparing this expression of log-probability with equation (7) for the game without contemporaneous effects, we can see two important differences: the term $\gamma_2\ T^{(1,1)}$ instead of $\beta_{21}\ S_{21}$; and the term $\sum_{t=1}^{T} \sigma_{\alpha 2}\left(y_{1t}, y_{2t-1}\right)$ instead of $\sum_{t=1}^{T} \sigma_{\alpha 2}\left(y_{1t-1}, y_{2t-1}\right)$. Both differences have implications for 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 (i.e., with $\sigma_{\alpha 1}$). Similarly, log-probability depends on β_{12} *only* through the statistic S_{12}, and now this statistic is also associated with the incidental parameters through the term $S_{12}\ [\sigma_{\alpha 2}(1,1) - \sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,1) + \sigma_{\alpha 2}(0,0)]$.[6] This implies that the parameters γ_2 and β_{12} cannot be identified.

[5] In particular, with $\beta_{21} \neq 0$ we still have that β_{21} and β_{21} are not identified but β_{11} and β_{22} are identified.

[6] Note that $\sum_{t=1}^{T} \sigma_{\alpha 2}\left(y_{1t}, y_{2t-1}\right)$ 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 $T_1^{(1)}\ [\sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,0)] + \left[T_2^{(1)} + y_{20} - y_{2T}\right]\ [\sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0)] + S_{12}\ [\sigma_{\alpha 2}(1,1) - \sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,1) + \sigma_{\alpha 2}(0,0)] + T\sigma_{\alpha 2}(0,0)$.

Based on this equation, we can write $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta)$ as $\mathbf{u}(\widetilde{\mathbf{y}})' \, \mathbf{g}_\alpha + \mathbf{s}(\widetilde{\mathbf{y}})' \, \beta^*$ as in the previous model, but now the vectors of statistics $\mathbf{u}(\widetilde{\mathbf{y}})$ and $\mathbf{s}(\widetilde{\mathbf{y}})$, and the vector of identified parameters β^* is different. More specifically,

$$\begin{cases} \mathbf{u}(\widetilde{\mathbf{y}}) & = & \left[(y_{10}, y_{20}),\ (y_{1T}, y_{2T}),\ T^{(j,k)} : (j,k) \in \{0,1\}^2,\ S_{12}\right]' \\ \mathbf{s}(\widetilde{\mathbf{y}}) & = & [S_{11},\ S_{22}]' \\ \beta^* & = & [\beta_{11},\ \beta_{22}]' \end{cases} \tag{14}$$

PROPOSITION 2. For the myopic dynamic game with partial contemporaneous effects described by equation (12): (A) The vector $\mathbf{u}(\widetilde{\mathbf{y}}) = [(y_{10}, y_{20}),\ (y_{1T}, y_{2T}),\ T^{(j,k)} : (j,k) \in \{0,1\}^2,\ S_{12}]'$ *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) Furthermore,* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{u}(\widetilde{\mathbf{y}}), \beta) = \mathbf{s}(\widetilde{\mathbf{y}})' \beta^* - \ln(\sum_{\widetilde{\mathbf{y}}' : \mathbf{u}(\widetilde{\mathbf{y}}') = \mathbf{u}(\widetilde{\mathbf{y}})} \exp\{\mathbf{s}(\widetilde{\mathbf{y}}')' \beta^*)\})$ *with* $\mathbf{s}(\widetilde{\mathbf{y}}) = [S_{11}, S_{22}]'$ *and* $\beta^* = [\beta_{11},\ \beta_{22}]'$*. (C) For* $T \geq 3$*, there are histories* $\widetilde{\mathbf{y}}$ *such that* $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{u}(\widetilde{\mathbf{y}}), \beta)$ *identifies the vector of parameters of interest* β^**.* ■

EXAMPLE 2. Consider the same framework as in Example 1: $T = 3$ and the pair of histories $A = \{\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\}$ and $B = \{\boldsymbol{y}_0, \boldsymbol{b}, \boldsymbol{a}, \boldsymbol{y}_3\}$. In Example 1, we showed that these histories have the same value for the statistics $\boldsymbol{y}_0$, $\boldsymbol{y}_3$, and $T^{(j,k)}$. Now, in this model with a contemporaneous effect, the sufficient statistic includes S_{12}, so we need to impose additional conditions on histories A and B such that $S_{12}(A) = S_{12}(B)$. In the example in table 1, we have that $S_{12}(A) = S_{12}(B)$ 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 2. In contrast, for case 3 we have that $S_{12}(A) = 1$ and $S_{12}(B)$, such that this history does not identify β_{12}. For case 4, we have that $S_{12}(A) = S_{12}(B) = 0$, but for this case we also have that $\mathbf{s}(A) = \mathbf{s}(B)$ such that $\ln \mathbb{P}(A \mid \mathbf{u}, \beta) = \ln \mathbb{P}(B \mid \mathbf{u}, \beta)$ and this pair of histories do not identify any parameter. ■

Table 2
Myopic Game of Complete Information
with contemporaneous effect $y_1 \to y_2$
Examples of histories and identified parameters with T=3
$A = \{\boldsymbol{y}_0,\ \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{y}_3\}$; $B = \{\boldsymbol{y}_0,\ \boldsymbol{b},\ \boldsymbol{a},\ \boldsymbol{y}_3\}$ with $S_{12}(A) = S_{12}(B)$

	$\boldsymbol{y}_0$	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{y}_3$	$\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$
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}

2.4 Complete information and full contemporaneous effects

Consider the game with contemporaneous effects of y_2 on y_1 (i.e., $\gamma_1 \neq 0$) and and of y_1 on y_2 (i.e., $\gamma_2 \neq 0$). We also eliminate from the model the effects of lagged y_2 on current y_1 (i.e., $\beta_{12} = 0$) and of lagged y_1 on current y_2 (i.e., $\beta_{21} = 0$).

$$\begin{cases} y_{1t} &= 1\{\alpha_1 + \gamma_1\ y_{2t} + \beta_{11}\ y_{1t-1} - \varepsilon_{1t} \geq 0\} \\ y_{2t} &= 1\{\alpha_2 + \gamma_2\ y_{1t} + \beta_{22}\ y_{2t-1} - \varepsilon_{2t} \geq 0\} \end{cases} \tag{15}$$

Given the result for the model with triangular contemporaneous effect, it is clear that the parameters γ_1 and γ_2 cannot be identified in this model. Therefore, here we are interested in the identification of the "switching cost" parameters β_{11} and β_{22}.

This model is dynamic panel data version of the two-player binary choice games in Bresnahan and Reiss (1991) and Tamer (2003). Relative to the dynamic panel binary choice models considered above, this model has the additional complication of having multiple predictions, e.g. multiple equilibria. The model implies a partition of the space of the unobservables $(\varepsilon_1, \varepsilon_2)$ with the corresponding predictions of players' choice probabilities. 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 $\gamma_2 \leq 0$. In this case, 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,

$$\begin{cases} P(0,0 \mid \mathbf{y}_{t-1}; \alpha) &= [1 - \Lambda(\alpha_1 + \beta_{11}\ y_{1t-1})]\ \ [1 - \Lambda(\alpha_2 + \beta_{22}\ y_{2t-1})] \\ P(1,1 \mid \mathbf{y}_{t-1}; \alpha) &= \Lambda(\alpha_1 + \beta_{11}\ y_{1t-1} + \gamma_1)\ \ \Lambda(\alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2) \end{cases} \tag{16}$$

where $\Lambda(x)$ is the logistic function $\exp\{x\}/[1+\exp\{x\}]$. For the other probabilities, $P(0,1 \mid \mathbf{y}_{t-1};\alpha)$ and $P(1,0 \mid \mathbf{y}_{t-1};\alpha)$, the model does not provide unique predictions but it provides bounds.

$$\begin{cases} L(0,1|\mathbf{y}_{t-1};\alpha) & \leq P(0,1|\mathbf{y}_{t-1};\alpha) \leq \quad U(0,1|\mathbf{y}_{t-1};\alpha) \\ L(1,0|\mathbf{y}_{t-1};\alpha) & \leq P(1,0|\mathbf{y}_{t-1};\alpha) \leq \quad U(1,0|\mathbf{y}_{t-1};\alpha) \end{cases} \tag{17}$$

where the upper bounds are products of logit probabilities,

$$\begin{cases} U(0,1|\mathbf{y}_{t-1};\alpha) & \equiv \quad [1-\Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\right)]\ \Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}\right) \\ U(1,0|\mathbf{y}_{t-1};\alpha) & \equiv \quad \Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}\right)\ [1-\Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\right)] \end{cases} \tag{18}$$

and the lower bounds are $L(0,1|\mathbf{y}_{t-1};\alpha) = U(0,1|\mathbf{y}_{t-1};\alpha) - P^*(\mathbf{y}_{t-1})$ and $L(1,0|\mathbf{y}_{t-1};\alpha) = U(1,0|\mathbf{y}_{t-1};\alpha) - P^*(\mathbf{y}_{t-1})$ with

$$\begin{aligned} P^*(\mathbf{y}_{t-1}) \qquad & [\Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}\right)-\Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\right)] \\ & [\Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}\right)-\Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\right)] \end{aligned} \tag{19}$$

For this model, we consider two approaches for the identification of the parameters β_{11} and β_{22}. The first approach exploits 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. We show that this approach cannot provide point identification of any of the structural parameters, including β_{11} or β_{22}. The second approach exploits information also from histories with $y_{1t} \neq y_{2t}$ and takes into account the bounds of probabilities that model implies. This approach provides additional inequality constraints on the structural parameters.

2.4.1 Approach using only point predictions

Consider a market history $\widetilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0,1,..,T)$ with $y_{1t} = y_{2t}$ for any t such that we have well-defined probability functions.[7] For the moment, we maintain the player subindex i to facilitate the connection with the expressions derived above for the other models. For this type of histories, the model implies the following log-probability:

$$\begin{aligned} \ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right) \quad &= \quad \ln p\left(y_{10}, y_{20} \mid \alpha\right) + \alpha_1\ T_1^{(1)} + \alpha_2\ T_2^{(1)} + \textstyle\sum_{t=1}^{T} \sigma_{\alpha 1}\left(y_{1t-1}, y_{2t}\right) + \sigma_{\alpha 2}\left(y_{1t}, y_{2t-1}\right) \\ &+ \quad \beta_{11}\ S_{11} + \beta_{22}\ S_{22} + \left(\gamma_{12}+\gamma_{21}\right) T^{(1,1)} \end{aligned} \tag{20}$$

We now impose the restriction that the market history $\widetilde{\mathbf{y}}$ is such that $y_{1t} = y_{2t} \equiv y_t$ for any $t \geq 1$. This has the following implications: (a) $T^{(0,1)} = T^{(1,0)} = 0$; (b) $T_1^{(1)} = T_2^{(1)} = T^{(1,1)}$; (c) an

[7] In principle, we only need to impose the restriction $y_{1t} = y_{2t}$ for $t \geq 1$. However, the initial condition (y_{10}, y_{20}) is always part of the sufficient statistic for the incidental parameters. This implies that there is not any gain in identification by imposing or not the restriction $y_{10} = y_{20}$.

implication of (a) and (b) is that $T^{(0,0)} = T - T^{(1,1)}$; and (d) $S_{12} = S_{21} = S_{11} = S_{22} = \sum_{t=1}^{T} y_t y_{t-1}$. Define $\sigma_\alpha(y_t, y_{t-1}) = \sigma_{\alpha 1}(y_t, y_{t-1}) + \sigma_{\alpha 2}(y_t, y_{t-1})$. Taking into account these restrictions, we can re-write the log-probability function as follows:

$$\begin{aligned} \ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) &= \ln p(y_{10}, y_{20} \mid \alpha) + T\ \sigma_\alpha(0,0) + (y_0 - y_T)\ [\sigma_\alpha(0,1) - \sigma_\alpha(0,0)] \\ &+ T^{(1,1)}\ [\alpha_1 + \alpha_2 + \sigma_\alpha(0,1) + \sigma_\alpha(1,0) - 2\sigma_\alpha(0,0) + \gamma_{12} + \gamma_{21}] \\ &+ S_{11}\ [\sigma_\alpha(1,1) - \sigma_\alpha(0,1) - \sigma_\alpha(1,0) + \sigma_\alpha(0,0) + \beta_{11} + \beta_{22}] \end{aligned} \tag{21}$$

This expression shows already that neither the parameters γ_{12} and γ_{21} not the parameters β_{11} and β_{22} can be identified from these market histories. The parameters γ_{12} and γ_{21} are only related to the statistic $T^{(1,1)}$, and this statistic is also associated with the incidental parameters through the term $\alpha_1 + \alpha_2 + \sigma_\alpha(0,1) + \sigma_\alpha(1,0) - 2\sigma_\alpha(0,0)$. Similarly, the parameters β_{11} and β_{22} are only related to the statistic S_{11}, and this statistic is also associated with the incidental parameters through the term $\sigma_\alpha(1,1) - \sigma_\alpha(0,1) - \sigma_\alpha(1,0) + \sigma_\alpha(0,0)$.

PROPOSITION 3. For the myopic dynamic game with contemporaneous effects in equation (15), if we restrict the market histories $\widetilde{\mathbf{y}}$ to those with $y_{1t} = y_{2t}$ (such that the model has unique predictions on the probabilities), the structural parameters of the model cannot be identified. ■

2.4.2 Bounds approach

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

$$\begin{cases} \ln \mathbb{P}_L(\widetilde{\mathbf{y}} \mid \alpha, \beta) \equiv \ln p(y_{10}, y_{20} \mid \alpha) + \sum_{t=1}^{T} \ln L^*(\mathbf{y}_t | \mathbf{y}_{t-1}; \alpha, \beta) \\ \ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \alpha, \beta) \equiv \ln p(y_{10}, y_{20} \mid \alpha) + \sum_{t=1}^{T} \ln U(\mathbf{y}_t | \mathbf{y}_{t-1}; \alpha, \beta) \end{cases} \tag{22}$$

The upper bound probabilities $U(\mathbf{y}_t | \mathbf{y}_{t-1}; \alpha, \beta)$ are the ones presented in equation (16) for the unique predictions, and in equation (18) for the interval predictions. However, the lower probabilities $L^*(\mathbf{y}_t | \mathbf{y}_{t-1}; \alpha, \beta)$ that we use here are not the sharp bounds described above. This is because the sharp bounds do not have a logit structure. Without the logit structure it is not possible to derive sufficient statistics for α (Chamberlain, 2010). Instead, we can use non-sharp lower bounds that have a logit structure. There are two sets of lower bounds with a logit structure:

$$\begin{cases} L^*(0,1 | \mathbf{y}_{t-1}; \alpha) \equiv [1 - \Lambda(\alpha_1 + \beta_{11}\ y_{1t-1})]\ \Lambda(\alpha_2 + \beta_{22}\ y_{2t-1}) \\ L^*(1,0 | \mathbf{y}_{t-1}; \alpha) \equiv \Lambda(\alpha_1 + \beta_{11}\ y_{1t-1})\ [1 - \Lambda(\alpha_2 + \beta_{22}\ y_{2t-1})] \end{cases} \tag{23}$$

and

$$\begin{cases} L^{**}(0,1|\mathbf{y}_{t-1};\alpha) & \equiv & [1-\Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\right)]\ \Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\right) \\ L^{**}(1,0|\mathbf{y}_{t-1};\alpha) & \equiv & \Lambda\left(\alpha_1+\beta_{11}\ y_{1t-1}+\gamma_1\right)\ [1-\Lambda\left(\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2\right)] \end{cases} \tag{24}$$

Taking into account these bounds for the choice probabilities, we show in the Appendix the following result for the bounds on the log-probability of a market history.

PROPOSITION 4. For the myopic dynamic game with contemporaneous effects in equation (15), the log-probability of a market history has lower bounds $\ln \mathbb{P}_{L(1)}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right)$ *and* $\ln \mathbb{P}_{L(2)}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right)$ *and upper bound* $\ln \mathbb{P}_{U}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right)$ *which have the following expressions:*

$$\begin{aligned} \ln \mathbb{P}_{L(1)}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right) &= \mathbf{u}\left(\widetilde{\mathbf{y}}\right)'\mathbf{g}_\alpha + T_2^{(1)}\ \gamma_2 + S_{11}\ \beta_{11} + S_{22}\ \beta_{22} + T^{(1,1)}\ \gamma_1 \\ \ln \mathbb{P}_{L(2)}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right) &= \mathbf{u}\left(\widetilde{\mathbf{y}}\right)'\mathbf{g}_\alpha + T_1^{(1)}\ \gamma_1 + S_{11}\ \beta_{11} + S_{22}\ \beta_{22} + T^{(1,1)}\ \gamma_2 \\ \ln \mathbb{P}_{U}\left(\widetilde{\mathbf{y}} \mid \alpha,\beta\right) &= \mathbf{u}\left(\widetilde{\mathbf{y}}\right)'\mathbf{g}_\alpha + S_{11}\ \beta_{11} + S_{22}\ \beta_{22} + T^{(1,1)}\ [\gamma_1+\gamma_2] \end{aligned} \tag{25}$$

with $\mathbf{u}\left(\widetilde{\mathbf{y}}\right) \equiv [y_{10},\ y_{20},\ y_{1T},\ y_{2T},\ T_1^{(1)},\ T_2^{(1)}]'$. ■

Proposition 4 implies that $\mathbf{u}\left(\widetilde{\mathbf{y}}\right) \equiv [y_{10},\ y_{20},\ y_{1T},\ y_{2T},\ T_1^{(1)},\ T_2^{(1)}]'$ is a sufficient statistic for the lower bound and the upper bound probabilities of a market history. Let A and B be two histories such that $\mathbf{u}\left(A\right) = \mathbf{u}\left(B\right)$. This implies that: $\ln \mathbb{P}_{L(1)}\left(A \mid \alpha,\beta\right) - \ln \mathbb{P}_{U}\left(B \mid \alpha,\beta\right) = [S_{11}(A) - S_{11}(B)]\ \beta_{11} + [S_{22}(A) - S_{22}(B)]\ \beta_{22} + [T^{(1,1)}(A) - T^{(1,1)}(B)]\ \gamma_1 - T^{(1,1)}(B)\ \gamma_2$. Similarly, we have that $\ln \mathbb{P}_{U}\left(A \mid \alpha,\beta\right) - \ln \mathbb{P}_{L(1)}\left(B \mid \alpha,\beta\right) = [S_{11}(A) - S_{11}(B)]\ \beta_{11} + [S_{22}(A) - S_{22}(B)]\ \beta_{22} + [T^{(1,1)}(A) - T^{(1,1)}(B)]\ \gamma_1 + T^{(1,1)}(A)\ \gamma_2$. Therefore, we have the following bounds on the log-probability difference $\ln \mathbb{P}\left(A\right) - \ln \mathbb{P}\left(B\right)$:

$$\begin{aligned} & [S_{11}(A) - S_{11}(B)]\beta_{11} + [S_{22}(A) - S_{22}(B)]\beta_{22} + [T^{(1,1)}(A) - T^{(1,1)}(B)]\gamma_1 - T^{(1,1)}(B)\gamma_2 \\ & \leq \ln \mathbb{P}\left(A\right) - \ln \mathbb{P}\left(B\right) \leq \\ & [S_{11}(A) - S_{11}(B)]\beta_{11} + [S_{22}(A) - S_{22}(B)]\beta_{22} + [T^{(1,1)}(A) - T^{(1,1)}(B)]\gamma_1 - T^{(1,1)}(A)\gamma_2 \end{aligned} \tag{26}$$

Proposition 5 establishes that there are pairs of histories such that these bounds are informative an provide an informative identified set on the structural parameters γ_1, γ_2, β_{11}, and β_{22}.

PROPOSITION 5. For the myopic dynamic game with contemporaneous effects in equation (15), there are pairs of histories, A and B, such that the bounds for $\ln \mathbb{P}\left(A\right) - \ln \mathbb{P}\left(B\right)$ *in equation (26) depend on the structural parameters and provide an informative identified set for the vector of structural parameters* $(\gamma_1,\ \gamma_2,\ \beta_{11},\ \beta_{22})$. ■

EXAMPLE 3. Consider data with $T = 3$ and the pair of market histories. $A = [(0,0), (0,0), (1,1), (1,1)]$ and $B = [(0,0), (0,1), (1,0), (1,1)]$. These two histories have the same initial condition $(y_{10}, y_{20}) = (0,0)$, the same final choices $(y_{1T}, y_{2T}) = (1,1)$, and the same value for $T_1^{(1)} = 2$, and $T_2^{(1)} = 2$, such that $\mathbf{u}(A) = \mathbf{u}(B)$. The two histories have different values for the statistics S_{11}, S_{22}, and $T^{(1,1)}$. In particular, $S_{11}(A) = 1$, $S_{11}(B) = 1$, $S_{22}(A) = 1$, $S_{22}(B) = 0$, $T^{(1,1)}(A) = 2$, and $T^{(1,1)}(B) = 1$. Therefore,

$$\beta_{22} + \gamma_1 - \gamma_2 \leq \ln \mathbb{P}(A) - \ln \mathbb{P}(B) \leq \beta_{22} + \gamma_1 - 2\gamma_2. \qquad \blacksquare$$

3 Identification of dynamic games with forward-looking players

3.1 Framework

We first present a version of the model with two players, that we index by $i, j \in \{1, 2\}$. Subindex m represents a realization of the game or *market* where a pair of agents play the game. Time is discrete and indexed by t that belongs to $\{1, 2, ..., T\}$. The time horizon T is infinite. Every period t, the two players choose simultaneously their actions in market m, y_{1mt} and y_{2mt}, from the set of choice alternatives $\mathcal{Y} = \{0, 1, ..., J\}$ 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 for agent i if she chooses action y_i and the other player chooses y_j in market m at period t. This one-period payoff has the following structure:

$$U_{imt} = \alpha_{im}(y_{imt}) + \beta_i(y_{imt}, \mathbf{y}_{mt-1}) + \varepsilon_{imt}(y_{it}). \tag{27}$$

$\alpha_{im}(0)$, ..., $\alpha_{im}(J)$ are market and player characteristics that are unobservable to the researcher. The vector $\alpha_m \equiv \{\alpha_{1m}(j), \alpha_{2m}(j) : j \in \mathcal{Y}\}$ represents the fixed effects for market m, and all the fixed effects $\{\alpha_m : m = 1, 2, ..., M\}$ are the *incidental parameters* of the model. Vector $\mathbf{y}_{mt-1}$ contains the players' endogenous state variables (y_{1mt-1}, y_{2mt-1}). Functions β_i are the structural parameters of the model and the capture state dependence: dynamic strategic interactions between the two players, e.g., competition, social, or peer effects;[8] state dependence with respect to the action of the own player, e.g., adjustment costs, switching costs. We could extend this specification to

[8] Following the dynamic game framework in Ericson and Pakes (1995), we assume that there is *time-to-build* in the dynamics of the strategic interactions between the players. Players make their respective "investment" decisions at period t and pay the corresponding investment cost at that period, but it is not until next period $t+1$ that they obtain the returns from these investments. This implies that the payoff function of player i at period t depends on her action at period t (i.e., that affects her investment costs), and on all the players' actions at $t-1$ (i.e., that determine revenues and variable costs at period t).

include exogenous state variables in these functions. The unobservable variables $\{\varepsilon_{imt}(j) : j \in \mathcal{Y}\}$ are *i.i.d.* over (i, m, t, j) with a extreme value type I distribution. 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 firm i but unknown to the other players in the game. Note that the fixed effects are assumed common knowledge in both cases.

3.1.1 Dynamic games of incomplete information

Following Ericson and Pakes (1995) and the empirical literature on estimation of dynamic discrete choice games, we assume that players' decisions come from a Markov Perfect Equilibrium (MPE). This implies that players' strategies depend only of payoff relevant state variables. At any period t, the action of player i, y_{imt}, is a function of the variables known by this player that affects her payoff, or the payoff of other players, at period t. In this model, this means that $y_{imt} = \sigma_{im}(\varepsilon_{imt}, \mathbf{y}_{mt-1})$, where $\sigma_i(.)$ represents the strategy function of player i.

A player, say i, takes the strategy functions of the other players as given and chooses the strategy function that maximizes her own intertemporal value. This best response function is the solution to a single-agent dynamic programming problem. Let $V_{im}^{\sigma_{-i}}(\mathbf{y}_{mt-1})$ be the value function for player i (integrated over the i.i.d. unobservables ε_i) given that the other player behaves according to the strategy function σ_{-i}. This value function is the unique solution to the Bellman equation:

$$V_{im}^{\sigma_{-i}}(\mathbf{y}_{mt-1}) = \ln\left(\sum_{y_i \in \mathcal{Y}} \exp\left\{\alpha_{im}(y_{it}) + \beta_i(y_i, \mathbf{y}_{mt-1}) + v_{im}(y_i, \mathbf{y}_{mt-1})\right\}\right) \tag{28}$$

where $v_{im}(y_i, \mathbf{y}_{mt-1})$ is a continuation value that is equal to the next period value $V_{im}^{\sigma_{-i}}(\mathbf{y}_{mt})$: (a) conditional on the choice of player i; (b) integrated over the distribution of the private information of the other player (i.e., over $f_\varepsilon(\varepsilon_{-imt})$) and therefore, over the the corresponding choices of other players. That is,

$$v_{im}(y_i, \mathbf{y}_{mt-1}) \equiv \delta_{im}\left[\int V_{im}^{\sigma_{-i}}(y_i, \sigma_{-i}(\varepsilon_{-i}, \mathbf{y}_{mt-1}))\ f_\varepsilon(\varepsilon_{-i})\ d\varepsilon_{-i}\right] \tag{29}$$

Player i chooses y_{imt} to maximize its expected and discounted intertemporal payoff. The optimal choice at period t can be represented as:

$$y_{imt} = \arg\max_{y_i \in \mathcal{Y}} \left(\alpha_{im}(y_i) + \beta_i(y_i, \mathbf{y}_{mt-1}) + \varepsilon_{imt}(y_i) + v_{im}(y_i, \mathbf{y}_{mt-1})\right) \tag{30}$$

And the CCP function for player i is:

$$P_i(y_i \mid \alpha_m, \mathbf{y}_{mt-1}) = \frac{\exp\left\{\ (\alpha_{im}(y_i) + \beta_i(y_i, \mathbf{y}_{mt-1}) + v_{im}(y_i, \mathbf{y}_{mt-1}))\ \right\}}{\sum_{j \in \mathcal{Y}} \exp\left\{\ \alpha_{im}(j) + \beta_i(j, \mathbf{y}_{mt-1}) + \varepsilon_{imt}(j) + v_{im}(j, \mathbf{y}_{mt-1})\ \right\}} \tag{31}$$

Note that the continuation value $v_{im}(y_i, \mathbf{y}_{mt-1})$ depends implicitly on the strategy function of the other player. Therefore, equation (31) establishes implicitly a relationship between the CCP functions of the two players, i.e., an equilibrium mapping in the space of CCPs.

3.1.2 Dynamic games of complete information

We assume that players' decisions come from a Markov Perfect Equilibrium (MPE). This implies that players' strategies depend only of payoff relevant state variables. At any period t, the action of player i, y_{imt}, is a function of the variables known by this player that affect her payoff or the payoff of other players at period t. The vector of state variables is $\mathbf{s}_{mt} \equiv (\varepsilon_{mt}, \mathbf{y}_{mt-1})$, where ε_{mt} is the vector with the ε variables of all the players. Therefore, $y_{imt} = \sigma_i(\mathbf{s}_{mt})$, where $\sigma_i(.)$ represents the strategy function of player i.

The equilibrium concept of MPE is based on the Nash conjecture or assumption about players' best responses. A player, say i, takes the strategy functions of the other players as given and then chooses a strategy function that maximizes her own intertemporal value. In the context of a dynamic game, this best response function is the solution to a single-agent dynamic programming problem. Let $V_{im}^{\sigma_{-i}}(\mathbf{s}_{mt})$ be the value function for player i given that the other player behaves according to the strategy function σ_{-i}. This value function is the unique solution to the Bellman equation:

$$V_{im}^{\sigma_{-i}}(\mathbf{s}_{mt}) = \max_{j \in \mathcal{Y}} \ \left\{ U_{imt}(j, \sigma_{-i}[\mathbf{s}_{mt}]) + \delta_{im} \ \mathbb{E}\left[V_{im}^{\sigma_{-i}}(\mathbf{s}_{mt+1}) \ \mid \mathbf{s}_{mt}, \ y_{imt} = j, \ y_{-imt} = \sigma_{-i}(\mathbf{s}_{mt}) \right] \right\} \tag{32}$$

The second term in the expression in brackets $\{.\}$ can be described as the *continuation value function*. It is the expected and discounted value of next period value, $V_{im}^{\sigma_{-i}}(\mathbf{s}_{mt+1})$, under the following conditions: (i) player i chooses alternative j at period t; and (ii) the other player behaves, at period t and at any future period, according to the decision rule σ_{-i}. The expectation is taken with respect to the next period's realization of unobservable payoff components ε_m. We use $v_{im}(y_{it}, \sigma_{-i}(\mathbf{s}_{mt}))$ to represent this function in a more compact way. We will also represent this continuation value function using the slightly more compact notation $v_{im}(y_{it}, y_{-imt})$. Note that the continuation value depends on ε_{mt} only through the action of the other player, $\sigma_{-i}(\mathbf{s}_{mt})$. This is because ε_{mt} is independently distributed over time and it does not enter explicitly in the transition probability function of the other state variables such that the expected value of future payoffs does not depend on current ε_{mt}.

A MPE can be described as a pair of strategy functions, σ_1 and σ_2, such that each strategy

function is a best response to the other. More formally, for $i \in \{1, 2\}$, we have that:

$$\sigma_i(\mathbf{s}_{mt}) \quad = \quad \arg\max_{j \in \mathcal{Y}} \{U_{imt}(j, \sigma_{-i}[\mathbf{s}_{mt}]) + v_{im}(j, \sigma_{-i}(\mathbf{s}_{mt}))\} \tag{33}$$

Therefore, given a state $\mathbf{s}_{mt}$, a vector of choices (y_{1mt}, y_{2mt}) is consistent with a MPE if and only if $(y_{1mt}, y_{2mt}) = (\sigma_1(\mathbf{s}_{mt}), \sigma_2(\mathbf{s}_{mt}))$, and this condition is equivalent to:

$$\begin{aligned} y_{1mt} &= \arg\max_{j \in \mathcal{Y}} \{\alpha_{1m}(j) + \beta_1(j, \mathbf{y}_{mt-1}) + \varepsilon_{1mt}(j) + v_{1m}(j, y_{2mt})\} \\ y_{2mt} &= \arg\max_{j \in \mathcal{Y}} \{\alpha_{2m}(j) + \beta_2(j, \mathbf{y}_{mt-1}) + \varepsilon_{2mt}(j) + v_{2m}(j, y_{1mt})\} \end{aligned} \tag{34}$$

In general, this dynamic game has multiple MPE. It is important to distinguish two implications of multiple equilibria in this model. First, multiplicity of equilibria implies that there are multiple continuation value functions, v_{1m} and v_{2m}, that are consistent with the same primitives of the model. This multiple equilibria issue is similar to the one in dynamic discrete games of incomplete information (see Aguirregabiria and Mira, 2007). Second, for fixed continuation value function v_{1m} and v_{2m} and for a given value of the state variables $\mathbf{s}_{mt}$, there may be (and in general there are) multiple vectors of (y_{1mt}, y_{2mt}) that satisfy the equilibrium conditions in (34). This is the same type problem with multiple equilibria that appears in static games of complete information (see Bresnahan and Reiss, 1990, Tamer, 2003).

3.2 Identification of dynamic games of complete information

In empirical applications of dynamic games, the researcher has panel data of M markets over T periods of time, $\{y_{imt} : m = 1, 2, ..., M;\ i = 1, 2, ..., N;\ t = 0, 1, ..., T\}$. The number of players, N, and the number of time periods, T, are small, and the number of markets is large. Our identification results and the asymptotic properties of the proposed estimators assume that M goes to infinity and N and T are small and fixed. The incidental parameters problem comes from the "parameters" α_m that grow in number as the sample size M increases.

Let $\mathbf{y}_{mt} = (y_{m0t}, y_{m1t}, ..., y_{mNt})$ be the vector with the choices of all the players in market m at period t. And let $\widetilde{\mathbf{y}}_m = \{\mathbf{y}_{m1}, \mathbf{y}_{m2}, ..., \mathbf{y}_{mT}\}$ be the history of players' choices in market m from period 1 to T. We look for an statistic u_m such that it is observable to the researcher and the probability $\mathbb{P}(\widetilde{\mathbf{y}}_m \mid \alpha_m,\ u_m)$ satisfies two conditions: (i) u_m is a sufficient statistic for α_m is this probability, i.e., $\mathbb{P}(\widetilde{\mathbf{y}}_m \mid \alpha_m,\ u_m) = \mathbb{P}(\widetilde{\mathbf{y}}_m \mid u_m)$; and (ii) the probability $\mathbb{P}(\widetilde{\mathbf{y}}_m \mid u_m)$ still depends on the structural parameters of interest.

We start considering the two-player binary choice model of complete information where the only endogenous state variables are the lagged choices of the players, $y_{1,t-1}$ and $y_{2,t-1}$. The best

response decisions are:

$$\begin{aligned} y_{1t} &= 1\left\{\alpha_1 + \beta_1\ y_{1,t-1} + \gamma_1\ y_{2t} + \widetilde{v}_{1\alpha}\left(y_{2t}\right) - \varepsilon_{1t} \geq 0\right\} \\ y_{2t} &= 1\left\{\alpha_2 + \beta_2\ y_{2,t-1} + \gamma_2\ y_{1t} + \widetilde{v}_{2\alpha}\left(y_{1t}\right) - \varepsilon_{2t} \geq 0\right\} \end{aligned} \tag{35}$$

where $\widetilde{v}_{i\alpha}\left(y_{jt}\right) \equiv v_{i\alpha}\left(1, y_{jt}\right) - v_{i\alpha}\left(0, y_{jt}\right)$ and we use the subindex α in the continuation value function to indicate that it depends on the incidental parameters α. It is convenient to represent this model as follows:

$$y_{it} = 1\left\{\widetilde{\alpha}_i + \beta_i\ y_{i,t-1} + \widetilde{\gamma}_{i\alpha}\ y_{jt} - \varepsilon_{it} \geq 0\right\} \tag{36}$$

where $\widetilde{\alpha}_i \equiv \alpha_i + \widetilde{v}_{i\alpha}\left(0\right)$ and $\widetilde{\gamma}_{i\alpha} \equiv \gamma_i + \widetilde{v}_{i\alpha}\left(1\right) - \widetilde{v}_{i\alpha}\left(0\right)$. Given this representation, it should be clear that it is not possible to identify γ_i because it always appears together with the incidental parameters $\widetilde{v}_{i\alpha}\left(1\right) - \widetilde{v}_{i\alpha}\left(0\right)$. Therefore, our purpose here is to study the identification of the switching cost parameters β_1 and β_2.

Proposition 3 provides above provides a negative identification result for the identification of β_1 and β_2. If we consider only market histories with $y_{1t} = y_{2t}$ (such that the model provides unique predictions on the probabilities), the parameters β_1 and β_2 are not identified. The identification of these parameters requires considering histories with $y_{1t} \neq y_{2t}$ and therefore dealing with the interval predictions of the model.

3.3 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:

$$\begin{cases} y_{1t} = 1\left\{\alpha_1 + \beta_{11}\ y_{1,t-1} + \beta_{12}\ y_{2,t-1} + \widetilde{v}_{1\alpha}\left(y_{1,t-1}, y_{2,t-1}\right) - \varepsilon_{1t} \geq 0\right\} \\ y_{2t} = 1\left\{\alpha_2 + \beta_{21}\ y_{1,t-1} + \beta_{22}\ y_{2,t-1} + \widetilde{v}_{2\alpha}\left(y_{1,t-1}, y_{2,t-1}\right) - \varepsilon_{2t} \geq 0\right\} \end{cases} \tag{37}$$

where $\widetilde{v}_{i\alpha}\left(y_{1,t-1}, y_{2,t-1}\right) \equiv v_{i\alpha}\left(1; y_{1,t-1}, y_{2,t-1}\right) - v_{i\alpha}\left(0; y_{1,t-1}, y_{2,t-1}\right)$. The continuation value function $\widetilde{v}_{i\alpha}\left(y_{1,t-1}, y_{2,t-1}\right)$ depends on $\left(y_{1,t-1}, y_{2,t-1}\right)$ 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 $\left(y_{1,t-1}, y_{2,t-1}\right)$ and on the incidental parameters.

Equation (37) 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 Proposition 2.

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

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

Using the 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 $\sigma_{\alpha 2}(y_{1t}, y_{2t-1}) \equiv -\ln[1+\exp\{\alpha_2+\beta_{22}\ y_{2t-1}+\gamma_2 y_{1t}\}]$, we have:

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

Grouping terms, we have:

$$\begin{aligned}
\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \alpha, \beta) &\geq \ln p(y_{10}, y_{20} \mid \alpha) \\
&+ \textstyle\sum_{t=1}^{T} (1-y_{1t}) \ \ [\sigma_{\alpha 1}(y_{1t-1},0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\
&+ \textstyle\sum_{t=1}^{T} (1-y_{1t})\, y_{2t} \ \ [\alpha_2 + \beta_{22}\ y_{2t-1}] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t} \ \ [\sigma_{\alpha 1}(y_{1t-1},1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_{11}\ y_{1t-1} + \gamma_1] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t}y_{2t} \ \ [\alpha_2 + \beta_{22}\ y_{2t-1} + \gamma_2]
\end{aligned} \tag{A.3}$$

Since $\gamma_1 \leq 0$ and $\gamma_2 \leq 0$, we have that $\sigma_{\alpha 1}(y_{1t-1},1) \geq \sigma_{\alpha 1}(y_{1t-1},0)$ and $\sigma_{\alpha 2}(1, y_{2t-1}) \geq \sigma_{\alpha 2}(0, y_{2t-1})$. Therefore, in equation (A.3) we can replace $\sigma_{\alpha 1}(y_{1t-1},1)$ with $\sigma_{\alpha 1}(y_{1t-1},0)$ and $\sigma_{\alpha 2}(1, y_{2t-1})$ with $\sigma_{\alpha 2}(0, y_{2t-1})$ and the inequality still holds. Making this replacements, we have:

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

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$,$T^{(1,1)}$, S_{11}, and S_{12}, we have the following expression

for the lower bound $\ln \mathbb{P}_{L(2)}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right)$:

$$
\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) &\geq \ln \mathbb{P}_{L(2)}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) \\
&\equiv \ln p\left(y_{10}, y_{20} \mid \alpha\right) \\
&+ \left(y_{10} - y_{1T}\right) \ [\sigma_{\alpha 1}(1,0) - \sigma_{\alpha 1}(0,0)] + \left(y_{20} - y_{2T}\right) \ [\sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0)] \\
&+ T_1^{(1)} \ [\alpha_1 + \gamma_1 + \sigma_{\alpha 1}(1,0) - \sigma_{\alpha 1}(0,0)] + T_2^{(1)} \ [\alpha_2 + \sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0)] \\
&+ S_{11} \ \beta_{11} + S_{22} \ \beta_{22} + T^{(1,1)} \ \gamma_2
\end{aligned}
\tag{A.5}
$$

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

$$
\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) &\geq \ln p\left(y_{10}, y_{20} \mid \alpha\right) \\
&+ \textstyle\sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \ \left(\ln\left[1 - \Lambda\left(\alpha_1 + \beta_{11} \ y_{1t-1}\right)\right] + \ln\left[1 - \Lambda\left(\alpha_2 + \beta_{22} \ y_{2t-1}\right)\right]\right) \\
&+ \textstyle\sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \ \left(\ln\left[1 - \Lambda\left(\alpha_1 + \beta_{11} \ y_{1t-1} + \gamma_1\right)\right] + \ln \Lambda\left(\alpha_2 + \beta_{22} \ y_{2t-1} + \gamma_2\right)\right) \\
&+ \textstyle\sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \ \left(\ln \Lambda\left(\alpha_1 + \beta_{11} \ y_{1t-1}\right) + \ln\left[1 - \Lambda\left(\alpha_2 + \beta_{22} \ y_{2t-1}\right)\right]\right) \\
&+ \textstyle\sum_{t=1}^{T} y_{1t} y_{2t} \ \left(\ln \Lambda\left(\alpha_1 + \beta_{11} \ y_{1t-1} + \gamma_1\right) + \ln \Lambda\left(\alpha_2 + \beta_{22} \ y_{2t-1} + \gamma_2\right)\right)
\end{aligned}
\tag{A.6}
$$

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

$$
\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) &\geq \ln p\left(y_{10}, y_{20} \mid \alpha\right) \\
&+ \textstyle\sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\
&+ \textstyle\sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \ [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_2 + \beta_{22} \ y_{2t-1} + \gamma_2] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_1 + \beta_{11} \ y_{1t-1}] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t} y_{2t} \ [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_{11} \ y_{1t-1} + \gamma_1 + \alpha_2 + \beta_{22} \ y_{2t-1} + \gamma_2]
\end{aligned}
\tag{A.7}
$$

Grouping terms, we have:

$$
\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) &\geq \ln p\left(y_{10}, y_{20} \mid \alpha\right) \\
&+ \textstyle\sum_{t=1}^{T} \left(1 - y_{2t}\right) \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \ [\alpha_1 + \beta_{11} \ y_{1t-1}] \\
&+ \textstyle\sum_{t=1}^{T} y_{2t} \ [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_2 + \beta_{22} \ y_{2t-1} + \gamma_2] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t} y_{2t} \ [\alpha_1 + \beta_{11} \ y_{1t-1} + \gamma_1]
\end{aligned}
\tag{A.8}
$$

Since $\gamma_1 \leq 0$ and $\gamma_2 \leq 0$, we can replace $\sigma_{\alpha 1}(y_{1t-1}, 1)$ with $\sigma_{\alpha 1}(y_{1t-1}, 0)$ and $\sigma_{\alpha 2}(1, y_{2t-1})$ with $\sigma_{\alpha 2}(0, y_{2t-1})$ and the inequality still holds. Making this replacements, we have:

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

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

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

(c) Upper Bound $\ln \mathbb{P}_U\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right)$. For the upper bounds, we use the bounds for the choice probabilities $U(0,1|\mathbf{y}_{t-1};\alpha) \equiv [1 - \Lambda\left(\alpha_1 + \beta_{11} \ y_{1t-1} + \gamma_1\right)] \ \Lambda\left(\alpha_2 + \beta_{22} \ y_{2t-1}\right)$ and $U(1,0|\mathbf{y}_{t-1};\alpha) \equiv \Lambda\left(\alpha_1 + \beta_{11} \ y_{1t-1}\right) [1 - \Lambda\left(\alpha_2 + \beta_{22} \ y_{2t-1} + \gamma_2\right)]$. Then,

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

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

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) &\leq \ln p\left(y_{10}, y_{20} \mid \alpha\right) \\
&+ \textstyle\sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})] \\
&+ \textstyle\sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \ [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_2 + \beta_{22} \ y_{2t-1}] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \ [\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_{11} \ y_{1t-1}] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t} y_{2t} \ [\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_1 + \beta_{11} \ y_{1t-1} + \gamma_1 + \alpha_2 + \beta_{22} \ y_{2t-1} + \gamma_2]
\end{aligned} \tag{A.12}$$

Since $\gamma_1 \leq 0$ and $\gamma_2 \leq 0$, we can replace $\sigma_{\alpha 1}(y_{1t-1}, 0)$ with $\sigma_{\alpha 1}(y_{1t-1}, 1)$ and $\sigma_{\alpha 2}(0, y_{2t-1})$ with $\sigma_{\alpha 2}(1, y_{2t-1})$ and the inequality still holds. Making this replacements and grouping terms, we have:

$$\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) &\leq \ln p\left(y_{10}, y_{20} \mid \alpha\right) \\
&+ \textstyle\sum_{t=1}^{T} \sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) \\
&+ \textstyle\sum_{t=1}^{T} y_{2t} \ [\alpha_2 + \beta_{22} \ y_{2t-1}] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t} \ [\alpha_1 + \beta_{11} \ y_{1t-1}] \\
&+ \textstyle\sum_{t=1}^{T} y_{1t} y_{2t} \ [\gamma_1 + \gamma_2]
\end{aligned} \tag{A.13}$$

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$,$T^{(1,1)}$, S_{11}, and S_{12}, we have the following expression

for the upper bound $\ln \mathbb{P}_U\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right)$:

$$
\begin{aligned}
\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) &\leq \ln \mathbb{P}_U\left(\widetilde{\mathbf{y}} \mid \alpha, \beta\right) \\
&\equiv \ln p\left(y_{10}, y_{20} \mid \alpha\right) \\
&+ \left(y_{10} - y_{1T}\right) \; [\sigma_{\alpha 1}(1,0) - \sigma_{\alpha 1}(0,0)] + \left(y_{20} - y_{2T}\right) \; [\sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0)] \\
&+ T_1^{(1)} \; [\alpha_1 + \sigma_{\alpha 1}(1,0) - \sigma_{\alpha 1}(0,0)] + T_2^{(1)} \; [\alpha_2 + \sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0)] \\
&+ S_{11} \; \beta_{11} + S_{22} \; \beta_{22} + T^{(1,1)} \; [\gamma_1 + \gamma_2]
\end{aligned}
\tag{A.14}
$$

References

[1] Aguirregabiria, V., J. Gu, Y. Luo (2018): "Sufficient Statistics for Unobserved Heterogeneity in Structural Dynamic Logit Models," manuscript. Department of Economics, University of Toronto.

[2] 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.

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

[4] Aguirregabiria, V. and P. Mira (2010): "Dynamic discrete choice structural models: a survey', *Journal of Econometrics*, 156 (1), 38–67.

[5] Aguirregabiria, V., P. Mira, and H. Roman (2007): "An estimable dynamic model of entry, exit and growth in oligopoly retail markets," *American Economic Review*, 97, 449-454.

[6] Aguirregabiria, V. and A. Nevo (2013): "Recent developments in empirical IO: dynamic demand and dynamic games," in D. Acemoglu, M. Arellano and E. Deckel (eds), *Advances in Economics and Econometrics*. Tenth World Congress. Vol. 3, Econometrics, Cambridge: Cambridge University Press, pp. 53-122.

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

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

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

[10] 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.

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

[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] 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.

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

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

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

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

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

[21] 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.

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

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

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

[25] 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.

[26] 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.

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

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

[29] 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.

[30] 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.

[31] 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.

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

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

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

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

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

[37] 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.

[38] 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.

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

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

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

[42] 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.

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

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

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

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

[47] 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.

[48] 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.

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

[50] 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.

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

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

[53] 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.

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

[55] 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.

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