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

Victor Aguirregabiria[†] University of Toronto CEPR Jiaying Gu[‡] University of Toronto Pedro Mira [§] CEMFI

November 11, 2024

Abstract

In the structural estimation of dynamic discrete choice games, misspecification of unobserved heterogeneity can introduce significant bias in two key categories of structural parameters: those related to dynamic state dependence, such as adjustment or switching costs, and those that reflect strategic interactions among players, like competitive or peer effects. This paper examines the identification of these parameters within models that account for unobserved heterogeneity using a fixed-effects framework, and with short panel data. Drawing on recent advances in functional differencing techniques, we establish identification results for different game types, including distinctions between simultaneous versus sequential moves, myopic versus forward-looking decision-makers, and one-directional versus two-directional strategic interactions. We illustrate the applicability of these findings through an empirical study of a dynamic game of price competition.

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

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

^{*}We thank the valuable feedback provided by seminar participants at the International Panel Data Conference, the IAAE Annual Conference in Nicosia, the 3rd Conference on Structural Dynamic Models in Chicago, the Banff Empirical Microeconomics Workshop, and the Bristol Econometrics Study Group.

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

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

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

1 Introduction

Dynamic games are powerful tools for analyzing economic and social phenomena involving intertemporal interactions between agents. The structural estimation of such games has garnered significant attention, particularly in the study of oligopoly competition dynamics, following the influential work of Ericson and Pakes (1995), with empirical applications spanning various industries.¹ In addition, econometric models of dynamic games have been applied to a wide range of contexts, such as dynamic interactions within households (Eckstein and Lifshitz, 2015), longterm care decisions (Sovinsky and Stern, 2016), electoral competition (Sieg and Yoon, 2017), and the ratification of international treaties (Wagner, 2016). Moreover, there is a substantial body of literature on dynamic discrete choice models with social interactions where agents do not exhibit forward-looking behavior, as explored by Brock and Durlauf (2007) and Blume, Brock, Durlauf, and Ioannides (2011).

In dynamic games, model predictions are primarily driven by two types of structural parameters: those capturing dynamic state dependence –such as costs associated with switching, adjustment, investment, or entry and exit (referred to as the *dynamic* parameters) –and those representing the influence of other players' actions on a player's payoff, which arise from competition, spillovers, peer effects, or social interactions (referred to as the *game* parameters). The identification of these parameters is highly sensitive to the model's assumptions about the stochastic properties of variables that are observable to the players but unobservable to the researcher, which we can refer to as *unobserved heterogeneity*.

In dynamic models, it is well-established that neglecting or misspecifying persistent unobserved heterogeneity can lead to significant biases when estimating structural parameters that reflect true dynamics (Heckman, 1981). Unobserved heterogeneity can create spurious dynamics that become entangled with genuine state dependence. Similarly, failing to account for correlated unobserved heterogeneity across players in the game estimation literature can result in substan-

¹For a recent survey of this literature, see Aguirregabiria, Collard-Wexler, and Ryan (2021).

tial biases when estimating parameters related to strategic or social interactions (Bresnahan and Reiss, 1991, Blume, Brock, Durlauf, and Ioannides, 2011). In such cases, common-across-players unobserved factors may become confounded with strategic, social, or peer effects.

In this paper, we study the identification of dynamic games in empirical settings where a few players are observed over a few periods but across many markets. Our focus is on identifying the model's parameters when there is time-invariant unobserved heterogeneity at the market or market-player level, which follows a nonparametric distribution with unrestricted support, and thus, using a fixed effects panel data framework.

We begin by extending the application of the fixed-effect conditional likelihood method, originally introduced by Cox (1958), Rasch (1961), Andersen (1970), and Chamberlain (1980), to dynamic discrete choice games. In cases where this method fails to identify all the structural parameters in our model, we employ a functional differencing approach as proposed by Bonhomme (2012). Specifically, we use a variant of the technique recently introduced by Dobronyi, Gu, and Kim (2021), which derives a set of moment conditions and moment inequalities implied by the fixed-effects dynamic model. Our analysis shows that this method successfully identifies critical parameters that remain unidentified with the conditional likelihood approach alone. By incorporating functional differencing, we enhance the identification of key structural parameters that would otherwise be overlooked.

Our paper contributes to the literature on the identification and estimation of dynamic games with unobserved heterogeneity. Previous studies in this field have used a random effects approach, applying a finite mixture model for unobserved heterogeneity and imposing restrictions on the initial conditions (e.g., Aguirregabiria and Mira, 2007; Kasahara and Shimotsu, 2009; Arcidiacono and Miller, 2011, among others). In contrast, our research focuses on identifying structural parameters without imposing constraints on the distribution of unobserved heterogeneity, its support, or the initial conditions. By relaxing these assumptions, we offer a more flexible framework for studying dynamic games with unobserved heterogeneity.

This paper also contributes to the literature on the identification and estimation of struc-

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

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

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

2 Model

2.1 Framework

We focus specifically on two-player binary choice games. For our analysis, we label the players as *i* and *j*, where $i, j \in \{1, 2\}$. To represent the temporal dimension, we use discrete time with the index *t*, ranging from 1 to *T*, for different periods. The game between the two players takes place within a defined *market*, which can vary depending on the empirical application. A market could refer to a geographic area, a school, a family, an industry, an election, or other contexts. We denote different markets using the index m, where $m \in \{1, 2, ..., M\}$. For simplicity, we temporarily omit the market subindex in our notation.

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

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

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

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

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

$$V_{i}^{\sigma}(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_{t}) = \max_{y_{it} \in \{0,1\}} \left\{ u_{i}\left(y_{it}, y_{jt}, \mathbf{y}_{t-1}\right) + \varepsilon_{it}(y_{it}) + \delta_{i} \int V_{i}^{\sigma}(\mathbf{y}_{t}, \boldsymbol{\varepsilon}_{t+1}) \ g(\boldsymbol{\varepsilon}_{t+1}) \ d\boldsymbol{\varepsilon}_{t+1} \right\}, \quad (2)$$

with $y_{jt} = \sigma_j(\mathbf{y}_{t-1}, \boldsymbol{\varepsilon}_t)$. The integral accounts for the expectation over $\boldsymbol{\varepsilon}_{t+1}$, and $g(\boldsymbol{\varepsilon}_{t+1})$ is the density function of $\boldsymbol{\varepsilon}_{t+1}$.

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

$$y_{1t} = 1 \left\{ \widetilde{u}_{1} (y_{2t}, \mathbf{y}_{t-1}) + \widetilde{V}_{1} (y_{2t}) - \varepsilon_{1t} \ge 0 \right\}$$

$$y_{2t} = 1 \left\{ \widetilde{u}_{2} (y_{1t}, \mathbf{y}_{t-1}) + \widetilde{V}_{2} (y_{1t}) - \varepsilon_{2t} \ge 0 \right\}$$
(3)

with $\varepsilon_{it} \equiv \varepsilon_{it}(0) - \varepsilon_{it}(1)$. Here, $\widetilde{u}_i(y_{jt}, \mathbf{y}_{t-1})$ represents the utility difference $u_i(1, y_{jt}, \mathbf{y}_{t-1}) - u_i(0, y_{jt}, \mathbf{y}_{t-1})$. The term $\widetilde{V}_i(y_{jt})$ captures the difference in continuation values:

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

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

2.2 Structural and incidental parameters

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

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

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

Here, α_{im} captures market and player characteristics that are not observable to the researcher. To account for these unobservable factors, we define the vector $\boldsymbol{\alpha}_m \equiv (\alpha_{1m}, \alpha_{2m})$, which represents the fixed effects specific to market m. These fixed effects are referred to as the incidental parameters of the model.

The term $\beta_i y_{im,t-1}$ captures state dependence with respect to the lagged value of the player's action. It incorporates factors such as adjustment costs or switching costs, influencing a player's current decision based on their previous action. The term $\gamma_i y_{jmt}$ captures contemporaneous strategic effects between the players' actions. It accounts for how the choice of one player, j, may influence the payoff of the other player, i, in the same period. The term $\lambda_i y_{jm,t-1}$ represents state dependence with respect to the lagged value of the other player's action. It captures the dynamic strategic interactions between the two players, reflecting how a player's previous choice may have an impact on the other player's payoff. Together, these components form the structure of the model, incorporating fixed effects, contemporaneous strategic effects, and state dependence, to capture the dynamics of the two-player binary choice game.

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

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

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

2.3 Multiple equilibria and probabilities of game outcomes

The model presents two forms of the multiple equilibrium problem. First, given the model's primitives, there may exist multiple strategy functions that satisfy the system of best response conditions characterizing the model's Markov Perfect Equilibrium (MPE). These distinct strategy functions imply different continuation value functions, $\tilde{V}_{1m}(.)$ and $\tilde{V}_{2m}(.)$. Second, even if we fix the continuation value functions, $\tilde{V}_{1m}(.)$ and $\tilde{V}_{2m}(.)$. Second, even if the state variables, $\mathbf{y}_{m,t-1}$ and $\boldsymbol{\varepsilon}_{mt}$, the model generates multiple predictions for the equilibrium values of (y_{1mt}, y_{2mt}) . This second issue parallels the multiple equilibria problem commonly observed in static games of complete information, as discussed in seminal works such as Bresnahan and Reiss (1991) and Tamer (2003).

The model implies a partition of the space of time-varying unobservables $(\varepsilon_{1mt}, \varepsilon_{2mt})$ such that each region in the partition corresponds to a prediction (or multiple predictions) about players' choices. For $y_j \in \{0, 1\}$, let $e_{imt}^{y_j}$ represent the value of ε_{imt} that makes player *i* indifferent between choosing action 0 or action 1. The model then defines this threshold value as follows:

$$e_{imt}^{y_j} \equiv \alpha_{im} + \widetilde{V}_{im}(y_j) + \gamma_i \ y_j + \beta_i \ y_{im,t-1} + \lambda_i \ y_{jm,t-1} \tag{7}$$

The sign of $e_{imt}^1 - e_{imt}^0$ corresponds to the sign of $\gamma_i + \tilde{V}_{im}(1) - \tilde{V}_{im}(0)$. Notably, $\tilde{V}_{im}(1) - \tilde{V}_{im}(0)$ can be rewritten as $V_{im}(1,1) - V_{im}(0,1) - V_{im}(1,0) + V_{im}(0,0)$, which captures the supermodularity or submodularity of the value function in relation to both players' actions. This implies that the sign of this endogenous term matches the sign of the parameter γ_i . When $\gamma_i \geq 0$ (or $\gamma_i \leq 0$), the dynamic game exhibits strategic complementarity (or substitutability), leading to a supermodular (or submodular) value function. In this case, $V_{im}(1,1) - V_{im}(0,1) - V_{im}(1,0) +$ $V_{im}(0,0) \geq 0$ (or ≤ 0), which ensures $e_{imt}^1 \geq e_{imt}^0$ (or $e_{imt}^1 \leq e_{imt}^0$). These threshold values – two for each player– define two vertical lines and two horizontal lines in the two-dimension space of $(\varepsilon_{1mt}, \varepsilon_{2mt})$. The space for the four possible equilibrium outcomes for the players' decisions – (0,0), (0,1), (1,0), and (1,1) – and the corresponding probabilities of these outcomes, can be described using these threshold values. The form of this partition depends on the sign of the parameters γ_1 and γ_2 . Figure 1 represents the threshold values and the regions for the different equilibrium outcomes for the case with $\gamma_1 \leq 0$ and $\gamma_2 \leq 0$. The model provides unique predictions for the probabilities of outcomes (1, 1) and (0, 0):

$$\begin{cases} \mathbb{P}(0,0 \mid \mathbf{y}_{m,t-1}, \boldsymbol{\alpha}_{m}) = \frac{1}{1 + \exp\{e_{1mt}^{0}\}} \frac{1}{1 + \exp\{e_{2mt}^{0}\}} \\ \mathbb{P}(1,1 \mid \mathbf{y}_{m,t-1}, \boldsymbol{\alpha}_{m}) = \frac{\exp\{e_{1mt}^{1}\}}{1 + \exp\{e_{1mt}^{1}\}} \frac{\exp\{e_{2mt}^{1}\}}{1 + \exp\{e_{2mt}^{1}\}} \end{cases}$$
(8)

The quadrangle in the center of Figure 1 is associated with two possible outcomes or equilibria of the game: (1,0) and (0,1). This region with multiple equilibria implies that the model does not have unique predictions on the probabilities of outcomes (0,1) and (1,0). However, the model establishes bounds on the values of these probabilities.

The (sharp) upper bound to the probability of outcome (1,0) is given by the region up and to the left of the blue right angle. The upper bound to the probability of outcome (0,1) is associated with the region down and to the right of the red right angle. These upper bounds are the product of two logit probabilities:

$$\begin{cases} U(0,1 \mid \mathbf{y}_{m,t-1}, \boldsymbol{\alpha}_{m}) \equiv \frac{1}{1 + \exp\{e_{1mt}^{1}\}} \frac{\exp\{e_{2mt}^{0}\}}{1 + \exp\{e_{2mt}^{0}\}} \\ U(1,0 \mid \mathbf{y}_{m,t-1}, \boldsymbol{\alpha}_{m}) \equiv \frac{\exp\{e_{1mt}^{0}\}}{1 + \exp\{e_{1mt}^{0}\}} \frac{1}{1 + \exp\{e_{2mt}^{1}\}} \end{cases}$$
(9)

The sharp lower bounds for the probabilities of outcomes (0,1) and (1,0) correspond to the regions defined by the upper bounds, excluding the central quadrangle shown in Figure 1.

Figure 1: Regions in the Space of $(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)$



Unlike the upper bounds, these lower bounds are not simple products of logit probabilities but rather sums of these products. This distinction has important implications for identifying the structural parameters in this fixed effects model.

For this reason, we also consider non-sharp lower bounds that follow a logit probability structure. For each outcome, (0, 1) and (1, 0), we can construct two non-sharp lower bounds based on the product of logit probabilities. For outcome (0, 1), the two bounds are:

$$\begin{cases}
L^{0,0}(0,1 \mid \mathbf{y}_{m,t-1}, \boldsymbol{\alpha}_{m}) \equiv \frac{1}{1 + \exp\{e_{1mt}^{0}\}} \frac{\exp\{e_{2mt}^{0}\}}{1 + \exp\{e_{mt}^{0}\}} \\
L^{1,1}(0,1 \mid \mathbf{y}_{m,t-1}, \boldsymbol{\alpha}_{m}) \equiv \frac{1}{1 + \exp\{e_{1mt}^{1}\}} \frac{\exp\{e_{2mt}^{1}\}}{1 + \exp\{e_{2mt}^{1}\}}
\end{cases} (10)$$

The super-index in the names of these lower bounds, $L^{0,0}$ and $L^{1,1}$, reflect the beliefs to which the players are best responding. In $L^{0,0}$, both players are responding to the belief that their opponent will choose action 0, while in $L^{1,1}$, both players are responding to the belief that their opponent will choose action 1.

Similarly, for the probability of outcome (1,0), the non-sharp lower bounds based on the product of logit probabilities are:

$$\begin{cases}
L^{0,0}(1,0 \mid \mathbf{y}_{m,t-1}, \boldsymbol{\alpha}_{m}) \equiv \frac{\exp\{e_{1mt}^{0}\}}{1 + \exp\{e_{1mt}^{0}\}} \frac{1}{1 + \exp\{e_{2mt}^{0}\}} \\
L^{1,1}(1,0 \mid \mathbf{y}_{m,t-1}, \boldsymbol{\alpha}_{m}) \equiv \frac{\exp\{e_{1mt}^{1}\}}{1 + \exp\{e_{1mt}^{1}\}} \frac{1}{1 + \exp\{e_{2mt}^{1}\}}
\end{cases}$$
(11)

Again, the super-index in the names of these lower bounds reflect the beliefs to which the players are best responding.

2.4 Different versions of the model

The identification results in this paper vary across different versions of the model, depending on three key criteria.

a. Myopic versus forward-looking players: A player is considered myopic if her discount factor, δ_i , is zero, meaning she does not take future periods into account when making decisions. In contrast, a player with a non-zero discount factor is classified as forward-looking, indicating that she considers future payoffs. For myopic players, the continuation value \tilde{V}_i is zero.

b. Strategic interactions – Contemporaneous versus lagged, and one- versus twodirectional: The scope and nature of contemporaneous strategic interactions between players are governed by the parameters γ_1 and γ_2 , while lagged strategic interactions are captured by λ_1 and λ_2 . Most empirical studies incorporate either contemporaneous or lagged interactions, but not both. Following this approach, we explore two models: one without contemporaneous interactions but with lagged ones (i.e., $\gamma_1 = \gamma_2 = 0$), and another without lagged interactions but with contemporaneous ones (i.e., $\lambda_1 = \lambda_2 = 0$). Identifying contemporaneous interactions is more complex than identifying lagged ones. For these more challenging models, we differentiate between cases where both γ_1 and γ_2 are non-zero, indicating two-directional interactions between players, and cases where one of these parameters is zero, implying one-directional interactions.

b. Sequential versus simultaneous moves: In each period t, players make decisions either simultaneously or sequentially. The nature of these moves affects the set of equilibria in the model. For games without, or with one-directional, contemporaneous interactions, the equilibrium is unique, regardless of whether moves are sequential or simultaneous. However, in games with two-directional interactions, this assumption becomes crucial. When moves are simultaneous, the model may have multiple equilibria. In contrast, sequential moves lead to a unique equilibrium.

For instance, consider a sequential-move game where player 1 moves first. In the central quadrangle of Figure 1, the unique equilibrium (known as the Subgame Perfect Nash equilibrium) is (1,0). Player 1 understands that if she chooses action 1, player 2 will choose 0, and if she selects 0, player 2 will respond with 1. Hence, player 1's choice determines whether equilibrium (0, 1) or (1, 0) is selected. With $\gamma_1 \leq 0$, the equilibrium that maximizes player 1's profit is (1, 0). Therefore, in this sequential-move game, the central quadrangle of Figure 1 uniquely corresponds to the outcome (1, 0).

Table 1 presents an overview of the different model versions analyzed in this paper, along with a summary of the identification results. The table reveals several key patterns in our findings. First, we observe that point identification of the dynamic parameters, β_1 and β_2 , is more general and attainable than the identification of the strategic interaction parameters, γ_1 and γ_2 . This suggests that precise estimates of the dynamic parameters can be obtained in a wider range of scenarios. Second, point identification is achievable only when certain restrictions are imposed, such as assumptions regarding myopic behavior, sequential moves, or the structure

No Contemporaneous Interactions $\gamma_1 = \gamma_2 = 0$ One-Direction Interactions $\gamma_1 = 0$		Two-Direction Interactions Sequential Move	Two-Direction Interactions Simultaneous Move				
MYOPIC PLAYERS: $\tilde{V}_{imt} = 0$							
Point identification $\beta_1, \beta_2, \lambda_1, \lambda_2$ Point identification $\beta_1, \beta_2, \gamma_2$ Point iden. β_1, β_2 Partial iden. γ_1, γ_2 Partial identification $\beta_1, \beta_2, \gamma_1, \gamma_2$							
FORWARD-LOOKING PLAYERS: $\widetilde{V}_{imt} \neq 0$							

Table 1: Different Models and Summary of Identification Results

Point iden. β_1, β_2 Point iden. β_1, β_2 Point iden. β_1, β_2 Point iden. β_1, β_2 Partial iden. β_1, β_2			1110 /	
	Point iden. β_1, β_2	Point iden. β_1, β_2 Partial iden. γ_2	Point iden. β_1, β_2	Partial iden. β_1, β_2

of strategic interactions. Third, under the assumption of sequential moves, point identification of the dynamic parameters β is possible without needing to impose restrictions on the players' discount factors δ or the strategic parameters γ . This finding indicates that the sequential move framework alone provides valuable identifying power for the dynamic parameters.

Overall, Table 1 underscores the varying identification outcomes across different model versions and highlights the trade-offs between robustness and precision in the identification process. It also emphasizes the importance of incorporating specific assumptions to obtain precise parameter estimates.

3 Identification

The sampling framework involves a random sample of M markets. Within each market, the data consists of the observed sequence of choices made between periods 1 and T, as well as the initial conditions (y_{1m0}, y_{2m0}) . The number of markets M is large and T is small. To simplify

notation, we will omit the market subindex m for the remainder of this section. To represent the complete history of choices within a market, we use the vector $\tilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, ..., T)$.

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

Before we present our identification results for different versions of the model, we describe the different methods we use to establish them. We also present a novel Proposition 1 that is key in implementing functional differencing to obtain our results.

a. Conditional likelihood without multiple equilibria. The versions of the model without multiple equilibria imply the following expression for the probability of a market history conditional of the fixed effects and the initial condition:

$$\mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{y}_{0}, \boldsymbol{\alpha}, \boldsymbol{\theta}\right) = \prod_{i=1}^{2} \prod_{t=1}^{T} \Lambda\left(y_{it} | \alpha_{i}, y_{jt}, \mathbf{y}_{t-1}, \boldsymbol{\theta}\right)$$
(12)

where $\Lambda\left(y_{it}|\alpha_i, y_{jt}, \boldsymbol{y}_{t-1}, \boldsymbol{\theta}\right)$ represents the logit probability of choice y_{it} .

The logit model possesses a crucial property that enables an additive separability of the log-likelihood with respect to the incidental parameters $\boldsymbol{\alpha}$ and the structural parameters $\boldsymbol{\theta}$. Specifically, the logarithm of the probability of a market history can be written as:

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{y}_{0}, \alpha, \boldsymbol{\theta}\right) = \mathbf{s}\left(\widetilde{\mathbf{y}}, \mathbf{y}_{0}\right)' \mathbf{g}_{\alpha} + \mathbf{c}\left(\widetilde{\mathbf{y}}\right)' \boldsymbol{\theta}$$
(13)

where $\mathbf{s}(\widetilde{\mathbf{y}}, \mathbf{y}_0)$ and $\mathbf{c}(\widetilde{\mathbf{y}})$ are vectors of statistics, \mathbf{g}_{α} is a vector of functions of the incidental parameters, and $\boldsymbol{\theta}$ is the vector of structural parameters. This structure has two key implications. First, the vector of statistics $\mathbf{s}(\widetilde{\mathbf{y}}, \mathbf{y}_0)$ is sufficient for the incidental parameters $\boldsymbol{\alpha}$: that is, $\mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}, \mathbf{y}_0), \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}, \mathbf{y}_0), \boldsymbol{\theta})$. Second, a parameter in the vector $\boldsymbol{\theta}$ is identified if the corresponding element in the vector $\mathbf{c}(\widetilde{\mathbf{y}})$ has variation across some market histories after we condition on the vector of sufficient statistics $\mathbf{s}(\widetilde{\mathbf{y}}, \boldsymbol{y}_0)$.

b. Functional differencing without multiple equilibria. Recent studies by Honoré and Weidner (2020) and Dobronyi, Gu, and Kim (2021) have employed a functional differencing approach inspired by Bonhomme (2012) to establish parameter identification in dynamic logit models that cannot be identified using the conditional likelihood method. In this paper, we apply a similar approach.

For models with a unique equilibrium –and thus unique predictions about the probability of a market history– the following Proposition 1 is key. This proposition mirrors Proposition 2 in Aguirregabiria and Carro (2024), which establishes necessary and sufficient conditions for identifying Average Marginal Effects in single-agent dynamic discrete choice models with fixed effects. Similarly, we demonstrate that this result also holds for the identification of structural parameters in dynamic games with fixed effects.

Proposition 1. Consider a fixed effects dynamic game with equilibrium uniqueness and represented by the probability function $\mathbb{P}(\tilde{\mathbf{y}} \mid \boldsymbol{y}_0, \boldsymbol{\alpha}, \boldsymbol{\theta})$ in equation (12). The vector of structural parameters $\boldsymbol{\theta}$ is identified if and only if the following conditions hold:

i. There is a vector of weighting functions $\boldsymbol{w}(\boldsymbol{y}_0, \widetilde{\mathbf{y}}, \boldsymbol{\theta})$ mapping from $\{0, 1\}^{2(T+1)} \times \Theta$ to $\mathbb{R}^{dim(\boldsymbol{\theta})}$, which satisfies the following system of equations:

$$\sum_{\widetilde{\mathbf{y}} \in \mathcal{Y}^{2T}} \boldsymbol{w}(\boldsymbol{y}_0, \widetilde{\mathbf{y}}, \boldsymbol{\theta}) \mathbb{P}(\widetilde{\mathbf{y}} \mid \boldsymbol{y}_0, \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbf{0}$$
(14)

for every $\boldsymbol{y}_0 \in \{0,1\}^2$ and $\boldsymbol{\alpha} \in \mathbb{R}^2$.

ii. Given the vector of weighting functions $\boldsymbol{w}(\boldsymbol{y}_0, \widetilde{\mathbf{y}}, \boldsymbol{\theta})$, the following system of dim $(\boldsymbol{\theta})$ equa-

tions has a unique solution in $\boldsymbol{\theta}$:

$$\sum_{(\boldsymbol{y}_0, \widetilde{\mathbf{y}}) \in \mathcal{Y}^{2(T+1)}} \boldsymbol{w}(\boldsymbol{y}_0, \widetilde{\mathbf{y}}, \boldsymbol{\theta}) \mathbb{P}(\boldsymbol{y}_0, \widetilde{\mathbf{y}}) = \mathbf{0}$$
(15)

where $\mathbb{P}(\boldsymbol{y}_0, \widetilde{\mathbf{y}})$ represents the empirical distribution of $(\boldsymbol{y}_0, \widetilde{\mathbf{y}})$ in the data.

The solution of the system of equations in (15) identifies the true value of θ .

Proof of Proposition 1. See Appendix A.1.

Proposition 1 imposes no restrictions on the distribution of the time-varying i.i.d. unobservables, ε_{it} . As a result, it applies to a broad class of fixed-effects dynamic discrete choice models, extending beyond the logit class.

The system of equations in (14), described in condition (i), involves infinitely many restrictions – one for each value of α – but only a finite number of unknown weights, corresponding to the 4^T possible market histories. Given this, one might perceive that satisfying the necessary and sufficient conditions for identification in condition (i) of Proposition 1 seems improbable. In other words, a solution is only likely if a specific structure exists, where a finite-dimensional vector of weights solves the system of infinite restrictions.

The propositions presented throughout this paper, which establish point identification of the structural parameters, rely on such a specific structure. In particular, the logit model implies that the system of equations in (14) can be expressed as a finite-order polynomial in the variables e^{α_1} and e^{α_2} . This finding shows that a solution exists if and only if the coefficients of every monomial term in this polynomial are set to zero. This property reduces the infinite system of equations to a finite linear system with a finite number of unknowns. Moreover, if a solution exists, it leads to a closed-form expression for the weights.

3.1 Myopic players

3.1.1 Model with no contemporaneous strategic interactions

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

$$\begin{cases} y_{1t} = 1 \{ \alpha_1 + \beta_1 \ y_{1t-1} + \lambda_1 \ y_{2t-1} - \varepsilon_{1t} \ge 0 \} \\ y_{2t} = 1 \{ \alpha_2 + \beta_2 \ y_{2t-1} + \lambda_2 \ y_{1t-1} - \varepsilon_{2t} \ge 0 \} \end{cases}$$
(16)

This is an autoregressive bivariate logit model. Narendranthan, Nickell, and Metcalf (1985) consider this model in their study of the joint dynamics of unemployment and sickness. They present a proof for the identification of the parameters using the same conditional likelihood approach as in our paper.² Consequently, the identification of this model is a well-established result in the literature. We include this result as it serves as a straightforward example for introducing notation and ensuring comprehensiveness.

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

$$\mathbb{P}\left(\widetilde{\boldsymbol{y}} \mid \boldsymbol{y}_{0}, \boldsymbol{\alpha}, \boldsymbol{\theta}\right) = \prod_{i=1}^{2} \prod_{t=1}^{T} \frac{\exp\left\{ y_{it} \left[\alpha_{i} + \beta_{i} \ y_{it-1} + \lambda_{i} \ y_{jt-1}\right] \right\}}{1 + \exp\left\{\alpha_{i} + \beta_{i} \ y_{it-} + \lambda_{i} \ y_{jt-1}\right\}}$$
(17)

Define, for $i \in \{1, 2\}$, function $\sigma_{\alpha_i}(y_1, y_2) \equiv -\ln[1 + \exp\{\alpha_i + \beta_i \ y_1 + \lambda_i \ y_2\}]$, and let $\sigma_{\alpha}(y_1, y_2) \equiv \sigma_{\alpha_1}(y_1, y_2) + \sigma_{\alpha_2}(y_1, y_2)$. Given a choice history $\widetilde{\mathbf{y}}$, define the statistics:

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

²See Honoré and Kyriazidou (2019) and Honoré and De Paula (2021) for their recent analysis of this model.

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

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{y}_{0}, \boldsymbol{\alpha}, \boldsymbol{\theta}\right) = \alpha_{1} T_{1}^{(1)} + \alpha_{2} T_{2}^{(1)} + \beta_{1} C_{11} + \lambda_{1} C_{12} + \lambda_{2} C_{21} + \beta_{2} C_{22} + \sum_{y_{1}, y_{2}} \sigma_{\boldsymbol{\alpha}} \left(y_{1}, y_{2}\right) \left[T^{(y_{1}, y_{2})} + 1\left\{\left(y_{10}, y_{20}\right) = \left(y_{1}, y_{2}\right)\right\} - 1\left\{\left(y_{1T}, y_{2T}\right) = \left(y_{1}, y_{2}\right)\right\}\right]$$

$$(18)$$

Or using the compact expression in equation (13), we have that:

$$\begin{aligned} \mathbf{s}(\boldsymbol{y}_{0}, \widetilde{\mathbf{y}}) &= [1, y_{10}, y_{20}, y_{10}y_{20} ; 1, y_{1T}, y_{2T}, y_{1T}y_{2T} ; T, T_{1}^{(1)}, T_{2}^{(1)}, T_{1}^{(1)}]' \\ \mathbf{c}(\widetilde{\mathbf{y}}) &= [C_{11}, C_{12}, C_{21}, C_{22}]' \\ \boldsymbol{\theta} &= [\beta_{1}, \lambda_{1}, \lambda_{2}, \beta_{2}]' \end{aligned}$$
(19)

Given (13) and (19) we can establish the following identification result.

Proposition 2. In the myopic dynamic game without contemporaneous interactions, as described in equation (16), the conditional likelihood approach establishes that structural parameters $\beta_1, \beta_2, \lambda_1, \text{ and } \lambda_2$ are point identified when $T \ge 3$.

Proof of Proposition 2. See Appendix A.2.

The proof of Proposition 1 demonstrates that $\mathbf{s}(\tilde{\mathbf{y}})$ is a sufficient statistic for $\boldsymbol{\alpha}$. It also shows that for each structural parameter — β_1 , β_2 , λ_1 , and λ_2 — there exists a pair of market histories, $A = \{\boldsymbol{y}_0, \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{y}_3\}$ and $B = \{\boldsymbol{y}_0, \boldsymbol{b}, \boldsymbol{a}, \boldsymbol{y}_3\}$, with $\boldsymbol{a} \neq \boldsymbol{b}$, such that the structural parameter is identified by the log-odds ratio $\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$. Table 2 provides examples of history pairs that identify each parameter for panels with T = 3.

3.1.2 Myopic players with one-direction strategic interactions

Now, we relax the condition of no contemporaneous strategic interactions and allow γ_2 to be different to zero: there is a contemporaneous effect of y_1 on y_2 . We still keep the restriction

Pairs of Histories Identifying Structural Parameters (T=3)							
$A = \{ y_0, a, b, y_3 \}; B = \{ y_0, b, a, y_3 \}$							
	$oldsymbol{y}_0$	a	b	$oldsymbol{y}_3$	$\ln \mathbb{P}\left(A\right) - \ln \mathbb{P}\left(B\right)$		
Case 1:	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c}1\\0\end{array}\right)$	$\left(\begin{array}{c}1\\0\end{array}\right)$	eta_1		
Case 2:	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\1\end{array}\right)$	$\left(\begin{array}{c} 0\\1\end{array}\right)$	β_2		
Case 3:	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\1\end{array}\right)$	$\left(\begin{array}{c}1\\0\end{array}\right)$	λ_1		
Case 4:	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c}1\\0\end{array}\right)$	$\left(\begin{array}{c}0\\1\end{array}\right)$	λ_2		

Table 2: Myopic Dynamic Game Without Contemporaneous Effects

 $\gamma_1 = 0$ – no contemporaneous effect of y_2 on y_1 , and include the restriction $\lambda_2 = 0$. That is, the model is defined by the following best response functions:

$$\begin{cases} y_{1t} = 1 \{ \alpha_1 + \beta_1 \ y_{1t-1} + \lambda_1 \ y_{2t-1} - \varepsilon_{1t} \ge 0 \} \\ y_{2t} = 1 \{ \alpha_2 + \gamma_2 \ y_{1t} + \beta_2 \ y_{2t-1} - \varepsilon_{2t} \ge 0 \} \end{cases}$$
(20)

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

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}}|\boldsymbol{y}_{0},\boldsymbol{\alpha},\boldsymbol{\theta}\right) = \alpha_{1} T_{1}^{(1)} + \alpha_{2} T_{2}^{(1)} + \sum_{t=1}^{T} \sigma_{\boldsymbol{\alpha}_{1}}\left(y_{1t-1}, y_{2t-1}\right) + \sigma_{\boldsymbol{\alpha}_{2}}\left(y_{1t}, y_{2t-1}\right) + \beta_{1} C_{11} + \lambda_{1} C_{12} + \beta_{2} C_{22} + \gamma_{2} T^{(1,1)}$$

$$(21)$$

with $\sigma_{\alpha_1}(y_1, y_2) \equiv -\ln[1 + \exp\{\alpha_1 + \beta_1 y_1 + \lambda_1 y_2\}]$ and $\sigma_{\alpha_2}(y_1, y_2) \equiv -\ln[1 + \exp\{\alpha_2 + \gamma_2 y_1 + \beta_2 y_2\}]$.

By comparing equations (21) and (18), we can find two important differences. Firstly, equation (21) includes the term $\gamma_2 T^{(1,1)}$, which is absent in (18). Secondly, in equation (21), the term that depends on incidental parameters includes not only the sum $\sum_{t=1}^{T} \sigma_{\alpha_1}(y_{1,t-1}, y_{2t-1})$ – which is also present in equation (18) – but also the sum $\sum_{t=1}^{T} \sigma_{\alpha_2}(y_{1t}, y_{2t-1})$, which did not appear in equation (18). These differences have implications for parameter identification.

Similarly as for the previous model, we can rewrite the right hand side of equation (21) as $\mathbf{s}(\boldsymbol{y}_0, \widetilde{\mathbf{y}})' \mathbf{g}_{\alpha} + \mathbf{c}(\widetilde{\mathbf{y}})' \boldsymbol{\theta}^*$, but now the vectors of statistics $\mathbf{s}(\boldsymbol{y}_0, \widetilde{\mathbf{y}})$ and $\mathbf{c}(\widetilde{\mathbf{y}})$, and the vector of identified parameters $\boldsymbol{\theta}^*$ are different. More specifically,³

$$\begin{cases} \mathbf{s}(\boldsymbol{y}_{0}, \widetilde{\mathbf{y}}) = [1, y_{10}, y_{20}, y_{10}y_{20} ; 1, y_{1T}, y_{2T}, y_{1T}y_{2T} ; T, T, T_{1}^{(1)}, T_{2}^{(1)}, T^{(1,1)} ; C_{12}]' \\ \mathbf{c}(\widetilde{\mathbf{y}}) = [C_{11}, C_{22}]' \\ \boldsymbol{\theta}^{*} = [\beta_{1}, \beta_{2}]' \end{cases}$$

$$(22)$$

There are some fundamental differences with respect to the model without contemporaneous strategic interactions. First, the statistic C_{12} and the structural parameter λ_1 appear in the logprobability of a choice history through the term C_{12} ($\Delta\sigma_{\alpha_2} + \lambda_1$). Without further restrictions, we have that the incidental parameter $\Delta\sigma_{\alpha_2}$ is not zero. This implies that this sufficient statistics approach cannot identify parameter λ_1 . Second, the statistic $T^{(1,1)}$ and the structural parameter γ_2 appear through the term $T^{(1,1)}$ ($\Delta\sigma_{\alpha_1} + \gamma_2$). Without further restrictions, the sufficient statistics approach does not identify parameter γ_2 .

Proposition 3 establishes the point identification of dynamic parameters β_1 and β_2 without further restrictions, as well as necessary and sufficient conditions for the identification of γ_2 and λ_1 when using a conditional likelihood approach.

Proposition 3. In the myopic dynamic game with one-direction contemporaneous interactions described in equation (20), using a conditional likelihood approach and with $T \ge 3$: (A) parameters β_1 and β_2 are point identified; (B) a necessary and sufficient condition for the identification of parameter γ_2 is that $\beta_1 = 0$ or $\lambda_1 = 0$; (C) a necessary and sufficient condition for the identification of parameter λ_1 is that $\gamma_2 = 0$ or $\beta_2 = 0$.

 $[\]overline{ \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}(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).$ Note that this expression is equal to $T \sigma_{\alpha 2}(0,0) + T_1^{(1)} \left[\sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,0) \right] + \left[T_2^{(1)} + y_{20} - y_{2T} \right] \left[\sigma_{\alpha 2}(0,1) - \sigma_{\alpha 2}(0,0) \right] + C_{12} \left[\sigma_{\alpha 2}(1,1) - \sigma_{\alpha 2}(0,0) + \sigma_{\alpha 2}(0,0) \right].$

Examples of histories and identified parameters with $T=3$						
$A = \{ \boldsymbol{y}_0 \}$, a, b, z	$oldsymbol{y}_3\}; \hspace{0.2cm} B$:	$= \{ \boldsymbol{y}_0, \ \boldsymbol{b}_0 \}$	$m{a}, m{y}_3\}$	with $C_{12}(A) = C_{12}(B)$	
	$oldsymbol{y}_0$	a	b	$oldsymbol{y}_3$	$\ln \mathbb{P}\left(A\right) - \ln \mathbb{P}\left(B\right)$	
Case 1:	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c}1\\0\end{array}\right)$	$\left(\begin{array}{c}1\\0\end{array}\right)$	β_1	
Case 2:	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\ 0\end{array}\right)$	$\left(\begin{array}{c} 0\\1\end{array}\right)$	$\left(\begin{array}{c} 0\\1\end{array}\right)$	β_2	

 Table 3: Myopic Dynamic Game Without Contemporaneous Effects

Proof of Proposition 3: See Appendix A.3.

Suppose that T = 3 and consider the pair of histories $A = \{y_0, a, b, y_3\}$ and $B = \{y_0, b, a, y_3\}$. In this model with a contemporaneous effect, the sufficient statistic includes C_{12} , so a pair of identifying histories A and B should satisfy the condition $C_{12}(A) = C_{12}(B)$. Table 3 presents examples of pairs of histories A and B which identify parameters β_1 and β_2 as the log-odds ratio $\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$.

Parameter γ_2 appears in the log-probability of a choice history only through the term $T^{(1,1)}$ $(\Delta\sigma_{\alpha_1} + \gamma_2)$. This implies that this parameter is point-identified if and only if $\Delta\sigma_{\alpha_1}$ is equal to zero 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 in turn $\sigma_{\alpha_1}(y_1,y_2)$ is defined as $-\ln[1 + \exp{\{\alpha_1 + \beta_1 \ y_1 + \lambda_1 \ y_2\}}]$. Taking this into account we have that $\Delta\sigma_{\alpha_1} = 0$ for every value of α_1 if and only if $\beta_1 = 0$ or $\lambda_1 = 0$.

Parameter λ_1 appears in the log-probability of a choice history only through the term C_{12} $(\Delta \sigma_{\alpha_2} + \lambda_1)$. This parameter is point-identified using a sufficient statistics approach if and only if $\Delta \sigma_{\alpha_2} = 0$ for every possible value of α_2 . This is the case if and only if $\gamma_2 = 0$ or $\beta_2 = 0$.

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

Proposition 4. Consider the myopic dynamic game with one-direction strategic interactions as described in equation (20) where the fixed effects of the two players are restricted to be the same: $\alpha_{1m} = \alpha_{2m}$. The functional differenting approach implies moment conditions that point identify all the structural parameters, β_1 , β_2 , λ_1 , and γ_2 .

Proof of Proposition 4: See Appendix A.4.

3.1.3 Myopic players, two-direction strategic interactions, sequential move

Consider the game with two-direction contemporaneous interactions such that $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$. We eliminate the lagged strategic interactions between players such that $\lambda_1 = \lambda_2 = 0$.

$$\begin{cases} y_{1t} = 1 \{ \alpha_1 + \gamma_1 \ y_{2t} + \beta_1 \ y_{1t-1} - \varepsilon_{1t} \ge 0 \} \\ y_{2t} = 1 \{ \alpha_2 + \gamma_2 \ y_{1t} + \beta_2 \ y_{2t-1} - \varepsilon_{2t} \ge 0 \} \end{cases}$$
(23)

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

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

$$\{e_{1t}^1 < \varepsilon_{1t} \le e_{1t}^0 \qquad \& \qquad e_{2t}^1 < \varepsilon_{2t} \le e_{2t}^0\}$$
(24)

where $e_{it}^1 \equiv \alpha_i + \gamma_i + \beta_i y_{i,t-1}$, and $e_{it}^0 \equiv \alpha_i + \beta_i y_{i,t-1}$. In the sequential move game, where player 1 moves first, the outcome (1,0) is the unique equilibrium (Subgame Perfect Nash equilibrium)

associated with this region of $(\varepsilon_{1t}, \varepsilon_{2t})$. Player 1 is aware that if she chooses $y_{1t} = 1$, player 2 will choose $y_{2t} = 0$, and if she chooses $y_{1t} = 0$, then player 2 will choose $y_{2t} = 1$. Therefore, player 1's decision determines which of the two Nash equilibria, (0, 1) or (1, 0), is selected. Player 1 selects the equilibrium that maximizes its payoff. Considering $\gamma_1 \leq 0$, the Nash equilibrium with the highest payoff for player 1 is (1, 0).

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

$$\mathbb{P}(y_{1t}, y_{2t}; \boldsymbol{\alpha}) = \Lambda([2y_{1t} - 1][\alpha_1 + \gamma_1 y_{2t} + \beta_1 y_{1t-1}]) \ \Lambda([2y_{2t} - 1][\alpha_2 + \gamma_2 y_{1t} + \beta_2 y_{2t-1}])$$
(25)

In contrast, the probability of the outcome (0, 1) cannot be expressed as a product of logits. This feature affects both the derivation of a sufficient statistic for $\boldsymbol{\alpha}$ and the implementation of a functional differencing approach.

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

$$\ln \mathbb{P}(\widetilde{\mathbf{y}}|\boldsymbol{\alpha},\boldsymbol{\theta}) = \ln p_{\alpha}(y_{10}, y_{20}) + \alpha_1 T_1^{(1)} + \alpha_2 T_2^{(1)} + \sum_{t=1}^T \sigma_{\alpha_1}(y_{1t-1}, y_{2t}) + \sigma_{\alpha_2}(y_{1t}, y_{2t-1}) + \beta_1 C_{11} + \beta_2 C_{22} + (\gamma_1 + \gamma_2) T^{(1,1)}$$
(26)

with $\sigma_{\alpha_i}(y_{i,t-1}, y_{jt}) = -\ln[1 + \exp\{\alpha_i + \beta_i y_{i,t-1} + \gamma_i y_{jt}\}]$. Similarly as for the previous models, we can rewrite the right hand side of equation (26) as $\mathbf{s}(\widetilde{\mathbf{y}})' \mathbf{g}_{\alpha} + \mathbf{c}(\widetilde{\mathbf{y}})' \boldsymbol{\theta}^*$, where now the vectors of statistics and parameters have the following form:

$$\ln \mathbb{P}(\widetilde{\mathbf{y}}|\boldsymbol{\alpha},\boldsymbol{\theta}) = \begin{bmatrix} 1, y_{10}, y_{20}, y_{10}y_{20}, y_{1T}, y_{2T}, y_{1T}y_{2T}, T, T_1^{(1)}, T_2^{(1)}, C_{12}, C_{21} \end{bmatrix}' \mathbf{g}_{\boldsymbol{\alpha}} + \beta_1 C_{11} + \beta_2 C_{22} + (\gamma_1 + \gamma_2) T^{(1,1)}$$
(27)

A preliminary examination of equation (27) might suggest that the parameter $\gamma_1 + \gamma_2$ is

identified, as it appears alongside the statistic $T^{(1,1)}$, which is not included in the vector of sufficient statistics $\mathbf{s}(\tilde{\mathbf{y}})$. However, due to the nature of the choice history $\tilde{\mathbf{y}}$, which does not include any outcome $(y_{1t}, y_{2t}) = (0, 1)$, we have that $(1 - y_{1t})y_{2t} = 0$ for every t, and it follows that $\sum_t y_{2t} = \sum_t y_{1t}y_{2t}$, implying that $T_2^{(1)} = T^{(1,1)}$. Consequently, given $\mathbf{s}(\tilde{\mathbf{y}})$ (which includes $T_2^{(1)}$ as an element), the statistic $T^{(1,1)}$ lacks variation, thereby rendering this approach of using sufficient statistics insufficient for identifying $\gamma_1 + \gamma_2$. Nonetheless, parameters β_1 and β_2 can still be point identified. This assertion is formalized in the following proposition.

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

Proof of Proposition 5. See Appendix A.5. \blacksquare

Partial Identification of γ_1, γ_2

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

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

or

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

and the following logit form upper bound:

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

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

3.1.4 Sharp Identified set

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

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

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

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

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

3.1.5 Myopic players, two-direction strategic interactions, simultaneous move

Here we concentrate on the (point) identification of the switching cost parameters $-\beta_1$ and β_2 - and on the partial identification of all the parameters.

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

Lemma 1. Suppose that the log-probability of a market history has lower and upper bounds

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

$$h(\mathbf{s}_{L}(\widetilde{\mathbf{y}})) + \mathbf{c}_{L}(\widetilde{\mathbf{y}})'\beta \leq \ln \mathbb{P}(\widetilde{\mathbf{y}}) \leq h(\mathbf{s}_{U}(\widetilde{\mathbf{y}})) + \mathbf{c}_{U}(\widetilde{\mathbf{y}})'\boldsymbol{\theta}$$
(28)

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

i. If $\mathbf{s}_{L}(A) = \mathbf{s}_{U}(B)$ and $\mathbf{c}_{L}(A) \neq \mathbf{c}_{U}(B)$, then: $[\mathbf{c}_{L}(A) - \mathbf{c}_{U}(B)]' \boldsymbol{\theta} \leq \ln \mathbb{P}(A) - \ln \mathbb{P}(B)$. *ii.* If $\mathbf{s}_{U}(A) = \mathbf{s}_{L}(B)$ and $\mathbf{c}_{U}(A) \neq \mathbf{c}_{L}(B)$, then: $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \leq [\mathbf{c}_{U}(A) - \mathbf{c}_{L}(B)]' \boldsymbol{\theta}$.

These inequalities imply partial identification of some structural parameters.

Proof of Lemma 1: See Appendix A.6.

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

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

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

$$\begin{cases} \ln \mathbb{P}_{L} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) \equiv \ln p_{\boldsymbol{\alpha}} \left(y_{10}, y_{20} \right) + \sum_{t=1}^{T} \ln L(\mathbf{y}_{t} | \mathbf{y}_{t-1}; \boldsymbol{\alpha}, \boldsymbol{\theta}) \\ \ln \mathbb{P}_{U} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) \equiv \ln p_{\boldsymbol{\alpha}} \left(y_{10}, y_{20} \right) + \sum_{t=1}^{T} \ln U(\mathbf{y}_{t} | \mathbf{y}_{t-1}; \boldsymbol{\alpha}, \boldsymbol{\theta}) \end{cases}$$
(29)

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

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

Lemma 2. For the myopic complete information dynamic game with contemporaneous effects in equation (23), the log-probability of a market history has lower bounds $\ln \mathbb{P}_{L\{E,W\}}(\widetilde{\mathbf{y}} \mid \alpha, \theta)$ and $\ln \mathbb{P}_{L\{S,N\}}(\widetilde{\mathbf{y}} \mid \alpha, \theta)$ and upper bound $\ln \mathbb{P}_{U}(\widetilde{\mathbf{y}} \mid \alpha, \theta)$ which have the following expressions:

$$\ln \mathbb{P}_{L\{E,W\}} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) = \mathbf{s}^{1} \left(\widetilde{\mathbf{y}} \right)' \mathbf{g}_{\boldsymbol{\alpha}}^{1} + \left[T_{1}^{(1)}, T_{1}^{(1)}, C_{11}, C_{12} \right] \mathbf{g}_{\boldsymbol{\alpha}}^{2} + C_{11} \beta_{1} + C_{22} \beta_{2} + T_{1}^{(1)} \gamma_{1} + T^{(1,1)} \gamma_{2} \ln \mathbb{P}_{L\{S,N\}} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) = \mathbf{s}^{1} \left(\widetilde{\mathbf{y}} \right)' \mathbf{g}_{\boldsymbol{\alpha}}^{1} + \left[T_{2}^{(1)}, T_{2}^{(1)}, C_{21}, C_{22} \right] \mathbf{g}_{\boldsymbol{\alpha}}^{2} + C_{11} \beta_{1} + C_{22} \beta_{2} + T^{(1,1)} \gamma_{1} + T_{2}^{(1)} \gamma_{2} \ln \mathbb{P}_{U} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) = \mathbf{s}^{1} \left(\widetilde{\mathbf{y}} \right)' \mathbf{g}_{\boldsymbol{\alpha}}^{1} + \left[T_{2}^{(1)}, T_{1}^{(1)}, C_{21}, C_{12} \right] \mathbf{g}_{\boldsymbol{\alpha}}^{2} + C_{11} \beta_{1} + C_{22} \beta_{2} + T^{(1,1)} \left[\gamma_{1} + \gamma_{2} \right]$$
(30)

where \mathbf{g}_{α}^{1} and \mathbf{g}_{α}^{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 $T_{2}^{(1)}$.

Proof of Lemma 2: See Appendix A.7.

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

Proposition 6. Consider the myopic complete information dynamic game with contemporaneous effects in equation (23). Define the vector of statistics $\mathbf{s}^1(\widetilde{\mathbf{y}}) \equiv [T, y_{10}, y_{20}, y_{1T}, y_{2T}, T_1^{(1)}, T_2^{(1)}]$. Let A and B be two market histories such that $\mathbf{s}^1(A) = \mathbf{s}^1(B)$ and $T_1^{(1)} = T_2^{(1)}$. Let $\Delta(A, B, \beta_1, \beta_2)$ be $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) - [C_{11}(A) - C_{11}(B)] \beta_1 - [C_{22}(A) - C_{22}(B)] \beta_2$.

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

$$\Delta(A, B, \beta_1, \beta_2) \geq \left[T_1^{(1)}(A) - T^{(1,1)}(B) \right] \gamma_1 + \left[T^{(1,1)}(A) - T^{(1,1)}(B) \right] \gamma_2$$
(31)

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

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

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

$$\Delta(A, B, \beta_1, \beta_2) \geq \left[T^{(1,1)}(A) - T^{(1,1)}(B) \right] \gamma_1 + \left[T_2^{(1)}(A) - T^{(1,1)}(B) \right] \gamma_2$$
(33)

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

$$\Delta(A, B, \beta_1, \beta_2) \leq \left[T^{(1,1)}(A) - T^{(1,1)}(B) \right] \gamma_1 + \left[T^{(1,1)}(A) - T^{(1)}_2(B) \right] \gamma_2$$
(34)

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

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

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

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

 $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \ge \beta_{22} + \gamma_1 + \gamma_2 \tag{35}$

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

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

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

$$\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \ge \beta_{22} + \gamma_1 + \gamma_2 \tag{37}$$

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 $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \ge \beta_1 + \gamma_1 + \gamma_2$ and $\ln \mathbb{P}(A) - \ln \mathbb{P}(B) \le \beta_1 + \gamma_2$

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

condition (iii) and (iv) are satisfied and imply informative bounds on the parameters.

3.2 Forward-looking players

3.2.1 Forward-looking players with one-direction strategic interactions

Consider the complete information game in equation (6). It is convenient to represent this model as follows:

$$y_{it} = 1 \{ \widetilde{\alpha}_i + \beta_i \ y_{i,t-1} + \widetilde{\gamma}_{i\alpha} \ y_{jt} - \varepsilon_{it} \ge 0 \}$$

$$(38)$$

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

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

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

$$y_{1t} = 1 \left\{ \varepsilon_{1t} \leq \alpha_1 + \beta_1 y_{1t-1} + \widetilde{v}_{1\alpha} \right\}$$
(39)

This identification of this forward-looking dynamic logit model – with fixed effects unobserved heterogeneity – has been established in Aguirregabiria, Gu, and Luo (2021). In this model: the incidental parameter is $\alpha_1 + \tilde{v}_{1\alpha}$; the vector of sufficient statistics is $\mathbf{s}(\tilde{\mathbf{y}}) = [y_{10}, y_{1T}, T_1^{(1)}]$; and the structural parameter β_1 is identified from the maximization of the conditional likelihood function. We now establish the point identification of parameters β_1 and β_2 . The best response of player 2 in this triangular model is:

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

$$(40)$$

where $\tilde{\alpha}_2 \equiv \alpha_2 + \tilde{v}_{2\alpha}(0)$, and $\tilde{\gamma}_{2\alpha} \equiv \gamma_2 + \tilde{v}_{2\alpha}(1) - \tilde{v}_{2\alpha}(0)$. Given equations (39) and (40), the log-probability of the market history $\tilde{\mathbf{y}} \equiv (y_{1t}, y_{2t} : t = 0, 1, .., T)$ has the following structure:

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}}|\boldsymbol{\alpha},\beta\right) = \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) + \alpha_{1} T_{1}^{(1)} + \alpha_{2} T_{2}^{(1)} + \widetilde{\gamma}_{2\alpha} T^{(1,1)} + \sum_{t=1}^{T} \sigma_{\boldsymbol{\alpha}1}\left(y_{1t-1}\right) + \sigma_{\boldsymbol{\alpha}2}\left(y_{1t}, y_{2t-1}\right) + \beta_{1} C_{11} + \beta_{2} C_{22}$$

$$(41)$$

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

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

$$(42)$$

Proposition 7. For the forward-looking dynamic game with one-direction strategic interactions as described in equations (39) and (40):

- A. Vector $\mathbf{s}(\widetilde{\mathbf{y}}) = [1, y_{10}, y_{20}, y_{10}y_{20}, y_{1T}, y_{2T}, y_{1T}y_{2T}, T, T_1^{(1)}, T_2^{(1)}, T_1^{(1,1)}, C_{12}]'$ is a minimal sufficient statistic for $\boldsymbol{\alpha}$ such that $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{u}(\widetilde{\mathbf{y}}), \alpha, \boldsymbol{\theta})$ does not depend on $\boldsymbol{\alpha}$.
- B. $\ln \mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{c}(\widetilde{\mathbf{y}}), \beta) = \mathbf{s}(\widetilde{\mathbf{y}})' \boldsymbol{\theta}^* \ln(\sum_{\widetilde{\mathbf{y}}':\mathbf{s}(\widetilde{\mathbf{y}}')=\mathbf{s}(\widetilde{\mathbf{y}})} \exp{\{\mathbf{c}(\widetilde{\mathbf{y}}')'\boldsymbol{\theta}^*\}})$ with $\mathbf{c}(\widetilde{\mathbf{y}}) = [C_{11}, C_{22}]'$ and $\boldsymbol{\theta}^* = [\beta_1, \beta_2]'.$
- C. For $T \ge 3$, 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 $\boldsymbol{\theta}^*$.

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

3.2.2 Forward-looking players with two-direction strategic interactions

Consider he forward-looking dynamic game where we do not restrict any γ parameter to be zero.

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

$$(43)$$

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 $(\tilde{\gamma}_{1\alpha}, \tilde{\gamma}_{2\alpha})$ 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 $(\tilde{\gamma}_{1\alpha}, \tilde{\gamma}_{2\alpha})$. Though this different is coneptually simple, it has substantial implications on the identification of the γ parameters. More specifically, we cannot point identify the switching cost parameters. Proposition 8 establishes that these parameters are partially identified.

Proposition 8. Consider the forward-looking complete information dynamic game with contemporaneous effects in equation (43). Under conditions $\beta_1 \ge 0$, $\beta_2 \ge 0$, $\tilde{\gamma}_{1\alpha} \le 0$, and $\tilde{\gamma}_{2\alpha} \le 0$, there are market histories that provide informative bounds on the parameters β_1 and β_2 . These parameters are partially identified.

Proof of Proposition 8: See Appendix A.8.

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

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

4 Empirical application

4.1 Framework and model

Let firms i = 1, 2, ..., N represent the manufacturers in an industry, each producing a single product. These firms compete by setting prices across M distinct geographic markets, indexed by m. Let p_{imt} denote the price set by firm i in market m at week t, and let p_{mt} represent the vector of prices across all firms in market m at time t. Each firm selects prices to maximize its expected intertemporal profits.

A firm's profit in a given market consists of two components: its variable profit function, $\pi_{imt}(\boldsymbol{p}_{mt})$, which represents revenue minus variable costs and depends on the prices set by all firms; and its price adjustment cost, represented by $ac_{imt}(p_{imt}, p_{im,t-1})$, which depends on the firm's current and previous period prices. Price adjustment costs encompass a broad range of expenses, including the costs of acquiring and processing information needed to set prices, attention costs, reputational costs, the expenses involved in communicating new prices to consumers, and fixed costs related to inventory and promotional efforts associated with a price change.

The magnitude of price adjustment costs significantly influences the speed at which inflationary shocks are passed through to prices (Rotemberg, 1982; Nakamura and Steinsson, 2008) and affects the welfare cost of inflation. Additionally, oligopolistic competition and its interaction with price adjustment costs can play a substantial role in amplifying price stickiness (Wang and Werning, 2022; Mongey, 2021).

In many industries, including the one relevant to this application, firms' pricing choices alternate between two levels: a regular (or high) price and a promotional (or low) price. While these high and low price levels can vary among firms, they remain stable over extended periods—typically months or even years. Given these established price levels, firms compete by deciding whether to set the high or low price at any given time. Our focus is on this high-low pricing competition, treating each firm's high and low price levels as predetermined.

Let $y_{imt} \in \{0, 1\}$ denote an indicator of whether a firm chooses a low or promotional price in market m at time t. Likewise, let \boldsymbol{y}_{mt} denote the vector of these indicators across all firms in market m at time t. For convenience, we keep using $\pi_{imt}(\boldsymbol{y}_{mt})$ and $ac_{imt}(y_{imt}, y_{im,t-1})$ to denote the variable profit and price adjustment cost functions, respectively.

In our application, we focus on a duopoly industry, N = 2. Consistent with our notation in previous sections, \tilde{u}_{imt} denotes the difference in firm *i*'s payoff between choosing y = 1 and y = 0—- specifically, the differential profit from setting a promotion price. According to discussion above:

$$\widetilde{u}_{imt} = \pi_{imt}(1, y_{jmt}) - \pi_{imt}(0, y_{jmt}) + ac_{imt}(1, y_{im,t-1}) - ac_{imt}(0, y_{im,t-1}) = \alpha_{imt} + \gamma_{imt} y_{jmt} + \beta_{imt} y_{im,t-1},$$
(44)

with:

$$\alpha_{imt} \equiv \pi_{imt}(1,0) - \pi_{imt}(0,0) + ac_{imt}(1,0) - ac_{imt}(0,0)$$

$$\gamma_{imt} \equiv \pi_{imt}(1,1) - \pi_{imt}(0,1) - \pi_{imt}(1,0) + \pi_{imt}(0,0)$$

$$\beta_{imt} \equiv ac_{imt}(1,1) - ac_{imt}(0,1) - ac_{imt}(1,0) + ac_{imt}(0,0)$$
(45)

Here, α_{imt} represents the profit difference when the competitor opts for a high price, γ_{imt} captures the complementarity, or supermodularity, of the variable profit function with respect to the prices of the two firms, and β_{imt} reflects the supermodularity of the firm's price adjustment cost function with respect to the its prices at two consecutive periods.

For identifying the adjustment cost parameters β , our econometric model imposes the restrictions $\alpha_{imt} = \alpha_{im} - \varepsilon_{imt}$, $\gamma_{imt} = \gamma_{im}$, and $\beta_{imt} = \beta_i$. This means that demand and marginal cost functions must remain stable over time, except for an i.i.d. additive shock ε_{imt} , and that the price adjustment cost function should remain stable over time and across markets. Naturally, the assumption of time-invariance for demand and marginal costs is more plausible with shorter panel durations.

In this application, the panel data spans more than two years (128 weeks) with weekly observations. To enhance the plausibility of the time-invariance condition, we partition this long time series into shorter subperiods, each lasting 8 weeks. Thus, our time-invariance restriction requires that demand and marginal costs remain constant (up to the i.i.d. shock) within each 8-week period.

4.2 Data and Estimates

The dataset comes from A.C. Nielsen scanner panel data for the ketchup product category in the geographic market of Sioux Falls, South Dakota. It contains price data from 45 participating stores and covers a 128-week period from mid-1986 to mid-1988.⁴ For our analysis, a period is a week. As mentioned above, we partition the price time of each store into shorter subperiods, each lasting 8 weeks, and treat each subperiod-store as a different market. The number of markets is M = 45 * (128/8) = 720.

There are four brands in this market: three national brands, Heinz, Hunt's, and Del Monte; and a store brand. Here we focus on competition between the two leading brands, Heinz and Hunt's, which account for 69% and 18% of the market, respectively.

Table 4 presents estimates of the price adjustment cost parameters, β_{Heinz} and β_{Hunts} , for the model where firms are forward-looking and the strategic interaction parameters are not restricted

⁴Our sample comes from Erdem, Imai, and Keane (Erdem, Imai, and Keane). We thank the authors for sharing the data with us.

Parameter	\mathbf{FE} –	SS	No Ur	nob. Het.	Dummy V	/ariables
β_{Heinz}	[0.3253],	0.5215]	[0.8145	, 0.8806]	[0.6412 ,	0.7074]
β_{Hunts}	[0.2628 ,	0.4571]	[0.7543	, 0.8173]	[0.6167 ,	0.6658]

Table 4: Estimates of Price Adjustment Cost Parameters95% Confidence Intervals

to be zero. In this version of the model, and with Fixed Effects unobserved heterogeneity, the β parameters are partially identified but not point identified.

We present three sets of 95% confidence interval estimates for the parameters: (1) our sufficient statistics-based bounds method; (2) a bounds method that assumes no time-invariant unobserved heterogeneity; and (3) a bounds method that accounts for time-invariant unobserved heterogeneity by including market dummies (i.e., using a dummy variable estimator). Confidence intervals are computed using the bootstrap inference method of Cox and Shi (2023).

The estimates indicate that failing to account for persistent unobserved heterogeneity, or attempting to control for it using the inconsistent dummy variable approach, leads to a substantial overestimation of price adjustment costs. Our Fixed Effect-Sufficient Statistics estimates reveal that price adjustments costs are statistically significant. There is also a small but statistically insignificant difference in the magnitude of price adjustment costs between the two firms.

5 Conclusions

This paper provides a comprehensive study on the identification of dynamic games with fixed effects, introducing a flexible framework that accommodates unobserved heterogeneity across markets and players without imposing parametric or support restrictions. The model accounts for forward-looking players and multiple equilibria, offering a more robust approach for empirical researchers studying markets with intertemporal dependencies and strategic decision-making.

A key contribution of this paper is the development of an identification method that extends

the Conditional Likelihood-Sufficient Statistics approach to models involving multiple equilibria and partial identification. We demonstrate that there exist sufficient statistics for the lower and upper bounds of market histories and that they can be combined to achieve partial identification of structural parameters, thus broadening the applicability of these techniques in dynamic game settings.

Furthermore, the paper advances the use of functional differencing methods in dynamic structural models. For certain versions of the model, this approach improves point identification of structural parameters that remain unidentified under the conditional likelihood approach alone.

An important outcome of these constructive identification results is the implication for estimation methods. Specifically, these results facilitate estimation techniques that avoid the complex solution of the dynamic game or the computation of present values in players' dynamic programming problems. In these fixed effects models, controlling for unobserved heterogeneity involves differencing out the continuation value component in the players' intertemporal payoffs. As a result, estimating dynamic games using these fixed effects methods becomes computationally comparable to estimating a single-agent static fixed effect discrete choice model.

The empirical application of these methods underscores their practical relevance in real-world market contexts. By analyzing panel data on supermarket prices that follow a High-Low pricing pattern, we illustrate how the identification results can be applied effectively to estimate firms' price adjustment costs. This practical demonstration strengthens the theoretical contributions and provides clear guidelines for future empirical work in dynamic market settings.

A APPENDIX

A.1 Proof of Proposition 1

(A) Sufficient condition. Multiplying equation (14) times $p^*(\boldsymbol{\alpha}|\boldsymbol{y}_0) \mathbb{P}(\boldsymbol{y}_0)$, integrating over $\boldsymbol{\alpha}$, and taking into account that $\int \mathbb{P}(\widetilde{\boldsymbol{y}}|\boldsymbol{y}_0, \boldsymbol{\alpha}, \boldsymbol{\theta}) p^*(\boldsymbol{\alpha}|\boldsymbol{y}_0) d\boldsymbol{\alpha}$ is equal to $\mathbb{P}(\widetilde{\boldsymbol{y}}|\boldsymbol{y}_0, \boldsymbol{\theta})$, we obtain:

$$\sum_{\boldsymbol{y}_{0},\widetilde{\boldsymbol{y}}} \boldsymbol{w} \left(\boldsymbol{y}_{0}, \widetilde{\boldsymbol{y}}, \boldsymbol{\theta} \right) \ \mathbb{P}(\boldsymbol{y}_{0}, \widetilde{\boldsymbol{y}} \mid \boldsymbol{\theta}) = \boldsymbol{0}$$
(46)

Evaluated at the true value of $\boldsymbol{\theta}$, the model probability $\mathbb{P}(\boldsymbol{y}_0, \widetilde{\boldsymbol{y}} \mid \boldsymbol{\theta})$ becomes the data frequency $\mathbb{P}(\boldsymbol{y}_0, \widetilde{\boldsymbol{y}})$. Thus, condition (ii) in Proposition 1 implies the point identification of $\boldsymbol{\theta}$.

(B) Necessary condition. The frequencies $\mathbb{P}(\boldsymbol{y}_0, \boldsymbol{\widetilde{y}})$ contain all the information in the data about the vector of structural parameters. Given that we do not impose any restriction on the weights $\boldsymbol{w}(\boldsymbol{y}_0, \boldsymbol{\widetilde{y}}, \boldsymbol{\theta})$, condition (ii) in Proposition 1 describes, w.l.o.g., necessary conditions for the identification of $\boldsymbol{\theta}$. Therefore, we need to establish that condition (i) is necessary to obtain the system of equations (15) in condition (ii).

We need to prove that, if equation $\sum_{\widetilde{y}} w(y_0, \widetilde{y}, \theta) \mathbb{P}(\widetilde{y} \mid y_0) = 0$ holds, then equation (14) should hold for every value α . The proof is by contradiction.

Suppose that:

- a. Equation $\sum_{\widetilde{\boldsymbol{y}}} \boldsymbol{w}(\boldsymbol{y}_0, \widetilde{\boldsymbol{y}}, \boldsymbol{\theta}) \mathbb{P}(\widetilde{\boldsymbol{y}} \mid \boldsymbol{y}_0) = \boldsymbol{0}$ holds for any distribution $p^*(\boldsymbol{\alpha} \mid \boldsymbol{y}_0)$ in the DGP.
- b. There is a value $\boldsymbol{\alpha} = c$ and a value of \boldsymbol{y}_0 such that equation (15) does not hold:

$$\sum_{\widetilde{\mathbf{y}}} \boldsymbol{w}(\boldsymbol{y}_0, \widetilde{\mathbf{y}}, \boldsymbol{\theta}) \mathbb{P}(\widetilde{\mathbf{y}} \mid \boldsymbol{y}_0, c) \neq \mathbf{0}$$

We show below that condition (b) implies that there is a density function $p^*(\boldsymbol{\alpha}|\boldsymbol{y}_0)$ (in fact, a continuum of density functions) such that condition (a) does not hold.

W.l.o.g., consider distributions of $\boldsymbol{\alpha}$ with only two points support, c and c' with $p^*(c|\boldsymbol{y}_0) = q$. Define the following function $d(\boldsymbol{y}_0, \boldsymbol{\alpha})$ that measures the extent in which equation (15) is not satisfied:

$$d(\boldsymbol{y}_0, \boldsymbol{\alpha}) \equiv \sum_{\widetilde{\boldsymbol{y}}} \boldsymbol{w}(\boldsymbol{y}_0, \widetilde{\boldsymbol{y}}) \mathbb{P}(\widetilde{\boldsymbol{y}} \mid \boldsymbol{y}_0, \boldsymbol{\alpha})$$
(47)

Condition (b) implies that $d(\mathbf{y}_0, c) \neq \mathbf{0}$. For notational simplicity but w.l.o.g., consider that the initial condition \mathbf{y}_0 has binary support $\{0, 1\}$. By definition, we have that:

$$\sum_{\widetilde{\mathbf{y}}} \boldsymbol{w}(\boldsymbol{y}_0, \widetilde{\mathbf{y}}) \mathbb{P}(\widetilde{\mathbf{y}} \mid \boldsymbol{y}_0) = q \ d(\boldsymbol{y}_0, c) + (1-q) \ d(\boldsymbol{y}_0, c')$$

By definition, each value $d(\boldsymbol{y}_0, \boldsymbol{\alpha})$ is for a particular value of $\boldsymbol{\alpha}$, and therefore, it does not depend on the distribution of $\boldsymbol{\alpha}$. More specifically, $d(\boldsymbol{y}_0, c)$ and $d(\boldsymbol{y}_0, c')$ do not depend on the value of q. Therefore, there always exists (a continuum of) values of q such that the right-hand side of (A.1) is different from zero, and condition (a) does not hold.

A.2 Proof of Proposition 2

Given that, $\mathbb{P}(\widetilde{\mathbf{y}} \mid \mathbf{s}(\widetilde{\mathbf{y}}), \boldsymbol{\alpha}, \boldsymbol{\theta}) = \mathbb{P}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}) / \mathbb{P}(\mathbf{s}(\widetilde{\mathbf{y}}) \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$, and using the structure of $\mathbb{P}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$ in equation (13), we have that:

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \mathbf{s}\left(\widetilde{\mathbf{y}}\right), \beta\right) = \mathbf{c}\left(\widetilde{\mathbf{y}}\right)' \boldsymbol{\theta} - \ln \left(\sum_{\widetilde{\mathbf{y}}': \mathbf{s}(\widetilde{\mathbf{y}}') = \mathbf{s}(\widetilde{\mathbf{y}})} \exp\left\{\mathbf{c}\left(\widetilde{\mathbf{y}}'\right)' \boldsymbol{\theta}\right)\right)$$
(48)

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

Suppose that T = 3, let $\boldsymbol{y}_t \equiv (y_{1t}, y_{2t})$, and consider the following pair of histories: $A = \{\boldsymbol{y}_0, \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{y}_3\}$ and $B = \{\boldsymbol{y}_0, \boldsymbol{b}, \boldsymbol{a}, \boldsymbol{y}_3\}$. We first verify that histories A and B have the same sufficient statistic \mathbf{s} . It is clear that the two histories have the same initial condition \boldsymbol{y}_0 , and last

period choices, y_3 . And it is also clear that the frequency of choices in $\{a, b, y_3\}$ is the same as in $\{b, a, y_3\}$ 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 $\mathbf{a} \neq \mathbf{b}$ 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,

$$C_{11}(A) - C_{11}(B) = (a_1 - b_1) (y_{10} - y_{13})$$

$$C_{12}(A) - C_{12}(B) = (a_1 - b_1)y_{20} - (a_2 - b_2)y_{13} + a_2b_1 - a_1b_2$$

$$C_{21}(A) - C_{21}(B) = (a_2 - b_2)y_{10} - (a_1 - b_1)y_{23} + a_1b_2 - a_2b_1$$

$$C_{22}(A) - C_{22}(B) = (a_2 - b_2) (y_{20} - y_{23})$$
(49)

Using the expressions in (49), Table 2 presents four pairs of histories, with each pair identifying one of the structural parameters. The corresponding parameter that is identified by $\ln \mathbb{P}(A) - \ln \mathbb{P}(B)$. In cases 1 and 2, we identify the parameter β_i by keeping constant the choice of the other player $-j \neq i$ – and comparing the frequency of the history where player i "switches" -(0,1,0,1) – with the frequency of the history where she "stays" – (0,0,1,1). In cases 3 and 4, we compare the probability of history (0,0,0,1) for player i when the other player chooses alternative 1 at period t = 2 - (0,0,1,0) – and when this choice is at period t = 1 - (0,1,0,0). There are other values for y_0 , a, b, and y_3 that identify linear combinations of the several parameters in θ .

A.3 Proof of Proposition 3

Consider the same framework as in the proof of Proposition 1: T = 3 and the pair of histories $A = \{y_0, a, b, y_3\}$ and $B = \{y_0, b, a, y_3\}$. In the proof of Proposition 1, we showed that these histories have the same value for the statistics y_0 , y_3 , and $T^{(y_1,y_2)}$. Now, in this model with a contemporaneous effect, the sufficient statistic includes C_{12} , so we need to impose additional conditions on histories A and B such that $C_{12}(A) = C_{12}(B)$. In the histories in Table 2, we have that $C_{12}(A) = C_{12}(B)$ for cases 1 and 2. Therefore, these two pairs of market histories still

identify the parameters β_1 and β_2 , respectively, in this dynamic game. We present this result in Table 3.

A.4 Proof of Proposition 4

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

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

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

Define $g(\alpha, \theta)$ as the minimum common denominator (MCD) of all the ratios in the table

above. It is simple to verify that this MCD has the following expression.

$$g(\alpha, \theta) \equiv (1+A)^4 \ (1+AC)^2 \ (1+AB_{11}) \ (1+AB_{12}) \ (1+AB_{22}) \ (1+AB_{11}B_{12}) \ (1+ACB_{22}) \ (50)$$

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

$$P(\widetilde{\mathbf{y}} \mid \boldsymbol{\theta}, y_0) = \int P(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}, y_0) f(\alpha \mid y_0) d\alpha = \int P(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}, y_0) g(\alpha, \boldsymbol{\theta}) q(\alpha \mid \boldsymbol{\theta}, y_0) d\alpha$$
(51)

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

Given equation (51) and after some calculations, we can write the following system of 16 equations relating probabilities of choice histories with the vector of parameters $\boldsymbol{\theta}$ and moments in the distribution of A. Using matrix notation, this system is:

$$\mathbf{P}_{\widetilde{\mathbf{y}}} = \mathbf{G}(\boldsymbol{\theta}) \mathbf{m}_A \tag{52}$$

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

$$\mathbf{m}_A \equiv \int \left(1 \quad A \quad A^2 \quad \dots \quad A^{11} \right)' q(\alpha \mid \boldsymbol{\theta}) \ d\alpha$$
(53)

Given the system of equations in (52), we can construct a moment condition for θ by finding

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

In our specific case here with T = 2 and $(y_{10}, y_{20}) = (0, 0)$, the rank of $\mathbf{G}(\boldsymbol{\theta})$ is 4, hence the dimension of the left null space of $\mathbf{G}(\boldsymbol{\theta})$ is 4, and we can find 4 linearly independent moment conditions for $\boldsymbol{\theta}$. In particular, they take the form

$$-C P_{(1,0),(0,0)} - B_{11} C P_{(1,0),(0,1)} + C P_{(0,0),(1,0)} + P_{(0,0),(1,1)} = 0$$
(54)

$$\frac{B_{22} - 1}{B_{22} - C} \left(P_{(1,0),(0,0)} - P_{(0,1),(0,0)} \right) - \frac{B_{22}}{C} P_{(1,1),(0,0)} - \frac{B_{12}(B_{22} - 1)}{B_{22}(B_{22} - C)} P_{(0,1),(0,1)}
- \frac{B_{11} - B_{22} + C - B_{11}C}{B_{22} - C} P_{(1,0)(0,1)} - \frac{B_{11}B_{12}}{C} P_{(1,1),(0,1)} + P_{(1,0),(1,0)} + P_{(1,0),(1,1)} = 0$$

$$\frac{B_{22}(C - 1)}{B_{22} - C} \left(P_{(1,0),(0,0)} - P_{(0,1),(0,0)} \right) - B_{22} P_{(1,1),(0,0)}
+ \frac{B_{12}C - B_{22}C + B_{22}^2 - B_{12}B_{22}C}{B_{22} - C} P_{(0,1),(0,1)}
+ \frac{B_{11}B_{22}(C - 1)}{B_{22} - C} P_{(1,0),(0,1)} - B_{11}B_{12} P_{(1,1),(0,1)} + P_{(0,1),(1,0)} + P_{(0,1),(1,1)} = 0$$
(56)

and

$$\frac{B_{11}B_{12}(C-1)^2}{C(B_{22}-C)(B_{11}B_{12}-B_{22}C)(B_{22}-1)}(P_{(1,0),(0,0)}-P_{(0,1),(0,0)}) - \frac{B_{11}B_{12}(C-1)}{B_{22}C^2-B_{11}B_{12}C}P_{(1,1),(0,0)}$$

$$-\frac{B_{11}B_{12}(B_{12}C - B_{22}C - B_{22}^3C + B_{22}^2 + B_{22}^2C^2 - 2B_{12}B_{22}C + B_{12}B_{22}^2C)}{B_{22}^2C(B_{22} - C)(B_{11}B_{12} - B_{22}C)(B_{22} - 1)}P_{(0,1),(0,1)}$$

+
$$\frac{B_{11}^2 B_{12} (C-1)^2}{C(B_{22}-C)(B_{11}B_{12}-B_{22}C)(B_{22}-1)} P_{(1,0),(0,1)}$$

$$-\frac{B_{11}B_{12}}{B_{22}C}P_{(1,1),(0,1)} + \frac{B_{11}(B_{22}C - 1)(B_{12} - B_{22}C)}{B_{22}C(B_{11}B_{12} - B_{22}C)(B_{22} - 1)}P_{(0,1),(1,0)} + P_{(1,1),(1,0)} = 0$$
(57)

For general T > 2, then if we keep $(y_{10}, y_{20}) = (0, 0)$, the matrix $G(\theta)$ is going to be of dimension $2^T \times (8T - 4)$, which implies we will have $2^T - (8T - 4)$ number of moment conditions to overidentify the parameter θ .

A.5 Proof of Proposition 5

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

 $C_{11}(B) = 0$. This indicates that statistic $\ln P(A) - \ln P(B)$ identifies β_1 .

To identify β_2 , let's consider two choice histories: $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$. In these histories, the values of the sufficient statistics y_0 , y_3 , $T_1^{(1)}$, $T_2^{(1)}$, C_{21} , and C_{12} are iden-

tical, as well as the value of C_{11} . However, it is noteworthy that $C_{22}(A) = 0$ while $C_{22}(B) = 1$. Consequently, statistic $\ln P(A) - \ln P(B)$ identifies β_2 .

A.6 Proof of Lemma 1

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

$$\exp\left\{\mathbf{s}_{L}'\mathbf{g}_{\alpha} + \mathbf{c}_{L}'\boldsymbol{\theta}\right\} \leq \mathbb{P}\left(\mathbf{\widetilde{y}} \mid \boldsymbol{\alpha}\right) \leq \exp\left\{\mathbf{s}_{U}'\mathbf{g}_{\alpha} + \mathbf{c}_{U}'\boldsymbol{\theta}\right\}$$
(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:

$$\left[\int \exp\left\{\mathbf{s}_{L}^{\prime}\mathbf{g}_{\alpha}\right\} f(\alpha) \ d\alpha\right] \ \exp\left\{\mathbf{c}_{L}^{\prime}\boldsymbol{\theta}\right\} \leq \mathbb{P}\left(\widetilde{\mathbf{y}}\right) \leq \left[\int \exp\left\{\mathbf{s}_{U}^{\prime}\mathbf{g}_{\alpha}\right\} f(\alpha) \ d\alpha\right] \ \exp\left\{\mathbf{c}_{U}^{\prime}\boldsymbol{\theta}\right\} \quad (A.1.2)$$

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

$$h(\mathbf{s}_{L}) + \mathbf{c}_{L}^{\prime}\beta \leq \ln \mathbb{P}(\widetilde{\mathbf{y}}) \leq h(\mathbf{s}_{U}) + \mathbf{c}_{U}^{\prime}\boldsymbol{\theta}$$
(A.1.3)

A.7 Proof of Lemma 2

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

$$\mathbf{s}^{1}(\widetilde{\mathbf{y}})' \mathbf{g}_{\alpha}^{1} \equiv \ln p_{\alpha} (y_{10}, y_{20}) + T [\sigma_{\alpha 1}(0, 0) + \sigma_{\alpha 2}(0, 0)] + (y_{10} - y_{1T}) \Delta \sigma_{\alpha 1}(1, 0) + (y_{20} - y_{2T}) \Delta \sigma_{\alpha 2}(0, 1) + T_{1}^{(1)} [\alpha_{1} + \Delta \sigma_{\alpha 1}(0, 1)] + T_{2}^{(1)} [\alpha_{2} + \Delta \sigma_{\alpha 2}(0, 1)]$$
(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:

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

where $\Delta \sigma_{\alpha 1}(0,1) \equiv \sigma_{\alpha 1}(0,1) - \sigma_{\alpha 1}(0,0); \ \Delta \sigma_{\alpha 2}(1,0) \equiv \sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,0); \ \Delta^2 \sigma_{\alpha 1} \equiv \sigma_{\alpha 1}(1,1) - \sigma_{\alpha 1}(1,0) - \sigma_{\alpha 1}(0,1) + \sigma_{\alpha 1}(0,0); \ \text{and} \ \Delta^2 \sigma_{\alpha 2} \equiv \sigma_{\alpha 2}(1,1) - \sigma_{\alpha 2}(1,0) - \sigma_{\alpha 2}(0,1) + \sigma_{\alpha 2}(0,0).$ And the statistics $R_1^{(1,1)} \equiv \sum_{t=1}^T y_{1t-1} \ y_{1t} \ y_{2t} \ \text{and} \ R_2^{(1,1)} \equiv \sum_{t=1}^T y_{2t-1} \ y_{1t} \ y_{2t}.$

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

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \left(\ln \left[1 - \Lambda\left(\alpha_{1} + \beta_{1} \; y_{1t-1}\right)\right] + \ln \left[1 - \Lambda\left(\alpha_{2} + \beta_{22} \; y_{2t-1}\right)\right]\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \left(\ln \left[1 - \Lambda\left(\alpha_{1} + \beta_{1} \; y_{1t-1}\right)\right] + \ln \Lambda\left(\alpha_{2} + \beta_{2} \; y_{2t-1}\right)\right) \\
+ \sum_{t=1}^{T} y_{1t} \left(1 - y_{2t}\right) \left(\ln \Lambda\left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1}\right) + \ln \left[1 - \Lambda\left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2}\right)\right]\right) \\
+ \sum_{t=1}^{T} y_{1t} y_{2t} \left(\ln \Lambda\left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1}\right) + \ln \Lambda\left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2}\right)\right) \\$$
(A.3)

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

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})\right] \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)y_{2t} \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_{2} + \beta_{2} y_{2t-1}\right] \\
+ \sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_{1} + \beta_{1} y_{1t-1} + \gamma_{1}\right] \\
+ \sum_{t=1}^{T} y_{1t}y_{2t} \left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_{1} + \beta_{1} y_{1t-1} + \gamma_{1} + \alpha_{2} + \beta_{2} y_{2t-1} + \gamma_{2}\right] \\$$
(A.4)

Grouping terms, we have:

$$\ln \mathbb{P} \left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta} \right) \geq \ln p_{\alpha} \left(y_{10}, y_{20} \right) \\ + \sum_{t=1}^{T} \left(1 - y_{1t} \right) \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) \right] \\ + \sum_{t=1}^{T} y_{2t} \left[\alpha_{2} + \beta_{2} y_{2t-1} \right] \\ + \sum_{t=1}^{T} y_{1t} \left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_{1} + \beta_{1} y_{1t-1} + \gamma_{1} \right] \\ + \sum_{t=1}^{T} y_{1t} y_{2t} \left[\gamma_{2} \right]$$
(A.5)

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

 $\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln \mathbb{P}_{L\{E,W\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$

$$= \ln p_{\alpha} (y_{10}, y_{20}) + T [\sigma_{\alpha 1}(0, 0) + \sigma_{\alpha 2}(0, 0)] + (y_{10} - y_{1T}) \Delta \sigma_{\alpha 1}(1, 0) + (y_{20} - y_{2T}) \Delta \sigma_{\alpha 2}(0, 1) + T_1^{(1)} [\alpha_1 + \Delta \sigma_{\alpha 1}(1, 0)] + T_2^{(1)} [\alpha_2 + \Delta \sigma_{\alpha 2}(0, 1)] + T_1^{(1)} \Delta \sigma_{\alpha 1}(0, 1) + T_1^{(1)} \Delta \sigma_{\alpha 2}(1, 0) + C_{11} \Delta^2 \sigma_{\alpha 1} + C_{12} \Delta^2 \sigma_{\alpha 2} + C_{11} \beta_1 + C_{22} \beta_2 + T_1^{(1)} \gamma_1 + T^{(1,1)} \gamma_2$$
(A.6)

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

$$\ln \mathbb{P}_{L\{E,W\}} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) = \mathbf{s}^{1} \left(\widetilde{\mathbf{y}} \right)' \mathbf{g}_{\boldsymbol{\alpha}}^{1} + \left[T_{1}^{(1)}, T_{1}^{(1)}, C_{11}, C_{12} \right] \mathbf{g}_{\boldsymbol{\alpha}}^{2} + C_{11} \beta_{1} + C_{22} \beta_{2} + T_{1}^{(1)} \gamma_{1} + T^{(1,1)} \gamma_{2}$$
(A.7)

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

 $[1 - \Lambda(\alpha_2 + \beta_2 \ y_{2t-1})]$. Then,

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \left(\ln \left[1 - \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1}\right)\right] + \ln \left[1 - \Lambda \left(\alpha_{2} + \beta_{22} \; y_{2t-1}\right)\right]\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \left(\ln \left[1 - \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1}\right)\right] + \ln \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2}\right)\right) \\
+ \sum_{t=1}^{T} y_{1t} \left(1 - y_{2t}\right) \left(\ln \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1}\right) + \ln \left[1 - \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1}\right)\right]\right) \\
+ \sum_{t=1}^{T} y_{1t} y_{2t} \left(\ln \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1}\right) + \ln \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2}\right)\right) \\$$
(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:

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

Grouping terms, we have:

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{2t}\right) \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})\right] \\
+ \sum_{t=1}^{T} y_{1t} \left[\alpha_{1} + \beta_{1} y_{1t-1}\right] \\
+ \sum_{t=1}^{T} y_{2t} \left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_{2} + \beta_{2} y_{2t-1} + \gamma_{2}\right] \\
+ \sum_{t=1}^{T} y_{1t}y_{2t} \left[\gamma_{1}\right]$$
(A.10)

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$, $T_{11}^{(1)}$, C_{11} , and C_{12} , we have the following ex-

pression for the lower bound $\ln \mathbb{P}_{L\{S,N\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$:

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

$$= \ln p_{\alpha} (y_{10}, y_{20}) + T [\sigma_{\alpha 1}(0, 0) + \sigma_{\alpha 2}(0, 0)] + (y_{10} - y_{1T}) \Delta \sigma_{\alpha 1}(1, 0) + (y_{20} - y_{2T}) \Delta \sigma_{\alpha 2}(0, 1) + T_1^{(1)} [\alpha_1 + \Delta \sigma_{\alpha 1}(1, 0)] + T_2^{(1)} [\alpha_2 + \Delta \sigma_{\alpha 2}(0, 1)] + T_2^{(1)} [\Delta \sigma_{\alpha 1}(0, 1) + \Delta \sigma_{\alpha 2}(1, 0)] + C_{21} \Delta^2 \sigma_{\alpha 1} + C_{22} \Delta^2 \sigma_{\alpha 2} + C_{11} \beta_1 + C_{22} \beta_2 + T^{(1,1)} \gamma_1 + T_2^{(1)} \gamma_2$$
(A.11)

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

$$\ln \mathbb{P}_{L\{S,N\}} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) = \mathbf{s}^{1} \left(\widetilde{\mathbf{y}} \right)' \mathbf{g}_{\boldsymbol{\alpha}}^{1} + \left[T_{2}^{(1)}, T_{2}^{(1)}, C_{21}, C_{22} \right] \mathbf{g}_{\boldsymbol{\alpha}}^{2} + C_{11} \beta_{1} + C_{22} \beta_{2} + T^{(1,1)} \gamma_{1} + T_{2}^{(1)} \gamma_{2}$$
(A.12)

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

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\ + \sum_{t=1}^{T} \left(1 - y_{1t}\right) \left(1 - y_{2t}\right) \left(\ln \left[1 - \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1}\right)\right] + \ln \left[1 - \Lambda \left(\alpha_{2} + \beta_{22} \; y_{2t-1}\right)\right]\right) \\ + \sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \left(\ln \left[1 - \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1}\right)\right] + \ln \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1}\right)\right) \\ + \sum_{t=1}^{T} y_{1t} \left(1 - y_{2t}\right) \left(\ln \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1}\right) + \ln \left[1 - \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1}\right)\right]\right) \\ + \sum_{t=1}^{T} y_{1t} y_{2t} \left(\ln \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1}\right) + \ln \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2}\right)\right)$$
(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:

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})\right] \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)y_{2t} \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_{2} + \beta_{2} y_{2t-1}\right] \\
+ \sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_{1} + \beta_{1} y_{1t-1}\right] \\
+ \sum_{t=1}^{T} y_{1t}y_{2t} \left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_{1} + \beta_{1} y_{1t-1} + \gamma_{1} + \alpha_{2} + \beta_{2} y_{2t-1} + \gamma_{2}\right] \\$$
(A.14)

Grouping terms, we have:

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\ + \sum_{t=1}^{T} \left(1 - y_{1t}y_{2t}\right) \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})\right] \\ + \sum_{t=1}^{T} y_{2t} \left[\alpha_{2} + \beta_{2} y_{2t-1}\right] \\ + \sum_{t=1}^{T} y_{1t} \left[\alpha_{1} + \beta_{1} y_{1t-1}\right] \\ + \sum_{t=1}^{T} y_{1t}y_{2t} \left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \gamma_{1} + \gamma_{2}\right]$$
(A.15)

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$, $T_{1}^{(1,1)}$, C_{11} , and C_{12} , we have the following ex-

pression for the lower bound $\ln \mathbb{P}_{L\{E,N\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$:

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln \mathbb{P}_{L\{E,N\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$$

$$= \ln p_{\alpha} (y_{10}, y_{20}) + T [\sigma_{\alpha 1}(0, 0) + \sigma_{\alpha 2}(0, 0)] + (y_{10} - y_{1T}) \Delta \sigma_{\alpha 1}(1, 0) + (y_{20} - y_{2T}) \Delta \sigma_{\alpha 2}(0, 1) + T_{1}^{(1)} [\alpha_{1} + \Delta \sigma_{\alpha 1}(1, 0)] + T_{2}^{(1)} [\alpha_{2} + \Delta \sigma_{\alpha 2}(0, 1)] + T^{(1,1)} [\Delta \sigma_{\alpha 1}(0, 1) + \Delta \sigma_{\alpha 2}(1, 0)] + R_{1}^{(1,1)} \Delta^{2} \sigma_{\alpha 1} + R_{2}^{(1,1)} \Delta^{2} \sigma_{\alpha 2} + C_{11} \beta_{1} + C_{22} \beta_{2} + T^{(1,1)} [\gamma_{1} + \gamma_{2}]$$
 (A.16)

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

$$\ln \mathbb{P}_{L\{E,N\}} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) = \mathbf{s}^{1} \left(\widetilde{\mathbf{y}} \right)' \mathbf{g}_{\boldsymbol{\alpha}}^{1} + \left[T^{(1,1)}, T^{(1,1)}, R_{1}^{(1,1)}, R_{2}^{(1,1)} \right] \mathbf{g}_{\boldsymbol{\alpha}}^{2} + C_{11} \beta_{1} + C_{22} \beta_{2} + T^{(1,1)} \left[\gamma_{1} + \gamma_{2} \right]$$
(A.17)

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

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \left(\ln \left[1 - \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1}\right)\right] + \ln \left[1 - \Lambda \left(\alpha_{2} + \beta_{22} \; y_{2t-1}\right)\right]\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right) y_{2t} \left(\ln \left[1 - \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1}\right)\right] + \ln \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2}\right)\right) \\
+ \sum_{t=1}^{T} y_{1t} \left(1 - y_{2t}\right) \left(\ln \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1}\right) + \ln \left[1 - \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2}\right)\right]\right) \\
+ \sum_{t=1}^{T} y_{1t} y_{2t} \left(\ln \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1}\right) + \ln \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2}\right)\right) \\$$
(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:

Grouping terms, we have:

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\
+ \sum_{t=1}^{T} \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})\right] \\
+ \sum_{t=1}^{T} y_{2t} \left[\alpha_{2} + \beta_{2} y_{2t-1} + \gamma_{2}\right] \\
+ \sum_{t=1}^{T} y_{1t} \left[\alpha_{1} + \beta_{1} y_{1t-1} + \gamma_{1}\right] \\
+ \sum_{t=1}^{T} \left[y_{1t} + y_{2t} - y_{1t}y_{2t}\right] \left[\sigma_{\alpha 1}(y_{1t-1}, 1) - \sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(1, y_{2t-1}) - \sigma_{\alpha 2}(0, y_{2t-1})\right] \\$$
(A.20)

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$, $T_{11}^{(1)}$, C_{11} , and C_{12} , we have the following ex-

pression for the lower bound $\ln \mathbb{P}_{L\{S,W\}}(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$:

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \geq \ln \mathbb{P}_{L\{S,W\}}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$$

$$= \ln p_{\alpha} (y_{10}, y_{20}) + T [\sigma_{\alpha 1}(0, 0) + \sigma_{\alpha 2}(0, 0)] + (y_{10} - y_{1T}) [\sigma_{\alpha 1}(1, 0) - \sigma_{\alpha 1}(0, 0)] + (y_{20} - y_{2T}) [\sigma_{\alpha 2}(0, 1) - \sigma_{\alpha 2}(0, 0)] + T_1^{(1)} [\alpha_1 + \Delta \sigma_{\alpha 1}(1, 0)] + T_2^{(1)} [\alpha_2 + \Delta \sigma_{\alpha 2}(0, 1)] + [T_1^{(1)} + T_2^{(1)} - T^{(1,1)}] [\Delta \sigma_{\alpha 1}(0, 1) + \Delta \sigma_{\alpha 2}(1, 0)] + [C_{11} + C_{21} - R_1^{(1,1)}] \Delta^2 \sigma_{\alpha 1} + [C_{12} + C_{22} - R_2^{(1,1)}] \Delta^2 \sigma_{\alpha 2} + C_{11} \beta_1 + C_{22} \beta_2 + T_1^{(1)} \gamma_1 + T_2^{(1)} \gamma_2$$
(A.21)

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

$$\ln \mathbb{P}_{L\{S,W\}} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) = \mathbf{s}^{1} \left(\widetilde{\mathbf{y}} \right)' \mathbf{g}_{\boldsymbol{\alpha}}^{1} \\ + \left[T_{1}^{(1)} + T_{2}^{(1)} - T^{(1,1)}, T_{1}^{(1)} + T_{2}^{(1)} - T^{(1,1)}, C_{11} + C_{21} - R_{1}^{(1,1)}, C_{12} + C_{22} - R_{2}^{(1,1)} \right] \\ + C_{11} \beta_{1} + C_{22} \beta_{2} + T_{1}^{(1)} \gamma_{1} + T_{2}^{(1)} \gamma_{2}$$
(A.22)

g

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

$$\ln \mathbb{P} \left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta} \right) \leq \ln p_{\boldsymbol{\alpha}} \left(y_{10}, y_{20} \right)$$

$$+ \sum_{t=1}^{T} \left(1 - y_{1t} \right) \left(1 - y_{2t} \right) \left(\ln \left[1 - \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} \right) \right] + \ln \left[1 - \Lambda \left(\alpha_{2} + \beta_{22} \; y_{2t-1} \right) \right] \right)$$

$$+ \sum_{t=1}^{T} \left(1 - y_{1t} \right) y_{2t} \left(\ln \left[1 - \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1} \right) \right] + \ln \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} \right) \right)$$

$$+ \sum_{t=1}^{T} y_{1t} \left(1 - y_{2t} \right) \left(\ln \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} \right) + \ln \left[1 - \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2} \right) \right] \right)$$

$$+ \sum_{t=1}^{T} y_{1t} y_{2t} \left(\ln \Lambda \left(\alpha_{1} + \beta_{1} \; y_{1t-1} + \gamma_{1} \right) + \ln \Lambda \left(\alpha_{2} + \beta_{2} \; y_{2t-1} + \gamma_{2} \right) \right)$$

$$(A.23)$$

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

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \leq \ln p_{\alpha}\left(y_{10}, y_{20}\right) \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)\left(1 - y_{2t}\right) \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(0, y_{2t-1})\right] \\
+ \sum_{t=1}^{T} \left(1 - y_{1t}\right)y_{2t} \left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(0, y_{2t-1}) + \alpha_{2} + \beta_{2} y_{2t-1}\right] \\
+ \sum_{t=1}^{T} y_{1t}\left(1 - y_{2t}\right) \left[\sigma_{\alpha 1}(y_{1t-1}, 0) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_{1} + \beta_{1} y_{1t-1}\right] \\
+ \sum_{t=1}^{T} y_{1t}y_{2t} \left[\sigma_{\alpha 1}(y_{1t-1}, 1) + \sigma_{\alpha 2}(1, y_{2t-1}) + \alpha_{1} + \beta_{1} y_{1t-1} + \gamma_{1} + \alpha_{2} + \beta_{2} y_{2t-1} + \gamma_{2}\right] \\$$
(A.24)

Grouping terms, we have:

$$\ln \mathbb{P}\left(\tilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \leq \ln p_{\boldsymbol{\alpha}}\left(y_{10}, y_{20}\right) + \sum_{t=1}^{T} \left(1 - y_{2t}\right) \, \sigma_{\alpha 1}(y_{1t-1}, 0) + \left(1 - y_{1t}\right) \, \sigma_{\alpha 2}(0, y_{2t-1}) + \sum_{t=1}^{T} y_{2t} \, \sigma_{\alpha 1}(y_{1t-1}, 1) + y_{2t} \, \left[\alpha_{2} + \beta_{2} \, y_{2t-1}\right] + \sum_{t=1}^{T} y_{1t} \, \sigma_{\alpha 2}(1, y_{2t-1}) + y_{1t} \, \left[\alpha_{1} + \beta_{1} \, y_{1t-1}\right] + \sum_{t=1}^{T} y_{1t} y_{2t} \, \left[\gamma_{1} + \gamma_{2}\right]$$
(A.25)

Using the definitions of the statistics $T_1^{(1)}$, $T_2^{(1)}$, $T_{1}^{(1,1)}$, C_{11} , and C_{12} , we have the following ex-

pression for the upper bound $\ln \mathbb{P}_U(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta})$:

$$\ln \mathbb{P}\left(\widetilde{\mathbf{y}} \mid \alpha, \boldsymbol{\theta}\right) \leq \ln \mathbb{P}_{U}\left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}\right)$$

$$= \ln p_{\alpha} (y_{10}, y_{20}) + T [\sigma_{\alpha 1}(0, 0) + \sigma_{\alpha 2}(0, 0)]$$

$$+ (y_{10} - y_{1T}) \Delta \sigma_{\alpha 1}(1, 0) + (y_{20} - y_{2T}) \Delta \sigma_{\alpha 2}(0, 1)$$

$$+ T_{1}^{(1)} [\alpha_{1} + \Delta \sigma_{\alpha 1}(1, 0)] + T_{2}^{(1)} [\alpha_{2} + \Delta \sigma_{\alpha 2}(0, 1)]$$

$$+ T_{2}^{(1)} \Delta \sigma_{\alpha 1}(0, 1) + T_{1}^{(1)} \Delta \sigma_{\alpha 2}(1, 0)$$

$$+ C_{21} \Delta^{2} \sigma_{\alpha 1} + C_{12} \Delta^{2} \sigma_{\alpha 2}$$

$$+ C_{11} \beta_{1} + C_{22} \beta_{2} + T^{(1,1)} [\gamma_{1} + \gamma_{2}]$$

$$(A.26)$$

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

$$\ln \mathbb{P}_{U} \left(\widetilde{\mathbf{y}} \mid \boldsymbol{\alpha}, \boldsymbol{\theta} \right) = \mathbf{s}^{1} \left(\widetilde{\mathbf{y}} \right)' \mathbf{g}_{\boldsymbol{\alpha}}^{1} + \left[T_{2}^{(1)}, T_{1}^{(1)}, C_{21}, C_{12} \right] \mathbf{g}_{\boldsymbol{\alpha}}^{2} + C_{11} \beta_{1} + C_{22} \beta_{2} + T^{(1,1)} [\gamma_{1} + \gamma_{2}]$$
(A.27)

A.8 Proof of Proposition 8

Denote the following terms:

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

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

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

$$\begin{split} \ln P_{U}(\tilde{y}) &= \ln P_{\alpha}(y_{10}, y_{20}) + s^{1}(\tilde{y})'g_{\alpha}^{1} + [y_{10} - y_{1T}, y_{20} - y_{2T}, T^{(1,1)}, T^{(1,1)}]'g_{\alpha}^{2} + [T_{2}^{(1)}, T_{1}^{(1)}, C_{21}, C_{12}]g_{\alpha}^{3} \\ &+ C_{11}\beta_{1} + C_{22}\beta_{2} \\ \ln P_{L\{E,W\}}(\tilde{y}) &= \ln P_{\alpha}(y_{10}, y_{20}) + s^{1}(\tilde{y})'g_{\alpha}^{1} + [y_{10} - y_{1T}, y_{20} - y_{2T}, T_{1}^{(1)}, T^{(1,1)}]'g_{\alpha}^{2} + [T_{1}^{(1)}, T_{1}^{(1)}, C_{11}, C_{12}]g_{\alpha}^{3} \\ &+ C_{11}\beta_{1} + C_{22}\beta_{2} \\ \ln P_{L\{S,N\}}(\tilde{y}) &= \ln P_{\alpha}(y_{10}, y_{20}) + s^{1}(\tilde{y})'g_{\alpha}^{1} + [y_{10} - y_{1T}, y_{20} - y_{2T}, T^{(1,1)}, T_{2}^{(1)}]'g_{\alpha}^{2} + [T_{2}^{(1)}, T_{2}^{(1)}, C_{21}, C_{22}]g_{\alpha}^{3} \\ &+ C_{11}\beta_{1} + C_{22}\beta_{2} \\ \ln P_{L\{E,N\}}(\tilde{y}) &= \ln P_{\alpha}(y_{10}, y_{20}) + s^{1}(\tilde{y})'g_{\alpha}^{1} + [y_{10} - y_{1T}, y_{20} - y_{2T}, T^{(1,1)}, T^{(1,1)}]'g_{\alpha}^{2} \\ &+ [T^{(1,1)}, T^{(1,1)}, R_{1}^{(1,1)}, R_{2}^{(1,1)}]g_{\alpha}^{3} \\ &+ C_{11}\beta_{1} + C_{22}\beta_{2} \\ \ln P_{L\{S,W\}}(\tilde{y}) &= \ln P_{\alpha}(y_{10}, y_{20}) + s^{1}(\tilde{y})'g_{\alpha}^{1} + [y_{10} - y_{1T}, y_{20} - y_{2T}, T_{1}^{(1)}, T_{2}^{(1)}]'g_{\alpha}^{2} \\ &+ [T^{(1,1)}, T^{(1,1)}, R_{1}^{(1,1)}, R_{2}^{(1,1)}]g_{\alpha}^{3} \\ &+ C_{11}\beta_{1} + C_{22}\beta_{2} \\ \ln P_{L\{S,W\}}(\tilde{y}) &= \ln P_{\alpha}(y_{10}, y_{20}) + s^{1}(\tilde{y})'g_{\alpha}^{1} + [y_{10} - y_{1T}, y_{20} - y_{2T}, T_{1}^{(1)}, T_{2}^{(1)}]'g_{\alpha}^{2} \\ &+ [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)}]g_{\alpha}^{3} \\ &+ C_{11}\beta_{1} + C_{22}\beta_{2} \\ \ln P_{L\{S,W\}}(\tilde{y}) &= \ln P_{\alpha}(y_{10}, y_{20}) + s^{1}(\tilde{y})'g_{\alpha}^{1} + [y_{10} - y_{1T}, y_{20} - y_{2T}, T_{1}^{(1)}, T_{2}^{(1)}]'g_{\alpha}^{2} \\ &+ [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)}]g_{\alpha}^{3} \\ \end{split}$$

$$+C_{11}\beta_1+C_{22}\beta_2$$

where $s^{1}(\tilde{y}) = [T, T_{1}^{(1)}, T_{2}^{(1)}], g_{\alpha}^{1} = [\sigma_{\alpha_{1}}(0, 0) + \sigma_{\alpha_{2}}(0, 0), \alpha_{1} + \Delta\sigma_{\alpha_{1}}(1, 0), \alpha_{2} + \Delta\sigma_{\alpha_{2}}(0, 1)]', g_{\alpha}^{2} = [\Delta\sigma_{\alpha_{1}}(1, 0), \Delta\sigma_{\alpha_{2}}(0, 1), \tilde{\gamma}_{1\alpha}, \tilde{\gamma}_{2\alpha}]', \text{ and } g_{\alpha}^{3} = [\Delta\sigma_{\alpha_{1}}(0, 1), \Delta\sigma_{\alpha_{2}}(1, 0), \Delta^{2}\sigma_{\alpha_{1}}, \Delta^{2}\sigma_{\alpha_{2}}]'$

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

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

For a pair of choice histroies A and B, define

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

Define the statistics $s^1(\tilde{y}) = [T, T_1^{(1)}, T_2^{(1)}]$ [1] Using upper bound and $L\{E, W\}$:

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

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

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

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

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

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

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

provided the following conditions hold: (i) $y_{10}(A) = y_{20}(B)$, (ii) $y_{20}(A) - y_{20}(B)$, (iii) $s^{1}(A) = s^{1}(B)$, (iv) element-wise, $[y_{10}(A) - y_{1T}(A), y_{20}(A) - y_{2T}(A), T^{(1,1)}(A), T^{(1,1)}(A)] - [y_{10}(B) - y_{1T}(A), y_{20}(A) - y_{2T}(A), T^{(1,1)}(A)] - [y_{10}(B) - y_{1T}(A), y_{20}(A) - y_{2T}(A), T^{(1,1)}(A)] - [y_{10}(B) - y$

 $y_{1T}(B), y_{20}(B) - y_{2T}(B), T^{(1,1)}(B), T^{(1,1)}(B)] \ge 0, \text{ (v) element-wise, } [T_2^{(1)}(A), T_1^{(1)}(A), C_{21}(A), C_{12}(A)] - [T^{(1,1)}(B), T^{(1,1)}(B), R_1^{(1,1)}(B), R_2^{(1,1)}(B)] \le 0.$

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

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

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

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

References

- AGUIRREGABIRIA, V., AND J. CARRO (2024): "Identification of average marginal effects in fixed effects dynamic discrete choice models," *Review of Economics and Statistics*, Forthcoming.
- AGUIRREGABIRIA, V., A. COLLARD-WEXLER, AND S. P. RYAN (2021): "Dynamic games in empirical industrial organization," in *Handbook of industrial organization*, vol. 4, pp. 225–343. Elsevier.
- AGUIRREGABIRIA, V., J. GU, AND Y. LUO (2021): "Sufficient statistics for unobserved heterogeneity in dynamic structural logit models," *Journal of Econometrics*, 223(2), 280–311.
- AGUIRREGABIRIA, V., AND P. MIRA (2007): "Sequential estimation of dynamic discrete games," *Econometrica*, 75(1), 1–53.
- ANDERSEN, E. (1970): "Asymptotic properties of conditional maximum likelihood estimators," Journal of the Royal Statistical Society, Series B, 32, 283–301.
- ARCIDIACONO, P., AND R. A. MILLER (2011): "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity," *Econometrica*, 79(6), 1823– 1867.
- BLUME, L., W. BROCK, S. DURLAUF, AND Y. IOANNIDES (2011): "Identification of social interactions," in *Handbook of Social Economics, Vol. 1*, pp. 853–964. North-Holland.
- BONHOMME, S. (2012): "Functional differencing," *Econometrica*, 80(4), 1337–1385.
- BRESNAHAN, T., AND P. REISS (1991): "Empirical models of discrete games," Journal of Econometrics, 48(1), 57–81.
- BROCK, W., AND S. DURLAUF (2007): "Identification of binary choice models with social interactions," *Journal of Econometrics*, 140(1), 52–75.

- CHAMBERLAIN, G. (1980): "Analysis of covariance with qualitative data," *Review of Economic* Studies, 47(1), 225–238.
- Cox, D. (1958): "The regression analysis of binary sequences," Journal of the Royal Statistical Society. Series B, 20(2), 215–242.
- COX, G., AND X. SHI (2023): "Simple adaptive size-exact testing for full-vector and subvector inference in moment inequality models," *The Review of Economic Studies*, 90(1), 201–228.
- DOBRONYI, C., J. GU, AND K. KIM (2021): "Identification of Dynamic Panel Logit Models with Fixed Effects," arXiv:2104.04590.
- ECKSTEIN, Z., AND O. LIFSHITZ (2015): "Household interaction and the labor supply of married women," *International Economic Review*, 56(2), 427–455.
- ERDEM, T., S. IMAI, AND M. KEANE (2003): "Brand and quantity choice dynamics under price uncertainty," *Quantitative Marketing and Economics*, 1(1), 5–64.
- ERICSON, R., AND A. PAKES (1995): "Markov-perfect industry dynamics: a framework for empirical work," *The Review of Economic Studies*, 62(1), 53–82.
- HECKMAN, J. (1981): "The incidental parameters problem and the problem of initial conditions in estimating a discrete time - discrete data stochastic process," in *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press.
- HONORÉ, B., AND A. DE PAULA (2021): "Identification in simple binary outcome panel data models," *The Econometrics Journal*, 24(2), C78–C93.
- HONORÉ, B., AND E. KYRIAZIDOU (2019): "Identification in Binary Response Panel Data Models: Is Point-Identification More Common Than We Thought?," Annals of Economics and Statistics, 134, 207–226.

- HONORÉ, B., AND M. WEIDNER (2020): "Dynamic Panel Logit Models with Fixed Effects," arXiv preprint arXiv:2005.05942.
- KASAHARA, H., AND K. SHIMOTSU (2009): "Nonparametric identification of finite mixture models of dynamic discrete choices," *Econometrica*, 77(1), 135–175.
- MONGEY, S. (2021): "Market structure and monetary non-neutrality," Discussion paper, National Bureau of Economic Research.
- NAKAMURA, E., AND J. STEINSSON (2008): "Five facts about prices: A reevaluation of menu cost models," *The Quarterly Journal of Economics*, 123(4), 1415–1464.
- NARENDRANTHAN, W., S. NICKELL, AND D. METCALF (1985): "An investigation into the incidence and dynamic structure of sickness and unemployment in Britain, 1965–75," *Journal of the Royal Statistical Society: Series A*, 148(3), 254–267.
- RASCH, G. (1961): "On General Laws and the Meaning of Measurement in Psychology," Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 4, 321–333.
- ROTEMBERG, J. J. (1982): "Sticky prices in the United States," *Journal of Political Economy*, 90(6), 1187–1211.
- SIEG, H., AND C. YOON (2017): "Estimating dynamic games of electoral competition to evaluate term limits in us gubernatorial elections," *American Economic Review*, 107(7), 1824–1857.
- SOVINSKY, M., AND S. STERN (2016): "Dynamic modelling of long-term care decisions," *Review* of *Economics of the Household*, 14(2), 463–488.
- TAMER, E. (2003): "Incomplete simultaneous discrete response model with multiple equilibria," The Review of Economic Studies, 70(1), 147–165.

- WAGNER, U. (2016): "Estimating strategic models of international treaty formation," The Review of Economic Studies, 83(4), 1741–1778.
- WANG, O., AND I. WERNING (2022): "Dynamic oligopoly and price stickiness," American Economic Review, 112(8), 2815–2849.