

Identification of Dynamic Discrete Choice Games with Fixed Effects Unobserved Heterogeneity

Victor Aguirregabiria[†] Jiaying Gu[‡] Pedro Mira[§]

October 17, 2026

Abstract

Persistent unobserved heterogeneity can confound estimates of state dependence and strategic interactions in dynamic discrete choice games. We study identification in two-player binary-choice games observed in short panels, allowing player-market fixed effects to be arbitrarily correlated across players and with initial conditions. In models with a unique equilibrium, conditional likelihood (CL) provides closed-form identification formulas for state-dependence parameters. Functional differencing (FD) extends these results by delivering point identification of contemporaneous strategic interactions that the CL approach does not identify. For models with multiple equilibria, we extend both approaches to bounds on choice-history probabilities, allowing unrestricted equilibrium selection. CL-Bounds yields bounds on state-dependence parameters and, in myopic games, contemporaneous strategic interactions. FD-Bounds provides additional restrictions that can tighten these bounds. Numerical experiments illustrate the incremental identification gains from FD and FD-Bounds. We also assess the empirical identifying power of these methods through three applications: couples' employment decisions using Current Population Survey data, gasoline price adjustment using data from French gas stations, and cooperation using data from a laboratory experiment.

Keywords: Dynamic games; Fixed effects; State dependence; Conditional likelihood; Functional differencing; Multiple equilibria; Partial identification.

JEL codes: C23; C25; C51; C61.

Acknowledgments. We thank Stéphane Bonhomme, Jesús Carro, Andrew Chesher, Áureo de Paula, Bo Honoré, and Ariel Pakes for helpful comments and suggestions. We also thank seminar participants at Michigan State and participants at the International Panel Data Conference; the IAAE conference; the DSE Conference in Chicago; the Banff Empirical Microeconomics Workshop; the Bristol Econometrics Study Group; the SITE 2026 workshop, "Empirical Implementation of Theoretical Models of Strategic Interaction and Dynamic Behavior"; and the 2025 Workshop on Econometrics & Models of Strategic Interactions at Queen Mary University of London.

[†] University of Toronto and CEPR, victor.aguirregabiria@utoronto.ca.

[‡] University of Toronto, jiaying.gu@utoronto.ca.

[§] CEMFI, mira@cemfi.es.

1 Introduction

Persistence in economic decisions need not reflect true state dependence. A firm may remain out of a market because entry is costly or because the market is permanently unprofitable. Likewise, correlated decisions need not reflect strategic interactions: two firms may respond to correlated, persistent market conditions. These alternative explanations have substantially different implications for economic behavior and the effects of policy interventions. However, distinguishing between them is difficult in short panels. To obtain robust evidence on state dependence and strategic interactions, we adopt a fixed effects approach that leaves the joint distribution of player-market heterogeneity unrestricted. This flexibility poses identification challenges and calls for methods that address both the incidental parameters problem in short panels and the initial conditions problem arising from dependence between initial choices and unobserved heterogeneity.

In this paper, we study the identification of state dependence and strategic interactions in two-player binary-choice dynamic games observed across many markets for a fixed number of periods. Payoffs contain additive player-market fixed effects and independent logit shocks, which both players observe. The fixed effects may exhibit arbitrary dependence across players and with the initial outcome. We impose neither a parametric mixing distribution nor a finite-support restriction in the identification analysis.

Dynamic games introduce two further identification challenges that do not arise in reduced-form dynamic discrete-choice panel models. First, forward-looking players respond to continuation values that depend on their unobserved characteristics and on equilibrium behavior. Because these values are unknown, endogenous functions of the fixed effects, eliminating the fixed effects also requires eliminating the continuation values. The challenge is to do so while preserving enough variation to identify the structural parameters. Second, in games with simultaneous moves, a given state and realization of payoff shocks may support multiple equilibrium outcomes. Without a specified equilibrium-selection rule, the model therefore predicts a set of choice probabilities. This incompleteness creates an additional identification challenge.

We ask which features of short choice histories distinguish state dependence and strategic interactions from persistent heterogeneity, and how this identifying

information changes when we allow for forward-looking behavior and equilibrium multiplicity. We distinguish between complete models, with a unique equilibrium, and incomplete models, with multiple equilibria. For complete models, we begin with the standard conditional likelihood approach based on sufficient statistics (CL-SS) (Chamberlain, 1980, 1985, 2010; Honoré and Kyriazidou, 2000, 2019). We show that CL-SS delivers point identification of the parameters governing dependence on a player’s own lagged choice, but does not identify those governing contemporaneous or lagged strategic interactions. We then apply functional differencing (FD) to derive additional moment conditions that identify all model parameters, including those governing contemporaneous strategic interactions.

For incomplete models, we develop two partial identification methods that eliminate incidental parameters from the model’s bounds on choice-history probabilities. The first extends conditional likelihood (CL-Bounds), and the second extends functional differencing (FD-Bounds). In a complete model, matching the incidental components in the exact likelihoods of two histories yields a probability ratio that identifies a combination of structural parameters. In an incomplete model, we match the incidental component of an upper probability bound for one history with that of a lower bound for another. The resulting inequality holds for every value of the fixed effects and therefore remains valid after integrating over their distribution. When exact matching is not possible, uniform sign restrictions can yield valid inequalities without exact cancellation. Neither approach requires specifying an equilibrium-selection rule in the multiplicity region.

Functional differencing (FD) supplies restrictions beyond these pairwise comparisons. For a complete model, it searches for weighted sums of history probabilities that vanish for every fixed effect. For incomplete models, nonnegative combinations of lower and upper bounds produce further inequalities. We show how to construct these combinations as a cone defined by polynomial coefficient restrictions. The construction contains the simple CL-Bounds comparisons and can improve their bounds substantially.

Three empirical applications illustrate the implementation of these methods and its possible limits. Couples’ employment provides a familiar setting for separating state dependence from correlated household heterogeneity. Daily gasoline prices provide many reversible decisions and many geographic pairs. A repeated prisoner’s-dilemma experiment provides clearly defined pairs and choices, but

relatively few histories informative for some sequential specifications. We report exact-history estimates, sequential FD estimates, and the empirical support for selection-robust bounds. These are empirical illustrations of maintained restrictions, not evidence that a single timing or information structure is appropriate in all three settings.

The paper relates to three literatures. Conditional likelihood methods for fixed-effects discrete-choice panels originate in the work of [Chamberlain \(1980, 1985\)](#), with subsequent extensions including [Honoré and Kyriazidou \(2000\)](#). The special case of our model without contemporaneous strategic interactions corresponds to the bivariate dynamic-logit benchmark established in [Narendranathan et al. \(1985\)](#) and subsequent work, including [Honoré and Kyriazidou \(2019\)](#). [Aguirregabiria et al. \(2021\)](#) show how sufficient statistics can eliminate continuation values in single-agent structural models. We extend this logic to models with strategic interactions and to bounds on choice-history probabilities.

Our FD analysis builds on [Bonhomme \(2012\)](#), the systematic moment construction of [Honoré and Weidner \(2025\)](#), and the polynomial and moment-feasibility approach of [Dobronyi et al. \(2026\)](#). The latter distinguishes equality restrictions from the shape restrictions needed for a feasible mixing distribution. [Aguirregabiria and Carro \(2026\)](#) study identification of average marginal effects, and [Pakel and Weidner \(2026\)](#) develop tractable outer bounds for average effects. Our focus is on common payoff coefficients in games and on cancellation that remains valid when equilibrium selection is unspecified.

Finally, our framework complements research on dynamic games with unobserved heterogeneity, including [Aguirregabiria and Mira \(2007\)](#); [Kasahara and Shimotsu \(2009\)](#); [Arcidiacono and Miller \(2011\)](#); [Aguirregabiria and Mira \(2019\)](#). We allow an unrestricted distribution of permanent heterogeneity, while maintaining a deliberately simple state space and shock structure. The bounds approach also connects to the analysis of incomplete simultaneous discrete-response models in [Tamer \(2003\)](#). We do not equate a collection of valid inequalities with a sharp characterization.

Section 2 sets out the model and summarizes the results. Section 3 studies identification in complete models using CL-SS and FD. Section 4 develops CL and FD bounds for incomplete models. Section 5 presents the applications. Section 6 concludes. The appendices contain the likelihood algebra, recursive-model exten-

sions, numerical implementation, and the distinction between outer regions and sharp identified sets.

2 Model and overview of results

2.1 Data, payoffs, information, and equilibrium

Each market m contains two players, indexed by $i \in \{1, 2\}$. For a given player i , let $j \neq i$ denote the other player. Let $y_{imt} \in \{0, 1\}$ denote player i 's choice in market m at time t . For each market, we observe the initial joint outcome (y_{1m0}, y_{2m0}) and T subsequent joint outcomes at $t = 1, \dots, T$. Thus, $T = 3$ corresponds to three transitions and four observed joint outcomes. We study identification as the number of markets grows while T remains fixed.

Player i has discount factor $\delta_i \in [0, 1)$ and flow payoff

$$U_{imt} = u_{im}(y_{imt}, y_{jmt}; \mathbf{y}_{m,t-1}) + \varepsilon_{imt}(y_{imt}), \quad (1)$$

where the flow-payoff difference at state $\mathbf{y}_{m,t-1} = (s_1, s_2)$ is

$$u_{im}(1, y_j; s) - u_{im}(0, y_j; s) = \alpha_{im} + \beta_i s_i + \lambda_i s_j + \gamma_i y_j. \quad (2)$$

The coefficient β_i measures own-state dependence, λ_i a lagged effect of the other player's action, and γ_i a contemporaneous payoff interaction.

The shocks $\varepsilon_{imt}(0)$ and $\varepsilon_{imt}(1)$ are independent type-I extreme-value draws across players, alternatives, markets, and dates, and independent of permanent heterogeneity and initial conditions, $(\alpha_{1m}, \alpha_{2m}, y_{1m0}, y_{2m0})$. The difference $\varepsilon_{imt} \equiv \varepsilon_{imt}(0) - \varepsilon_{imt}(1)$ is standard logistic, with cdf $\Lambda(v) = e^v / (1 + e^v)$. Both players observe the current shocks and permanent market characteristics. The researcher observes neither. The joint distribution of $(\alpha_{1m}, \alpha_{2m}, y_{1m0}, y_{2m0})$ is unrestricted. The coefficients (β, λ, γ) are common across markets.

We consider stationary Markov perfect equilibria. Player i 's strategy is a time-invariant function of the payoff-relevant state variables: $y_{imt} = \sigma_{im}(\mathbf{y}_{m,t-1}, \varepsilon_{mt})$. Given the equilibrium continuation values, players' choices satisfy the following

best-response equations:

$$\begin{aligned} y_{1mt} &= \mathbf{1} \left\{ \tilde{u}_{1m}(y_{2mt}, \mathbf{y}_{m,t-1}) + \tilde{V}_{1m}(y_{2mt}) - \varepsilon_{1mt} \geq 0 \right\} \\ y_{2mt} &= \mathbf{1} \left\{ \tilde{u}_{2m}(y_{1mt}, \mathbf{y}_{m,t-1}) + \tilde{V}_{2m}(y_{1mt}) - \varepsilon_{2mt} \geq 0 \right\}, \end{aligned} \quad (3)$$

where $\varepsilon_{imt} \equiv \varepsilon_{imt}(0) - \varepsilon_{imt}(1)$ is the difference in payoff shocks. Here and below $\mathbf{1}\{\cdot\}$ denotes the indicator function. The term $\tilde{u}_{im}(y_{jmt}, \mathbf{y}_{m,t-1})$ denotes the difference in flow utility between choices 1 and 0, excluding payoff shocks: $\tilde{u}_{im}(y_{jmt}, \mathbf{y}_{m,t-1}) \equiv u_{im}(1, y_{jmt}, \mathbf{y}_{m,t-1}) - u_{im}(0, y_{jmt}, \mathbf{y}_{m,t-1})$. Similarly, $\tilde{V}_{im}(y_{jmt})$ denotes the difference in continuation values holding the other player's current choice fixed: $\tilde{V}_{im}(y_{jmt}) = \delta_i [V_{im}(1, y_{jmt}) - V_{im}(0, y_{jmt})]$, where V_{im} represents the value function integrated over the i.i.d. ε shocks.

The selected pair of equilibrium continuation-value functions remains fixed over time within each market. These functions may vary across markets and may be correlated with unobserved heterogeneity and the initial state $\mathbf{y}_{m0}$.

2.2 A common representation

Define

$$a_{im} \equiv \alpha_{im} + \tilde{V}_{im}(0), \quad c_{im} \equiv \tilde{V}_{im}(1) - \tilde{V}_{im}(0), \quad \kappa_{im} \equiv \gamma_i + c_{im}. \quad (4)$$

The best responses become:

$$y_{imt} = \mathbf{1} \{ a_{im} + \beta_i y_{im,t-1} + \lambda_i y_{jm,t-1} + \kappa_{im} y_{jmt} - \varepsilon_{imt} \geq 0 \}. \quad (5)$$

For myopic players, $c_{im} = 0$, $a_{im} = \alpha_{im}$, and $\kappa_{im} = \gamma_i$.

Representation (5) leaves the coefficients on lagged choices explicit because the next-period state is the current joint action. It does not imply that these coefficients are identified. Nor does it imply that $\gamma_i = 0$ eliminates effective strategic interaction: the continuation increment can be nonzero when past actions affect the opponent's future incentives.

For the identification arguments, we use $\boldsymbol{\eta}_m$ to represent all time-invariant incidental values, including effective intercepts, heterogeneous effective interactions when applicable, and the selected continuation profile: $\boldsymbol{\eta}_m = (\mathbf{a}_m, \boldsymbol{\kappa}_m)$.

2.3 Exact predictions, multiplicity, and the model classification

Compared with reduced-form dynamic logit models, the dynamic game described by equation (5) raises two additional identification challenges. First, with forward-looking players, the continuation-value terms c_{im} enter the best-response functions as incidental parameters that interact with the rival's current choice. Second, with multiple equilibria, the model does not deliver unique predictions for players' choice probabilities, even conditional on the incidental parameters. We establish identification results that allow for both features at once. We also derive new results for the cases in which only one of them is present, since these intermediate cases are of independent interest. Studying these cases separately clarifies the role of each challenge and guides our classification of model specifications and identification results.

Table 1 organizes the seven main specifications according to three criteria: completeness, forward-looking behavior, and contemporaneous payoff interactions. A model is *complete* if, conditional on η_m , its timing and maintained equilibrium-selection restrictions determine a unique vector of choice-history probabilities. It is *incomplete* if multiple such vectors remain admissible. Completeness therefore concerns probability predictions conditional on the maintained incidental values. In our model, incompleteness arises from equilibrium multiplicity under simultaneous moves. Imposing sequential moves in every period eliminates within-period multiplicity and yields complete probability predictions conditional on the maintained continuation values.

The second criterion distinguishes myopic players, for whom $c_{im} = 0$, from forward-looking players, whose continuation values can generate $c_{im} \neq 0$. Finally, we obtain additional identification results by ruling out contemporaneous interactions in flow payoffs, $\gamma_1 = \gamma_2 = 0$. Under this restriction, direct strategic effects operate through lagged choices and the parameters λ , as in dynamic games with *time-to-build*.

Row A.1 of Table 1 presents our simplest benchmark: a model with sequential moves, myopic players, and no contemporaneous interactions. This specification has a unique equilibrium and corresponds to the models studied by [Narendranathan et al. \(1985\)](#) and [Honoré and Kyriazidou \(2019\)](#). Row B.3, at the bottom of the table, presents our most general specification, allowing for forward-looking players, two-way contemporaneous interactions, and unrestricted equilibrium se-

Table 1: Identification results by model and method

Model	— Identification Method —	
	CL	FD
A. Complete Models: Unique Equilibrium (Sequential move)		
No Contemporaneous Interactions: $\gamma = 0$		
A.1. Myopic	Point(β, λ)	—
A.2. Forward-looking	Point(β) if $\lambda = 0$	Point(β, λ) if $T \geq 5$
Two-way contemporaneous interactions		
A.3. Myopic	Point(β) if $\lambda = 0$	Point(β, λ, γ) if $T \geq 4$
A.4. Forward-looking	Point(β) if $\lambda = 0$	Point(β, λ) if $T \geq 5$
Model	— Identification Method —	
	CL-Bounds	FD-Bounds
B. Incomplete Models: Multiple Equilibria (Simultaneous move)		
No Contemporaneous Interactions: $\gamma = 0$		
B.1. Forward-looking	Point(β)	—
Two-way contemporaneous interactions		
B.2. Myopic	Set(β, γ)	RSet(β, γ)
B.3. Forward-looking	Set(β)	RSet(β)

$\beta = (\beta_1, \beta_2)$, with analogous vector notation for λ and γ . Point and Set denote point and set identification, respectively. RSet means a smaller identified set. Dash bar — means no additional identification.

Empirical or methodological applications:

(A.1) [Narendranathan et al. \(1985\)](#); [Honoré and Kyriazidou \(2019\)](#).

(A.2) [Igami \(2017\)](#); [Igami and Uetake \(2020\)](#); [Iskhakov et al. \(2016\)](#), alternating-move version.

(A.3) [Bresnahan and Reiss \(1991\)](#); [Kooreman \(1994\)](#); [Maruyama \(2014\)](#); [Blevins \(2015\)](#).

(A.4) [Schmidt-Dengler \(2025\)](#).

(B.1) [Iskhakov et al. \(2016\)](#), simultaneous-move version.

(B.2) [Bresnahan and Reiss \(1991\)](#); [Tamer \(2003\)](#); [Kooreman \(1994\)](#), simultaneous-move version; [Soetevent and Kooreman \(2007\)](#).

(B.3) [Besanko et al. \(2010\)](#); [Gallant et al. \(2018\)](#).

lection. Appendix D presents additional identification results for specifications with one-way contemporaneous interactions, in which either $\gamma_1 = 0$ or $\gamma_2 = 0$. For each version of the model, the notes to Table 1 cite representative empirical applications. With the exception of Honoré and Kyriazidou (2019) in group A.1, all of them rely on a random-effects approach: they specify a parametric distribution for the unobserved heterogeneity and either assume it is independent of the initial conditions or model their correlation parametrically.

2.4 Outcome probabilities and probability bounds

Fix a lagged state $s = (s_1, s_2)$ and incidental values (a, κ) . Define

$$e_i^z(s) \equiv a_i + \beta_i s_i + \lambda_i s_j + \kappa_i z, \quad F_i^z(s) \equiv \Lambda(e_i^z(s)), \quad \text{for } z \in \{0, 1\}. \quad (6)$$

Figure 1 partitions the space of $(\varepsilon_1, \varepsilon_2)$ according to the equilibrium outcome, using the threshold values $e_i^0(s)$ and $e_i^1(s)$ for $i = 1, 2$. The left panel shows the case of strategic substitutes, $\kappa_1, \kappa_2 \leq 0$. Here, the simultaneous-move game has two equilibria when $(\varepsilon_1, \varepsilon_2)$ lies in the rectangle $(e_1^1, e_1^0] \times (e_2^1, e_2^0]$, and a unique equilibrium outside it. The right panel shows the case of strategic complements, $\kappa_1, \kappa_2 \geq 0$. The thresholds are now reversed, and the region of multiple equilibria is the rectangle $(e_1^0, e_1^1] \times (e_2^0, e_2^1]$.

For notational simplicity, we suppress the state s and the incidental parameters as arguments. Let $p_{y_1 y_2}$ denote the probability that the outcome (y_1, y_2) is observed. When multiple equilibria can generate an outcome, the model does not pin down this probability, because it also depends on the equilibrium selection mechanism. For such outcomes, let $L_{y_1 y_2}$ and $U_{y_1 y_2}$ denote the sharp lower and upper bounds on $p_{y_1 y_2}$, respectively. Under strategic substitutes, we have

$$\begin{aligned} p_{00} &= (1 - F_1^0)(1 - F_2^0), & p_{11} &= F_1^1 F_2^1, \\ U_{10} &= F_1^0(1 - F_2^1), & U_{01} &= (1 - F_1^1)F_2^0, \\ L_{10} &= U_{10} - D, & L_{01} &= U_{01} - D. \end{aligned} \quad (7)$$

where $D \equiv (F_1^0 - F_1^1)(F_2^0 - F_2^1) \geq 0$.

These bounds are sharp for *individual one-period* choice probabilities, conditional on the state and incidental parameters. For a choice history $\mathbf{h} = (y_{1t}, y_{2t} : t =$

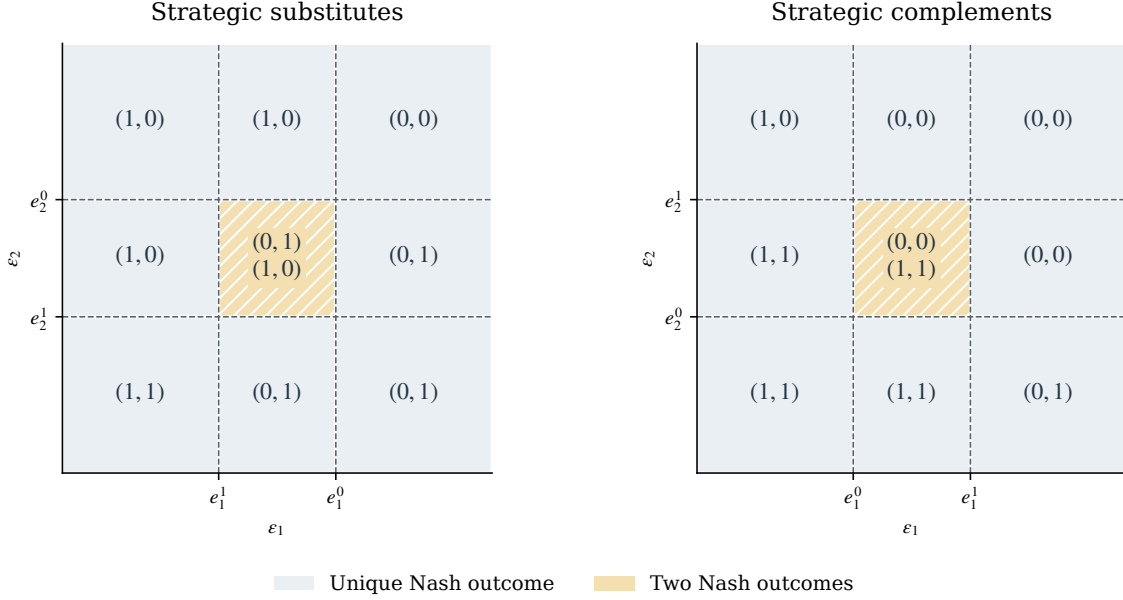


Figure 1: Equilibrium regions conditional on the state and continuation values
 Axes are the current payoff-shock differences. Labels give simultaneous Nash equilibrium outcomes. In the hatched region, both displayed outcomes are equilibria; elsewhere the displayed outcome is unique. The left panel has $\kappa_1, \kappa_2 < 0$ and the right panel has $\kappa_1, \kappa_2 > 0$.

$0, 1, \dots, T$), conditional on the initial condition (y_{10}, y_{20}) and incidental parameters, the sharp lower and upper probability bounds are obtained by multiplying the corresponding one-period bounds over the states visited along that history. Our CL-Bounds approach uses weaker bounds for tractability, retaining a product-of-logit-probabilities structure. The FD-Bounds approach accommodates the sharp one-period lower bounds directly, including their sums-of-products structure. Section 4 details how each construction eliminates the incidental parameters from the bounds.

3 Identification in Complete Models

In this section, we study the identification of Models A.1–A.4 in Table 1. In each of these models, the equilibrium is unique, so the probability of every possible equilibrium outcome is a well-defined function of the state variables and of the structural and incidental parameters.

Let $\theta = (\beta, \lambda, \gamma)$ denote the vector of structural parameters and $\eta = (a, \kappa)$ the vector of incidental parameters. For a joint history $\mathbf{h} = (y_0, y_1, \dots, y_T)$, let $p_h(\eta, \theta)$

denote the model-implied probability of $(\mathbf{y}_1, \dots, \mathbf{y}_T)$ conditional on the initial state $\mathbf{y}_0$ and the incidental parameters $\boldsymbol{\eta}$. The corresponding observable population probability averages over the distribution of incidental parameters across markets conditional on $\mathbf{y}_0$:

$$P_{\mathbf{h}} = P(\mathbf{h} \mid y_0) = \int p_{\mathbf{h}}(\boldsymbol{\eta}, \boldsymbol{\theta}) dF(\boldsymbol{\eta} \mid y_0). \quad (8)$$

Comparisons always use the same initial condition unless a separate argument justifies pooling. This ensures that we integrate over the same incidental distribution. Empirical frequencies estimate consistently $P_{\mathbf{h}}$.

3.1 Conditional likelihood approach

Proposition 1 (Exact incidental cancellation). *For any Models A.1-A.4:*

- (a) Factorization. *There exist vectors of statistics $s(\mathbf{h})$ and $r(\mathbf{h})$, and a function g such that for every history $\mathbf{h}$,*

$$\ln p_{\mathbf{h}}(\boldsymbol{\eta}, \boldsymbol{\theta}) = s(\mathbf{h})' g(\boldsymbol{\eta}, \boldsymbol{\theta}) + r(\mathbf{h})' \boldsymbol{\theta}. \quad (9)$$

- (b) Exact ratio. *For any pair of histories A, B with $s(A) = s(B)$, $P_A > 0$, $P_B > 0$,*

$$\frac{P_A}{P_B} = e^{[r(A) - r(B)]' \boldsymbol{\theta}} \quad (10)$$

- (c) Identification. *Comparisons of pairs of histories (A, B) as in part (b) identify exactly the linear combinations of $\boldsymbol{\theta}$ in the row span of the vectors $[r(A) - r(B)]'$. In particular, if these rows have full column rank, $\boldsymbol{\theta}$ is point identified.*

- (d) Conditional likelihood. *Conditional on $s(\mathbf{h}) = s$,*

$$P(\mathbf{h} \mid s) = \frac{e^{r(\mathbf{h})' \boldsymbol{\theta}}}{\sum_{\mathbf{h}': s(\mathbf{h}') = s} e^{r(\mathbf{h}')' \boldsymbol{\theta}}}, \quad (11)$$

which does not depend on the incidental parameters $\boldsymbol{\eta}$.

Proof. Parts (b)–(d) follow from the factorization in part (a) alone. The specific

structure of Models A.1–A.4 matters only for part (a), which we verify model by model at the end of the proof.

A key identity. Suppose (9) holds. The term $e^{r(\mathbf{h})'\boldsymbol{\theta}}$ does not depend on $\boldsymbol{\eta}$. So exponentiating (9) and integrating over $F(\boldsymbol{\eta} \mid y_0)$ gives, for every history $\mathbf{h}$,

$$P_{\mathbf{h}} = e^{r(\mathbf{h})'\boldsymbol{\theta}} \Phi(s(\mathbf{h}), \boldsymbol{\theta}), \quad \Phi(s, \boldsymbol{\theta}) \equiv \int e^{s'g(\boldsymbol{\eta}, \boldsymbol{\theta})} dF(\boldsymbol{\eta} \mid y_0). \quad (12)$$

The incidental parameters, and their distribution F , affect $P_{\mathbf{h}}$ only through Φ . Moreover, Φ depends on the history only through $s(\mathbf{h})$. Because $e^{r(\mathbf{h})'\boldsymbol{\theta}} \in (0, \infty)$, it follows that $P_{\mathbf{h}} > 0$ implies $0 < \Phi(s(\mathbf{h}), \boldsymbol{\theta}) = P_{\mathbf{h}} e^{-r(\mathbf{h})'\boldsymbol{\theta}} < \infty$.

Part (b). Let $s(A) = s(B) = s$ with $P_A, P_B > 0$. By (12), both probabilities contain the same factor $\Phi(s, \boldsymbol{\theta})$, which is strictly positive and finite. Dividing one by the other cancels it and delivers equation (10).

Part (c). Taking the logarithm of equation (10), we have $\ln P_A - \ln P_B = [r(A) - r(B)]' \boldsymbol{\theta}$. Stacking this equation over the set of pairs (A, B) satisfying the conditions of part (b) gives the linear system $\boldsymbol{\Delta} = R \boldsymbol{\theta}$, where $\boldsymbol{\Delta}$ collects the log-probability differences $\ln P_A - \ln P_B$ and R collects the corresponding rows $[r(A) - r(B)]'$. The vector $\boldsymbol{\Delta}$ and the matrix R are known. It follows that full column rank of matrix R gives a unique solution: $\boldsymbol{\theta} = (R'R)^{-1} R' \boldsymbol{\Delta}$. If that rank condition fails, the system identifies only its row span.

Part (d). Fix s with $\mathbb{P}\{s(\mathbf{h}) = s\} > 0$, and let $\mathbf{h}$ satisfy $s(\mathbf{h}) = s$. Then

$$P(\mathbf{h} \mid s) = \frac{P_{\mathbf{h}}}{\sum_{\mathbf{h}': s(\mathbf{h}') = s} P_{\mathbf{h}'}} = \frac{e^{r(\mathbf{h})'\boldsymbol{\theta}} \Phi(s, \boldsymbol{\theta})}{\Phi(s, \boldsymbol{\theta}) \sum_{\mathbf{h}': s(\mathbf{h}') = s} e^{r(\mathbf{h}')'\boldsymbol{\theta}}} = \frac{e^{r(\mathbf{h})'\boldsymbol{\theta}}}{\sum_{\mathbf{h}': s(\mathbf{h}') = s} e^{r(\mathbf{h}')'\boldsymbol{\theta}}}.$$

The second equality uses (12). The cancellation is valid because $\Phi(s, \boldsymbol{\theta}) > 0$ whenever the conditioning event has positive probability. The resulting expression depends on neither $\boldsymbol{\eta}$ nor $F(\cdot \mid y_0)$. This is the conditional-likelihood argument of Chamberlain (1980), applied to the general factorization (9). More generally, by the Fisher–Neyman factorization theorem, equation (9) implies that for fixed $\boldsymbol{\theta}$ the statistic $s(\mathbf{h})$ is sufficient for $\boldsymbol{\eta}$. This is the formal sense in which conditioning on s removes the incidental parameters.

Part (a). For the rest of this section, we show that each of Models A.1–A.4 admits the factorization (9) and derive the statistics $s(\mathbf{h})$ and $r(\mathbf{h})$ and the function g explicitly

for each model. □

3.1.1 Model A.1: No contemporaneous interaction

Consider the model with no contemporaneous interactions and with myopic players, $\gamma_i = 0$ and $c_{im} = 0$, for $i = 1, 2$:

$$y_{imt} = \mathbf{1} \{a_{im} + \beta_i y_{im,t-1} + \lambda_i y_{jm,t-1} - \varepsilon_{imt} \geq 0\}. \quad (13)$$

As shown in Appendix A, the likelihood of a history in this model, $p_h(\boldsymbol{\eta}, \boldsymbol{\theta})$, depends on the data only through the following statistics: y_0, y_T , and

$$N_i(\mathbf{h}) = \sum_{t=1}^T y_{it}, \quad N_{11}(\mathbf{h}) = \sum_{t=1}^T y_{1t}y_{2t}, \quad C_{ij}(\mathbf{h}) = \sum_{t=1}^T y_{it}y_{j,t-1}. \quad (14)$$

We show that the sufficient vector is $s(\mathbf{h}) = (y_0, y_T, N_1, N_2, N_{11})$. It fixes the number of visits to each lagged joint state, and hence every logit-denominator term, as well as the intercept contributions. The remaining statistic is $r(\mathbf{h}) = (C_{11}, C_{22}, C_{12}, C_{21})'$, associated with $(\beta_1, \beta_2, \lambda_1, \lambda_2)'$.

Corollary 1 (No contemporaneous interaction: CL identification). *In the myopic model with $\gamma = 0$, the coefficients (β, λ) are point identified by CL-SS for $T \geq 3$.*

Proof. In Appendix A. □

This identification result is established in the panel-logit literature, including Narendranathan et al. (1985); Honoré and Kyriazidou (2019). We use one running example to show its economic content. Let

$$\begin{aligned} A &= ((0,0), (0,0), (1,0), (1,0)), \\ B &= ((0,0), (1,0), (0,0), (1,0)). \end{aligned} \quad (15)$$

Player 2 always chooses zero. Player 1 has the same initial and terminal action and chooses one twice, but those choices are adjacent only in A . The two histories have the same s , C_{12} , and C_{21} , while $C_{11}(A) - C_{11}(B) = 1$. Therefore

$$\ln P_A - \ln P_B = \beta_1. \quad (16)$$

The comparison holds fixed the features that absorb permanent heterogeneity and changes the ordering that captures persistence. Appendix A gives pairs of histories A, B that identify the other parameters, β_2 , λ_1 , and λ_2 .

3.1.2 Models A.2-A.4: Contemporaneous interactions and sequential moves

Models A.2–A.4 depart from Model A.1 by allowing contemporaneous strategic interactions ($\gamma_i \neq 0$), forward-looking players ($c_{im} \neq 0$), or both. In each case, $\kappa_{im} \neq 0$. To guarantee a unique equilibrium, we assume that players move sequentially within each period: player 1 moves first, and player 2 moves after observing player 1's decision. The three models differ only in the composition of κ_{im} :

$$\kappa_{im} = c_{im} \text{ (Model A.2),} \quad \kappa_{im} = \gamma_i \text{ (Model A.3),} \quad \kappa_{im} = \gamma_i + c_{im} \text{ (Model A.4).}$$

This difference is immaterial for identification under the CL-SS approach. As shown in Appendix B, the three models share the same statistics $s(\mathbf{h})$ and $r(\mathbf{h})$. Proposition 1 therefore applies to all three models in identical form, with κ_{im} playing the same role in each.

Player 2 observes player 1's current action and responds to it, and player 1 anticipates this response. To describe the resulting outcome probabilities, define

$$F_i^z(s) \equiv \Lambda(a_i + \beta_i s_i + \lambda_i s_j + \kappa_i z), \quad z \in \{0, 1\},$$

the probability that player i chooses one when the opponent's current action is z . For player 2, z is player 1's observed action. For player 1, z is the response it anticipates from player 2. In what follows, we omit the lagged state s and the incidental parameters, and condition on both throughout.

Because both of player 2's possible responses are driven by the same shock, player 2 is one of three *response types*. Under strategic substitutes ($\kappa_2 < 0$, so $F_2^1 < F_2^0$), these types are:

- player 2 enters regardless of player 1's action (probability F_2^1);
- player 2 enters only if player 1 stays out (probability $F_2^0 - F_2^1$);
- player 2 stays out regardless (probability $1 - F_2^0$).

Under complements ($\kappa_2 > 0$), the middle type instead enters only if player 1 enters (probability $F_2^1 - F_2^0$). Under complete information, player 1 knows player 2's type and enters if its payoff, given player 2's response to entry, is nonnegative. Summing over types gives the probability $p_{y_1 y_2}$ of each period outcome $(y_{1t}, y_{2t}) = (y_1, y_2)$:

	Substitutes ($\kappa_2 < 0$)	Complements ($\kappa_2 > 0$)
p_{11}	$F_1^1 F_2^1$	$F_1^1 F_2^1$
p_{10}	$F_1^0 (1 - F_2^1)$	$F_1^0 (1 - F_2^1)$
p_{00}	$(1 - F_1^0)(1 - F_2^0)$	$(1 - F_1^0)(1 - F_2^1) + (1 - F_1^1)(F_2^1 - F_2^0)$
p_{01}	$(1 - F_1^1) F_2^1 + (1 - F_1^0)(F_2^0 - F_2^1)$	$(1 - F_1^1) F_2^0$

(17)

In each case, three of the four outcomes have the same product-logit form as in Model A.1. The only difference is that each player's logit is evaluated at the opponent's *realized current action*:

$$p_{y_1 y_2} = \underbrace{(F_1^{y_2})^{y_1} (1 - F_1^{y_2})^{1-y_1}}_{\text{player 1, at } z=y_2} \times \underbrace{(F_2^{y_1})^{y_2} (1 - F_2^{y_1})^{1-y_2}}_{\text{player 2, at } z=y_1}, \quad (y_1, y_2) \in \mathcal{Q}, \quad (18)$$

where $\mathcal{Q} = \{11, 10, 00\}$ for substitutes and $\mathcal{Q} = \{11, 10, 01\}$ for complements.

Let $\mathcal{H}_{\mathcal{Q}}$ denote the set of histories whose outcomes in periods $1, \dots, T$ all lie in $\mathcal{Q}$. As shown in Appendix B, when $\lambda_1 = \lambda_2 = 0$, on $\mathcal{H}_{\mathcal{Q}}$ the likelihood of a history admits the factorization (9) with

$$s(\mathbf{h}) = (y_0, y_T, N_1, N_2, C_{12}, C_{21}), \quad r(\mathbf{h}) = (C_{11}, C_{22})', \quad (19)$$

and $r(\mathbf{h})$ is associated with $(\beta_1, \beta_2)'$. If instead $\lambda_1 \neq 0$ or $\lambda_2 \neq 0$, the statistics C_{11} and C_{22} must also move into $s(\mathbf{h})$, so that $r(\mathbf{h})$ is empty: no statistic is left unconditioned on that varies with (β_1, β_2) , and CL-SS does not identify any parameter of the model. For the remainder of this subsection we therefore focus on the case $\lambda_1 = \lambda_2 = 0$, for which this structure has two important differences from Model A.1: statistics (C_{12}, C_{21}) are included in $s(\mathbf{h})$ instead of $r(\mathbf{h})$; and statistic N_{11} is not (explicitly) included in neither $s(\mathbf{h})$ or $r(\mathbf{h})$.

First, the logit denominators in (18) depend on the pair $(y_{i,t-1}, y_{jt})$, that is, own

lagged action and the opponent's current action, rather than on the lagged joint state. Fixing the number of periods with each such pair requires conditioning on the cross statistics C_{12} and C_{21} . These statistics therefore move from $r(\mathbf{h})$ to $s(\mathbf{h})$.

Second, conditioning on the set of histories $\mathcal{H}_Q$, then statistics N_{11} is fixed without having to explicitly condition on it. For substitutes, outcome 01 never occurs in $\mathcal{H}_Q$, such that $N_{11} = N_2$. For complements, outcome 00 never occurs in $\mathcal{H}_Q$, such that $N_{11} = N_1 + N_2 - T$.

This is why the distinction among Models A.2–A.4 does not matter for CL–SS. Proposition 1 applies to histories in $\mathcal{H}_Q$, and the remaining index is $\beta_1 C_{11} + \beta_2 C_{22}$.

Corollary 2 (Sequential move games: CL identification). *Consider Models A.2–A.4 with a known move order and nonzero action-interactions κ_{im} of the same sign in all markets. Then, for $T \geq 3$:*

- (i) *If $\lambda_1 = \lambda_2 = 0$, CL–SS on the restricted histories $\mathcal{H}_Q$ point identifies (β_1, β_2) .*
- (ii) *If $\lambda_1 \neq 0$ or $\lambda_2 \neq 0$, CL–SS on $\mathcal{H}_Q$ does not identify any parameter of the model.*

Proof. In Appendix B. □

We return to the running example (15). For substitutes, histories A and B belong to $\mathcal{H}_Q$, since their outcomes are 00 and 10. They have the same $s(\mathbf{h})$ in (19) and the same C_{22} , while $C_{11}(A) - C_{11}(B) = 1$. Therefore, exactly as in Model A.1, $\ln P_A - \ln P_B = \beta_1$. The analogous comparison for player 2 holds player 1 at one:

$$\begin{aligned} A' &= ((1, 0), (1, 0), (1, 1), (1, 1)), \\ B' &= ((1, 0), (1, 1), (1, 0), (1, 1)). \end{aligned} \tag{20}$$

Player 2 chooses one twice, and these choices are adjacent only in A' . Both histories lie in $\mathcal{H}_Q$, share the same $s(\mathbf{h})$ and C_{11} , and satisfy $C_{22}(A') - C_{22}(B') = 1$. Hence $\ln P_{A'} - \ln P_{B'} = \beta_2$.

For complements, outcome 00 is excluded, so A and B are not admissible. Holding player 2 at one gives the corresponding comparison for β_1 (see Appendix B).

3.2 Functional differencing beyond conditional likelihood

In this section, we show how a FD approach provides point identification of parameters λ in all the sequential-move models A2–A.4, and of parameters λ and γ in the

myopic sequential-move model A.3.

For a fixed initial state $\mathbf{y}_0$, collect the $J = 4^T$ post-initial history probabilities in the vector $\mathbf{p}_{\mathbf{y}_0}(\boldsymbol{\eta}, \boldsymbol{\theta})$. A vector $\mathbf{w}_{\mathbf{y}_0}(\boldsymbol{\theta}) \in \mathbb{R}^J$ is a uniformly valid FD equality if

$$\mathbf{w}_{\mathbf{y}_0}(\boldsymbol{\theta})' \mathbf{p}_{\mathbf{y}_0}(\boldsymbol{\eta}, \boldsymbol{\theta}) = 0 \quad \text{for every admissible } \boldsymbol{\eta}. \quad (21)$$

By integrating equation (21) over the distribution of $\boldsymbol{\eta}$ conditional on $\mathbf{y}_0$, we have the moment conditions:

$$\mathbf{w}_{\mathbf{y}_0}(\boldsymbol{\theta}_0)' \mathbf{P}_{\mathbf{y}_0} = 0 \quad (22)$$

where $\mathbf{P}_{\mathbf{y}_0}$ is the $J \times 1$ vector of population probabilities of choice histories given the initial condition $\mathbf{y}_0$.

Proposition 2 (Uniform equalities and identification). *If the distribution of $\boldsymbol{\eta}$ conditional on $\mathbf{y}_0$ is unrestricted and includes every point mass at $\mathbb{R}^4$ (Fixed Effects model):*

- (a) *A linear system of moment conditions $\mathbf{w}_{\mathbf{y}_0}(\boldsymbol{\theta})' \mathbf{P}_{\mathbf{y}_0} = 0$ is valid for every such distribution if and only if (21) holds for every $\boldsymbol{\eta}$.*
- (b) *A collection of moment equalities $\mathbf{w}_{\mathbf{y}_0}(\boldsymbol{\theta})' \mathbf{P}_{\mathbf{y}_0} = 0$ identifies $\boldsymbol{\theta}_0$ if its solution set is $\boldsymbol{\theta}_0$. If $\mathbf{w}_{\mathbf{y}_0}(\cdot)$ is differentiable, a full-column-rank Jacobian at $\boldsymbol{\theta}_0$ is a sufficient condition for local identification.*

Proof. Pointwise validity implies integrated validity. Conversely, integrate against a point mass at any admissible $\boldsymbol{\eta}$ to recover the pointwise equality. A unique common population solution must equal the true parameter because the true parameter satisfies every equality. A nonsingular square sub-Jacobian gives local uniqueness by the inverse function theorem. $\square$

Proposition 2 requires no logistic assumption. However, its practical value depends on our ability to find vectors $\mathbf{w}_{\mathbf{y}_0}(\boldsymbol{\theta})$ that satisfy (21). In the logit models of this paper this can be done exactly, because every history probability is a rational function of the exponentiated incidental parameters. We first describe this construction. Finally, we present separate results for the contemporaneous interactions γ and for the cross-lag parameters λ .

3.2.1 Constructing the uniform equalities

Notation. Recall that the per-period choice probabilities in the sequential game are built from $F_i^z(s) = \Lambda(a_i + \beta_i s_i + \lambda_i s_j + \kappa_i z)$, where $s = (s_1, s_2)$ is the lagged state and z is the opponent's current action. Define the exponentiated parameters

$$A_i = e^{a_i}, \quad B_i = e^{\beta_i}, \quad L_i = e^{\lambda_i}, \quad K_i = e^{\kappa_i},$$

with $K_i = G_i \equiv e^{\gamma_i}$ in Model A.3. Then

$$F_i^z(s) = \frac{A_i x_i(s, z)}{1 + A_i x_i(s, z)}, \quad x_i(s, z) = B_i^{s_i} L_i^{s_j} K_i^z. \quad (23)$$

In this subsection we treat $A = (A_1, A_2)$ as the vector of incidental parameters, which is the relevant case for Model A.3. The extension to Models A.2 and A.4, where $K = (K_1, K_2)$ is also incidental, is discussed in the next subsection and changes nothing in the logic below.

Step 1: Each kernel entry is a ratio of polynomials in A . Fix a lagged state s and let $u_i = A_i B_i^{s_i} L_i^{s_j}$ denote player i 's exponentiated index when the opponent's current action is zero. Each product-logit outcome in (18) has probability

$$p_{y_1 y_2} = \prod_{i=1,2} \frac{(u_i K_i^{y_j})^{y_i}}{1 + u_i K_i^{y_j}},$$

that is, a single monomial in A divided by a product of factors of the form $1 + A_i x_i(s, z)$. The exceptional outcome is a sum of two such terms. Over the common denominator $d(s) = (1 + u_1)(1 + u_1 K_1)(1 + u_2)(1 + u_2 K_2)$, its numerator is a polynomial with several monomials. Using the kernel for the exceptional outcomes, as defined by equations (17) and (18):

$$\begin{aligned} \text{Substitutes: } d(s) p_{01} &= u_2 [1 + (K_1 + K_2 - K_1 K_2) u_1 + K_2 u_2 + K_2 u_1 u_2], \\ \text{Complements: } d(s) p_{00} &= 1 + K_1 u_1 + K_2 u_2 + (K_1 + K_2 - 1) u_1 u_2. \end{aligned} \quad (24)$$

Two features of (24) will be important below. First, the exceptional outcome contributes several monomials, not one. Second, the coefficients of these monomials are linear in $(1, K_1, K_2, K_1 K_2)$.

Step 2: A common denominator turns history probabilities into polynomials.

The probability of a history h given y_0 is the product of T kernel entries, evaluated at successive lagged states. Every factor that can appear in a denominator has the form $1 + A_i x_i(s, z)$. There are at most eight such factors per player (four lagged states times two values of z), and each can appear at most T times. Hence

$$\mathcal{D}(A; \theta) = \prod_{i=1,2} \prod_{s \in \{0,1\}^2} \prod_{z \in \{0,1\}} (1 + A_i x_i(s, z))^T > 0 \quad (25)$$

is a common denominator for all histories. Multiplying by it clears every fraction, so $\mathcal{D}(A; \theta) p_h(A; \theta)$ is a polynomial in (A_1, A_2) whose coefficients depend on θ . Collecting the distinct monomials $A_1^{k_1} A_2^{k_2}$ in a vector $m(A)$, we can write, for each initial state,

$$\mathcal{D}(A; \theta) p_{y_0}(A; \theta) = \Gamma_{y_0}(\theta) m(A). \quad (26)$$

The matrix $\Gamma_{y_0}(\theta)$ has one row per history and one column per monomial. Row h contains the coefficients of the polynomial $\mathcal{D} p_h$.

Step 3: Uniform equalities are a left null space. Because $\mathcal{D} > 0$, condition (21) holds for all $A > 0$ if and only if $w' \Gamma_{y_0}(\theta) m(A) = 0$ for all $A > 0$. A polynomial that vanishes on an open set has all its coefficients equal to zero. Therefore

$$w_{y_0}(\theta)' p_{y_0}(A; \theta) = 0 \text{ for all } A > 0 \iff w_{y_0}(\theta)' \Gamma_{y_0}(\theta) = 0. \quad (27)$$

The set of all uniformly valid equalities for initial state y_0 is the left null space of $\Gamma_{y_0}(\theta)$, and its dimension is $J - \text{rank } \Gamma_{y_0}(\theta)$. The common denominator is needed to obtain polynomials, but it does not appear in the final restriction. In practice, Γ_{y_0} need not be expanded symbolically. Evaluating $p_{y_0}(A; \theta)$ exactly at more values of A than there are monomials, and computing the left null space of the resulting matrix, gives the same null space.

Three types of equalities. It is useful to distinguish three types of vectors in this null space.

- (i) *Trivial equalities.* Two histories that go through the same transitions $s \rightarrow y$ the same number of times have identical probabilities in *any* time-homogeneous

first-order Markov model. Their rows in Γ_{y_0} are identical, and the weights $(1, -1)$ do not depend on θ . For example, given $y_0 = (0, 0)$, $P_{11,00,00} = P_{00,11,00}$. These equalities carry no information about θ . For $T = 3$ there are three of them per initial state; for $T = 4$ there are 42.

- (ii) *Conditional-likelihood equalities.* Two rows of Γ_{y_0} are proportional, with a factor of proportionality that depends only on θ . These are exactly the pairwise comparisons of Proposition 1(b). For instance, with strategic substitutes, $\lambda = 0$, and the histories A and B of the running example (15),

$$p_A = \frac{B_1 A_1^2}{R(A)}, \quad p_B = \frac{A_1^2}{R(A)},$$

where $R(A) = (1 + A_1)^2(1 + A_1 B_1)(1 + A_2)(1 + A_2 K_2)^2$. Hence the row of A equals B_1 times the row of B , and the weights $w = (1, -B_1)'$ deliver (16).

- (iii) *Genuine FD equalities.* Histories that contain the exceptional outcome have rows with several monomials and interaction-dependent coefficients, by (24). Such a row is not proportional to any other row, but a combination of several rows can still cancel every coefficient. The weights of these combinations depend on the interaction parameters, so they can carry information that the restricted conditional likelihood, which discards these histories, does not use.

From equalities to identification. The dimension of the left null space counts the valid equalities. Whether these equalities identify θ depends on how the weights vary with θ , which is the Jacobian condition in Proposition 2(b). This Jacobian can be computed without differentiating the weights. Differentiating the identity $w(\theta)' p_{y_0}(\eta, \theta) = 0$ with respect to θ and integrating over $F(\eta | y_0)$ gives

$$\left. \frac{\partial w(\theta)' P_{y_0}}{\partial \theta'} \right|_{\theta_0} = -w(\theta_0)' \int \frac{\partial p_{y_0}(\eta, \theta_0)}{\partial \theta'} dF(\eta | y_0). \quad (28)$$

Let W_{y_0} be a basis of the left null space. Local identification holds if the matrix obtained by stacking $W_{y_0}' \int \partial p_{y_0} / \partial \theta' dF$ over initial states has full column rank. Its entries are rational functions of θ and of the support points and weights of F . Therefore, if it has full rank at one point, it has full rank everywhere outside a lower-dimensional set; we then say that the parameters are *generically* locally

identified. We use this property below to verify identification by exact computation at randomly drawn parameter values and mixing distributions.

3.2.2 Models A.2–A.4: which parameters are incidental

The three sequential models share the kernel (17), but they differ in what is incidental. This difference matters for FD, even though it does not matter for the restricted conditional likelihood.

- *Model A.3 (myopic players).* Here $\kappa_{im} = \gamma_i$ is common across markets. The incidental parameters are $\eta_m = (a_{1m}, a_{2m})$, and $G_i = e^{\gamma_i}$ are structural parameters. FD equalities are polynomial identities in (A_1, A_2) , exactly as above.
- *Models A.2 and A.4 (forward-looking players).* Here κ_{im} contains c_{im} , which depends on the market's continuation values and hence varies across markets with the fixed effects. Then κ_{im} is itself incidental, and $\eta_m = (a_{1m}, a_{2m}, \kappa_{1m}, \kappa_{2m}) \in \mathbb{R}^4$. FD equalities must hold for all (A, K) with K in the given sign regime, so they are polynomial identities in (A_1, A_2, K_1, K_2) . The construction is the same, with monomials in all four variables.

Two consequences follow. First, the question of identifying γ concerns Model A.3 only. In Model A.2, $\gamma = 0$ by construction. In Model A.4, γ_i enters the model only through $\kappa_{im} = \gamma_i + c_{im}$; with c_{im} unrestricted across markets, any change in γ_i can be offset by a change in c_{im} , so γ_i is not identified by any method. Second, the question of identifying λ concerns all three models, but the answers differ. When the interactions are common across markets, they impose restrictions that disappear when they vary across markets.

3.2.3 Identification of γ in Model A.3

In this subsection we maintain $\lambda_1 = \lambda_2 = 0$, so that β is identified by conditional likelihood (Corollary 2). The question is whether the full distribution of histories, including the histories with the exceptional outcome that conditional likelihood discards, identifies γ . We return to unrestricted λ at the end of the next subsection.

An explicit equality. Consider strategic substitutes and the initial state $y_0 = (0, 0)$. Denote a history by its three post-initial joint outcomes, and define

$$\begin{aligned} E_1 &= B_1 P_{10,00,11} - P_{00,10,11}, \\ E_2 &= B_1 P_{10,01,10} - P_{01,10,10}, \\ E_3 &= B_1 P_{10,01,11} - B_2 P_{01,10,11}. \end{aligned} \tag{29}$$

Substituting the kernel (17) and applying (27) shows that, for every (A_1, A_2) ,

$$B_2(G_2 - 1) E_1 + B_2 G_1(G_2 - 1) E_2 + (B_2 G_2 - 1) E_3 = 0. \tag{30}$$

Since (30) holds at every value of the fixed effects, it also holds after integrating over their distribution, by Proposition 2(a).

Three remarks help interpret this equality. First, E_1 compares two histories without the exceptional outcome, whereas E_2 and E_3 involve histories that contain the exceptional outcome 01. Second, without interactions ($G_1 = G_2 = 1$) and with $\beta_2 \neq 0$, the equality reduces to $E_3 = 0$, which is a conditional-likelihood comparison of Model A.1. Third, the equality is linear in the data, and its weights are polynomials in (G_1, G_2) of degree at most one in each argument. Given β , it is one equation in two unknowns, so it does not by itself identify (γ_1, γ_2) . The equality also requires $\lambda = 0$: it fails when $\lambda \neq 0$.

A class of equalities that is linear in the interactions. The structure of (30) is not a coincidence. By (24), the exceptional outcome enters with coefficients that are linear in

$$z(G) = (1, G_1, G_2, G_1 G_2)'. \tag{31}$$

This suggests looking for equalities whose weights are linear in $z(G)$, with coefficients that depend on β but not on G . For a fixed β and sign regime, we look for coefficient arrays $\omega_{rkh}(\beta)$, indexed by equality r , element k of z , and history h , such that

$$\sum_h \sum_{k=0}^3 \omega_{rkh}(\beta) z_k(G) p_h(A; \beta, G) = 0 \quad \text{for all } A > 0 \text{ and all } G \text{ in the regime.} \tag{32}$$

Multiplying by the common denominator and setting the coefficient of every monomial in (A_1, A_2, G_1, G_2) to zero gives a finite, homogeneous linear system in the unknowns ω , as in Step 3. The displayed equality (30) belongs to this class. Expanding its weights gives the coefficients of $(1, G_1, G_2, G_1 G_2)$:

$$(-B_2 E_1 - E_3, -B_2 E_2, B_2 E_1 + B_2 E_3, B_2 E_2).$$

This class does not need to contain all FD equalities. Its advantage is that each equality in it becomes a *linear* restriction on $z(G)$. For each equality r , define the observable coefficients

$$\mathcal{C}(\mathbf{P}, \beta)_{rk} = \sum_h \omega_{rkh}(\beta) P_h, \quad (33)$$

where P_h is the population probability of history h . Integrating (32) over the distribution of the fixed effects gives $\mathcal{C}(\mathbf{P}, \beta_0) z(G_0) = \mathbf{0}$. The matrix $\mathcal{C}$ depends on the data and on β , which is identified, but not on the unknown interactions. Rows can be stacked across initial states.

Proposition 3 (Global identification of the interactions). *Consider Model A.3 with $\lambda_1 = \lambda_2 = 0$, $T \geq 3$, and a known sign regime. Suppose β is identified and that equalities of the form (32) give $\text{rank } \mathcal{C}(\mathbf{P}, \beta_0) = 3$. Then γ_1 and γ_2 are globally point identified within that regime. Specifically, the unique vector in the null space of $\mathcal{C}$ whose first coordinate equals one is $z(G_0)$, and its second and third coordinates are G_{10} and G_{20} .*

Proof. Any interaction vector G that is compatible with the data must satisfy $\mathcal{C}(\mathbf{P}, \beta_0) z(G) = \mathbf{0}$, and the true vector satisfies it by construction. Because $z(G_0) \neq \mathbf{0}$ always lies in the null space, the rank of $\mathcal{C}$ is at most three; rank three is the maximum possible and makes the null space one-dimensional. A one-dimensional space intersects the hyperplane $\{z_0 = 1\}$ in a single point, which must be $z(G_0)$. Its second and third coordinates give G_{10} and G_{20} , and $\gamma_{i0} = \ln G_{i0}$. The argument rules out every alternative value of G , not only nearby values, so identification is global. $\square$

The rank condition is a restriction on the population distribution of histories. It is not implied by the number of available equalities. We verified it by exact computation at $T = 3$. For each initial state and each sign regime, the equalities in (32) computed from that initial state alone yield $\text{rank } \mathcal{C} = 3$ at randomly drawn

values of (β, G) and of the mixing distribution. By the genericity argument following (28), the rank condition therefore holds generically, and it can fail only on a lower-dimensional set of parameter values and mixing distributions.

3.2.4 Identification of λ

By Corollary 2, conditional likelihood never identifies λ in the sequential game. Moreover, once $\lambda \neq 0$, it no longer identifies β either. We now ask what the complete set of FD equalities identifies, using the construction in (27) and the Jacobian (28).

Proposition 4 (FD identification with cross-lag effects). *Consider Models A.2–A.4 with fixed effects, a known sign regime, and the kernel (17). The following statements hold at generic parameter values and mixing distributions.*

- (a) Short panels with unrestricted cross-lags. *If $T = 3$ and λ_1, λ_2 are unrestricted, the only uniformly valid equalities are the trivial ones of type (i). No FD equality is informative about β , λ , or γ .*
- (b) Model A.3, one cross-lag. *If $T = 3$ and $\lambda_2 = 0$, the FD equalities locally identify $(\beta_1, \beta_2, \lambda_1, \gamma_1, \gamma_2)$. The case $\lambda_1 = 0$ is symmetric.*
- (c) Model A.3, both cross-lags. *If $T = 4$ and λ_1, λ_2 are unrestricted, the FD equalities locally identify all six parameters $(\beta_1, \beta_2, \lambda_1, \lambda_2, \gamma_1, \gamma_2)$.*
- (d) Models A.2 and A.4. *When κ_{im} varies across markets, FD equalities add nothing to conditional likelihood in short panels. They identify β when $\lambda_1 = \lambda_2 = 0$ (for $T = 3$ and $T = 4$), and β_i alone when $\lambda_j = 0$ (for $T = 3$). When λ_1, λ_2 are unrestricted, the only uniformly valid equalities are the trivial ones, both for $T = 3$ and for $T = 4$. For $T = 5$, a few nontrivial equalities appear (between one and five per initial state), and they locally identify $(\beta_1, \beta_2, \lambda_1, \lambda_2)$.*

Proof. The proof is computational and has two steps. First, for each initial state, we compute a basis of the left null space in (27) by exact evaluation of the history probabilities. For Models A.2 and A.4, the evaluation points include (K_1, K_2) . Second, we compute the rank of the stacked matrix in (28) at randomly drawn parameter values and mixing distributions, using exact arithmetic modulo a large prime.

For (a) and the last statement in (d), the computed null spaces have the same dimension as the space of trivial equalities (three per initial state for $T = 3$ and 42 for $T = 4$). Exact arithmetic cannot understate the dimension of a null space, and the trivial equalities are always valid, so the two spaces coincide. For (b), (c), and the remaining statements in (d), including the case $T = 5$, the stacked Jacobian has the stated rank at the drawn point. Because its entries are rational functions, it then has that rank generically. Appendix ?? gives details and code. $\square$

Table 2 summarizes the results of this section.

Table 2: Identification in the sequential game with fixed effects

Model	Cross-lags λ	T	Conditional likelihood	All FD equalities
A.3	$\lambda_1 = \lambda_2 = 0$	3	β	β, γ
A.3	$\lambda_2 = 0$	3	β_1	β, λ_1, γ
A.3	unrestricted	3	none	none
A.3	unrestricted	4	none	β, λ, γ
A.2, A.4	$\lambda_1 = \lambda_2 = 0$	3, 4	β	β
A.2, A.4	$\lambda_2 = 0$	3	β_1	β_1
A.2, A.4	unrestricted	3, 4	none	none
A.2, A.4	unrestricted	5	none	β, λ

Notes: “FD” refers to generic local identification by the complete set of uniformly valid equalities (Proposition 4). For γ in the first row, identification is global under the rank condition of Proposition 3. The case $\lambda_1 = 0$ is symmetric to $\lambda_2 = 0$, with the indices of the players exchanged. In Model A.4, γ is not identified because it enters only through $\kappa_{im} = \gamma_i + c_{im}$; in Model A.2, $\gamma = 0$.

Numerical illustration. Consider model A.3 with $\lambda_1 = \lambda_2 = 0$, $\beta = (\ln(3/2), \ln(4/5))$, $\gamma = (\ln(2/3), \ln(1/2))$, and $T = 3$. Conditional on each initial state, the fixed-effect pairs are $(\ln(1/2), \ln(4/3))$, $(\ln 2, \ln(2/3))$, and $(\ln 3, \ln(5/2))$, with probabilities 0.3, 0.4, 0.3. The initial states can have any positive probabilities. For the γ calculation we use the conditional history probabilities starting at 00. These discrete types specify the numerical population only and are not restrictions imposed by the identification method.

A numerical implementation of (32) with the four monomials in (31) gives a coefficient-evaluation rank of 235 out of 256 columns and 21 polynomial-weight identities. The nonzero singular values of $\mathcal{C}$ are approximately .02970, .001884,

and .000816; the remaining singular value is below 10^{-15} . Normalizing its null vector gives $(1, 2/3, 1/2, 1/3)$, recovering the two true interactions. Independent evaluations over a wider range of incidental values give a largest identity error below 4×10^{-16} . We then recover rational coefficient arrays at $B = (3/2, 4/5)$ and verify all 21 identities by exact polynomial cancellation. With the stated rational type probabilities, exact arithmetic also gives rank $\mathcal{C} = 3$ and the normalized null vector $(1, 2/3, 1/2, 1/3)$. Thus the rank condition is non-vacuous and holds exactly in this example; it is not inferred from a finite evaluation grid.

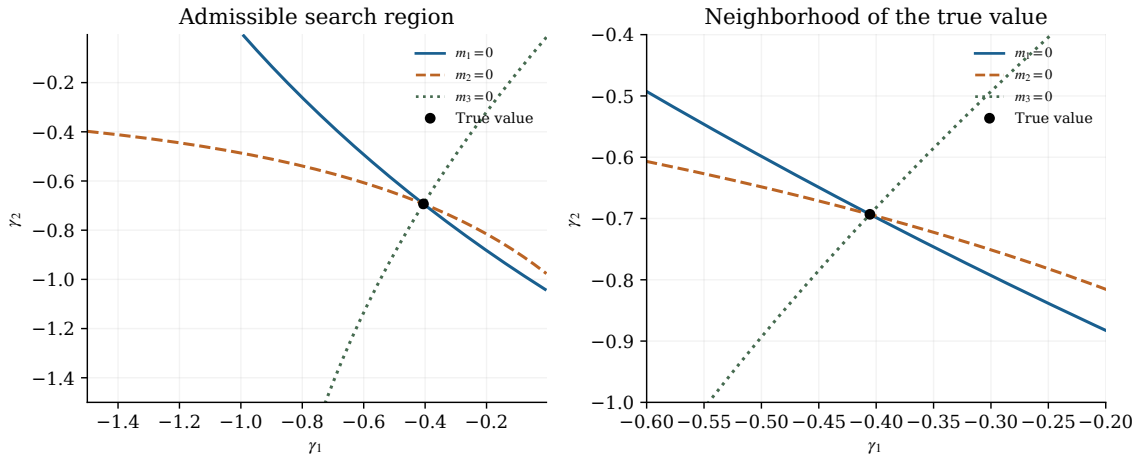


Figure 2: Sequential game: three FD zero contours and their common solution. Exact population probabilities. The curves use three independent linear combinations of the polynomial-weight moments. The point marks the true interactions. Plotting a finite region is illustrative; global uniqueness under Proposition 3 follows from the rank condition.

4 Identification in Incomplete Models

In an incomplete game, the observed conditional probability $p_h(\boldsymbol{\eta}, \boldsymbol{\theta})$ is unknown even at fixed incidental quantities because the selection rule is unspecified. Identification can nevertheless exploit bounds that hold for every admissible selection. We first match incidental terms across individual bounds. We then allow nonnegative combinations of bounds, which can use information unavailable to a pairwise comparison.

4.1 Sufficient-statistic bounds

Proposition 5 (Cancellation across probability bounds). *Consider two histories A, B with the same initial condition and bounds satisfying*

$$\begin{aligned} p_A(\boldsymbol{\eta}, \boldsymbol{\theta}) &\leq U_A(\boldsymbol{\eta}, \boldsymbol{\theta}), & p_B(\boldsymbol{\eta}, \boldsymbol{\theta}) &\geq L_B(\boldsymbol{\eta}, \boldsymbol{\theta}) > 0, \\ \ln U_A - \ln L_B &= d(A, B)' \boldsymbol{\theta} + q(A, B)' g(\boldsymbol{\eta}, \boldsymbol{\theta}). \end{aligned} \quad (34)$$

If $q(A, B)' g(\boldsymbol{\eta}, \boldsymbol{\theta}) \leq 0$ uniformly over admissible incidental quantities, then

$$P_A \leq \exp\{d(A, B)' \boldsymbol{\theta}_0\} P_B. \quad (35)$$

In particular, exact matching of all incidental components gives this result with $q = 0$. A sufficient signed-matching condition is zero coefficients on unrestricted components, non-negative coefficients on nonpositive components, and nonpositive coefficients on nonnegative components.

Proof. The assumptions imply $p_A \leq U_A \leq e^{d' \boldsymbol{\theta}} L_B \leq e^{d' \boldsymbol{\theta}} p_B$ pointwise. Integrate over the common conditional distribution of incidental quantities. The signed-matching conditions make every term in $q' g$ nonpositive. $\square$

When both population probabilities are positive, (35) can be written as a log-ratio inequality. For estimation the probability form is preferable: it remains meaningful when an empirical cell is empty and is linear in the observed history frequencies. Interchanging A and B yields the corresponding lower comparison.

The statistics in this argument summarize the incidental contribution to a *bound*. They are generally not sufficient statistics for the actual history probability, and there need not be an available conditional likelihood. We therefore call the method *sufficient-statistic bounds* (SS-Bounds). The resulting restrictions define an outer region containing the identified set. The presence of inequalities does not prove that the identified set is non-singleton, and their validity does not establish sharpness.

Product-logit bounds and their statistics. For the remainder of this subsection impose $\lambda = 0$ and $\kappa_i \leq 0$. Given an outcome $y = (y_1, y_2)$, define

$$q(y \mid s, b) = \prod_{i=1}^2 \frac{\exp\{y_i(a_i + \beta_i s_i + \kappa_i b_i)\}}{1 + \exp\{a_i + \beta_i s_i + \kappa_i b_i\}}. \quad (36)$$

The upper bound uses $b^U(y) = (y_2, y_1)$. A lower bound agrees with this rule at 00 and 11 and, at either mixed outcome, may use $b = (0, 0)$ or $b = (1, 1)$. Both rectangles lie in the unique-outcome region, so either choice gives a valid lower bound. Different periods may use different valid rectangles.

For a history and such a bound assignment, let

$$D_i^{uv}(h, b) = \sum_{t=1}^T \mathbf{1}\{y_{i,t-1} = u, b_{it} = v\}, \quad K_i(h, b) = \sum_{t=1}^T y_{it} b_{it}. \quad (37)$$

The log of its product bound is

$$\sum_i \left[a_i N_i + \beta_i C_{ii} + \kappa_i K_i - \sum_{u,v} D_i^{uv} \ln(1 + e^{a_i + \beta_i u + \kappa_i v}) \right]. \quad (38)$$

This formula constructs all the needed statistics directly. In the myopic model, $\kappa_i = \gamma_i$ is common, so equal (N_i, D_i^{uv}) across an upper and a lower bound cancel every intercept and denominator term. The remaining coefficient vector is the difference in $(C_{11}, C_{22}, K_1, K_2)$. No long expansion of each likelihood is needed.

Corollary 3 (Myopic simultaneous games). *In the myopic simultaneous game with $\lambda = 0$ and $\gamma \leq 0$, every upper/lower pair with a common initial state and matching (N_i, D_i^{uv}) satisfies*

$$\ln P_A - \ln P_B \leq \sum_i [C_{ii}(A) - C_{ii}(B)] \beta_i + \sum_i [K_i(A, b^U) - K_i(B, b^L)] \gamma_i. \quad (39)$$

The intersection of these restrictions is a valid outer region for (β, γ) , and includes nontrivial restrictions for $T = 3$.

For example, consider

$$A = (00, 00, 11, 11), \quad B = (00, 01, 10, 11). \quad (40)$$

They have the same initial and final states and totals $N_1 = N_2 = 2$. The own-lag counts differ only for player 2: $C_{22}(A) = 1, C_{22}(B) = 0$. Matching the relevant upper and lower bounds gives

$$\beta_2 + \gamma_1 + \gamma_2 \leq \ln P_A - \ln P_B \leq \beta_2 + \gamma_2. \quad (41)$$

Interchanging the players gives $\beta_1 + \gamma_1 + \gamma_2 \leq \ln P_A - \ln P_{B'} \leq \beta_1 + \gamma_1$, where $B' = (00, 10, 01, 11)$. Both the own-lag and interaction coefficients must be interchanged. These inequalities constrain combinations of parameters; they are not point-identification formulas for switching costs. Appendix E gives a compact catalogue of matching conditions and a complete enumeration algorithm.

The running comparison (15) illustrates the change in identifying logic. Its terminal and intermediate mixed outcomes are no longer assigned exact probabilities under unrestricted simultaneous selection. Keeping the same action totals and endpoints therefore does not preserve the equality (16). The researcher must instead match an upper bound to a lower bound, possibly using a different pair such as (40).

Forward-looking players and signed incidental terms. When κ_{im} is market specific, the terms $\kappa_i K_i$ also become incidental components. Their sign can still be used. Write

$$\begin{aligned}\sigma_i(u, v) &= -\ln(1 + e^{a_i + \beta_i u + \kappa_i v}), \\ d_i &= \sigma_i(1, 0) - \sigma_i(0, 0), \quad e_i = \sigma_i(0, 1) - \sigma_i(0, 0), \\ f_i &= \sigma_i(1, 1) - \sigma_i(1, 0) - \sigma_i(0, 1) + \sigma_i(0, 0).\end{aligned}\tag{42}$$

Also set $B_i(h, b) = \sum_t b_{it}$ and $R_i(h, b) = \sum_t y_{i,t-1} b_{it}$. The bound in (38) equals

$$T \sum_i \sigma_i(0, 0) + \sum_i (a_i + d_i) N_i + \sum_i [\beta_i C_{ii} + d_i (y_{i0} - y_{iT}) + \kappa_i K_i + e_i B_i + f_i R_i]. \tag{43}$$

The first two terms have unrestricted signs and must be matched. If $\kappa_i \leq 0$, then $e_i \geq 0$, $\text{sgn}(d_i) = -\text{sgn}(\beta_i)$, and $\text{sgn}(f_i) = \text{sgn}(\beta_i)$ whenever both coefficients are nonzero. These signs follow from the monotonicity and negative cross derivative of $-\ln(1 + e^v)$; Appendix E gives the integral argument.

Proposition 6 (Bounds with heterogeneous effective interactions). *Suppose $\lambda = 0$, $\kappa_{im} \leq 0$ for each player and market, and the signs of the common coefficients β_i are known. Let A, B have the same initial state and the same (N_1, N_2) . Use an upper assignment for A and a valid lower assignment for B . Denote upper-minus-lower differences by Δ . If, for*

each i ,

$$-\text{sgn}(\beta_i)\Delta(y_{i0} - y_{iT}) \leq 0, \quad \Delta K_i \geq 0, \quad \Delta B_i \leq 0, \quad \text{sgn}(\beta_i)\Delta R_i \leq 0, \quad (44)$$

then

$$P_A \leq \exp\{\Delta C_{11}\beta_1 + \Delta C_{22}\beta_2\}P_B. \quad (45)$$

When a coefficient is zero, its associated incidental terms vanish and their sign conditions can be dropped. The intersection of these inequalities is an outer region for β in the effective-coefficient model and hence in any structural model satisfying these restrictions.

Proof. Subtract (43) for the two bounds. The unrestricted components cancel. Each remaining incidental contribution is nonpositive by (44). Proposition 5 applies. $\square$

For $\beta_1, \beta_2 \geq 0$, two explicit comparisons are

$$\begin{aligned} A_1 &= (10, 11, 00, 00), & B_1 &= (10, 01, 10, 00), \\ A_2 &= (01, 11, 00, 00), & B_2 &= (01, 10, 01, 00). \end{aligned} \quad (46)$$

They imply $\ln P_{A_i} - \ln P_{B_i} \leq \beta_i$. The proof uses effective substitutability, not merely $\gamma_i \leq 0$. More generally, Proposition 6 allows different known signs for the two own-lag coefficients. This matters for price-adjustment settings and for the numerical experiment below, where $\beta_1 < 0$.

The assumption $\kappa_{im} \leq 0$ is an equilibrium restriction. A conservative primitive sufficient condition is available if one has an independent uniform bound on the oscillation of integrated values: if $\max_y \bar{V}_{im}(y) - \min_y \bar{V}_{im}(y) \leq K_i$, then $|c_{im}| \leq 2\delta_i K_i$, and $\gamma_i + 2\delta_i K_i \leq 0$ ensures effective substitutability. We do not assert that such a uniform bound follows from unrestricted fixed effects. In the general results the effective-sign restriction is maintained directly.

4.2 Functional differencing of bounds

Proposition 7 (Nonnegative combinations of bounds). *For a fixed initial state, let $L_h(\boldsymbol{\eta}, \boldsymbol{\theta}) \leq p_h(\boldsymbol{\eta}, \boldsymbol{\theta}) \leq U_h(\boldsymbol{\eta}, \boldsymbol{\theta})$ hold under every admissible selection. Suppose nonnega-*

tive weights $\mu_h(\boldsymbol{\theta})$ and $\zeta_h(\boldsymbol{\theta})$ satisfy

$$\sum_h \mu_h L_h - \sum_h \zeta_h U_h \geq 0 \quad \text{for every admissible } \boldsymbol{\eta}. \quad (47)$$

Then the population restriction

$$\sum_h [\mu_h(\boldsymbol{\theta}_0) - \zeta_h(\boldsymbol{\theta}_0)] P_h \geq 0 \quad (48)$$

holds. Exact equality in (47) is a sufficient special case. Multiple valid lower bounds for the same history may be included, with its observable probability repeated accordingly.

Proof. Nonnegative weights preserve the individual bounds, giving $\sum_h \mu_h p_h \geq \sum_h \mu_h L_h \geq \sum_h \zeta_h U_h \geq \sum_h \zeta_h p_h$. Integration proves the claim. $\square$

A pairwise SS comparison is obtained by assigning positive weight to a single lower bound and a single upper bound. Allowing additional histories or sharper lower bounds can provide additional restrictions. We define the refined outer region as the intersection of the SS-Bounds region and the additional FD-Bounds inequalities. Refinement is therefore weakly valid by construction; strict improvement requires an example or an additional argument.

For logit bounds, clear a common positive denominator and write

$$DL = \underline{G}(\boldsymbol{\theta})m(A), \quad DU = \overline{G}(\boldsymbol{\theta})m(A). \quad (49)$$

Exact cancellation is equivalent to

$$\underline{G}(\boldsymbol{\theta})'\mu - \overline{G}(\boldsymbol{\theta})'\zeta = 0, \quad \mu \geq 0, \quad \zeta \geq 0. \quad (50)$$

This is a polyhedral cone, not an arbitrary linear null space. A basis of the null space need not have nonnegative entries and need not generate the cone. One can enumerate extreme rays, or impose a normalization such as $\mathbf{1}'(\mu + \zeta) = 1$ and solve a linear program to search for the most violated population inequality. The coefficient representation makes the restriction uniform in the fixed effects. A finite grid without a global certificate is only a numerical diagnostic.

The inequality version (47) searches more generally for a polynomial nonnegative over the incidental domain. Coefficientwise nonnegativity and suitable sum-of-

squares representations give sufficient certificates. In several incidental dimensions these convenient certificates need not exhaust all nonnegative polynomials. No claim of sharpness follows merely from constructing them. Appendix F gives the associated convex-feasibility characterization and its closure qualification.

An explicit additional inequality. For the myopic simultaneous model, $y_0 = 00$, and $B_i = e^{\beta_i}$, $G_i = e^{\gamma_i} \leq 1$, use the sharp one-period bounds in (??) and their history products. They satisfy

$$\begin{aligned} (1 - G_1)U_{00,00,11} + G_1U_{10,00,00} + G_1^2U_{10,00,01} + B_1G_1U_{10,00,11} \\ = G_1(L_{00,10,00} + L_{00,10,01} + L_{00,10,11}) \end{aligned} \quad (51)$$

for every fixed-effect pair. Every weight is nonnegative in the maintained sign regime. Thus replacing each bound by its observed history probability gives the inequality with the right side weakly larger than the left side. Its player-swapped counterpart is also valid. The replication code checks the displayed identity by symbolic cancellation after clearing denominators.

The same identity yields a useful uniform inequality when effective interactions vary across markets. At each fixed-effect/effective-interaction value, divide its probability inequality by $G_1 > 0$ and discard the nonnegative terms $(G_1^{-1} - 1)p_{00,00,11}$ and $G_1p_{10,00,01}$. Integrating only *after* this step gives

$$e^{\beta_1}P_{10,00,11} \leq P_{00,10,00} + P_{00,10,01} + P_{00,10,11} - P_{10,00,00}. \quad (52)$$

This restriction and its player-swapped version hold for heterogeneous $\kappa_m \leq 0$ and either sign of β_i . They therefore apply to forward-looking players without estimating their continuation values. A positive denominator permits an explicit upper bound on β_1 by taking the logarithm of the right-side ratio. The probability form remains the primitive moment restriction.

Numerical gains with myopic players. We reproduce the workshop experiment with $T = 3$, initial state 00, $\beta = (-1, .4)$, $\gamma = (-.2, -.2)$, and myopia. The fixed-effect pairs (1.5, 1.5) and (2.5, 2.5) each have probability one half. In the multiplicity rectangle, the data-generating process selects each of 01 and 10 with probability one half, independently over time conditional on the state. Identification does not

impose this selection rule.

We enumerate the histories and valid product-logit lower assignments. Exact matching gives 12 nontrivial SS inequalities for this initial state. Linear programming computes their coordinate projections. Adding (51) and its player-swapped counterpart gives the refined region in Table 3. The reported refined extrema are numerical optimizations with repeated starts; they are outer-region projections, not confidence intervals or claims of structural sharpness. The supplied code reproduces the slide values to their reported precision.

Table 3: Myopic simultaneous game: numerical refinement of outer regions

Parameter	Truth	SS-Bounds	SS + two FD-Bounds	Width reduction
β_1	-1.000	$[-1.356, -0.635]$	$[-1.356, -0.945]$	42.9%
β_2	0.400	$[0.005, 0.765]$	$[0.005, 0.413]$	46.4%
γ_1	-0.200	$[-0.405, -0.039]$	$[-0.405, -0.039]$	0.0%
γ_2	-0.200	$[-0.355, 0.000]$	$[-0.355, 0.000]$	0.0%

Exact population probabilities. The parameter search maintains $\gamma_1, \gamma_2 \leq 0$; the own-lag coefficients are unrestricted. These are coordinate projections of the specified outer regions. Adding two inequalities does not exhaust the FD-Bounds cone.

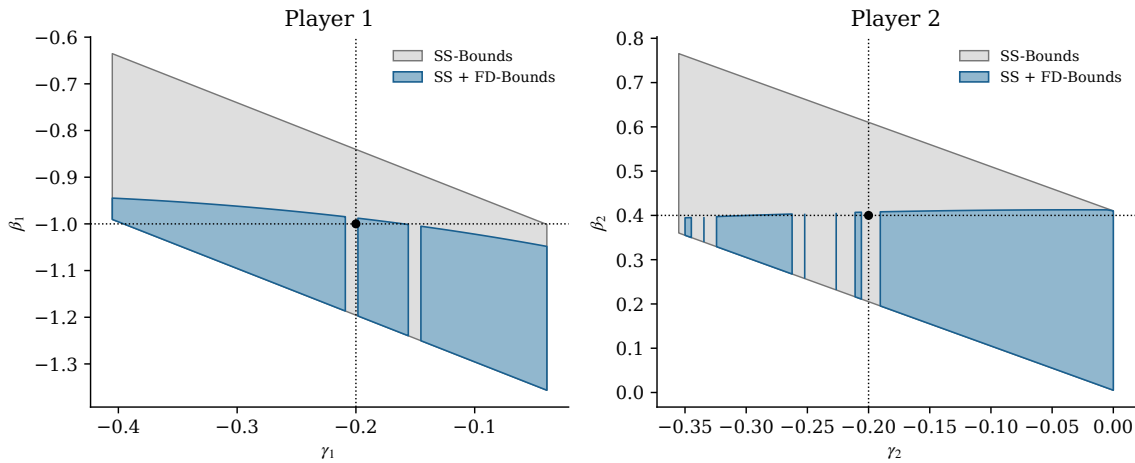


Figure 3: Myopic simultaneous game: joint profiles before and after FD-Bounds

For each plotted interaction coefficient, the remaining coefficients are allowed to vary subject to the specified inequalities. Gray denotes SS-Bounds and blue their intersection with the two additional FD-Bounds restrictions. Dots indicate the true values. The curves are numerical profile envelopes, not confidence bands.

Forward-looking numerical illustration. We use the same flow parameters, fixed-effect types, and selection rule, with $\delta_1 = \delta_2 = .8$. The integrated Bellman equations are solved separately for each type using the normalized level payoffs in (??), now with simultaneous moves. This yields effective interactions approximately $(-.19815, -.19826)$ for the first type and $(-.19941, -.19944)$ for the second. Both satisfy the effective-sign assumption. The largest Bellman residual is below 3×10^{-11} . Solving the game here constructs a numerical population; the identifying restrictions themselves do not require this solution.

The negative β_1 is relevant. A theorem requiring both own-lag coefficients to be nonnegative would not justify this experiment. We use the signed version of Proposition 6 in the orthant $\beta_1 \leq 0, \beta_2 \geq 0$. Exhaustive matching within the specified product-bound family gives 12 inequalities. They give more information than the lower-bound comparisons alone used in the slides. Table 4 reports the expanded, reproducible comparison and then adds (52) and its mirror.

Table 4: Forward-looking simultaneous game: signed SS and uniform FD bounds

Parameter	Truth	Signed SS-Bounds	Signed SS + uniform FD-Bounds
β_1	-1.000	$[-1.392, -0.604]$	$[-1.392, -0.604]$
β_2	0.400	$[0.008, 0.748]$	$[0.008, 0.534]$

Exact population probabilities, conditional on $y_0 = 00$. The added uniform FD upper bounds are approximately $-.291$ and $.534$, respectively; the first is redundant relative to the expanded signed SS family. The lower endpoints agree with the slides. Their original lower-bound-only SS comparison and different reported FD refinement are not the comparison family used in this table. All sets are outer regions under heterogeneous effective substitutability, not sharp sets for the primitive dynamic game.

These experiments distinguish the value of a stronger method from the value of using more comparisons within an existing method. FD-Bounds substantially tightens both myopic own-lag projections. In the forward experiment, extending signed SS comparisons already supplies an informative upper bound for the negative own-lag coefficient, while the uniform FD inequality further tightens the positive coefficient. Neither finding requires identifying the distribution of fixed effects or treating a primitive sign as a proved effective sign.

5 Empirical Applications

We use three panels of interacting decision makers: employment within couples, price adjustment by neighboring gasoline stations, and cooperation in a repeated-game lab experiment. Each records reversible binary decisions by many pairs of players. Their differences are useful: household employment has persistent outcomes, gasoline prices provide frequent adjustments, and the experiment has precisely observed choices and known incentives. We report the sample frequency of the choice histories underlying each estimate, since even a large panel may contain few observations that are informative for a particular identification argument.

For each dataset, we estimate several specifications. We begin with Model A.1, which assumes myopic players and no contemporaneous interactions ($\gamma = 0$). In this model, CL point identifies all structural parameters, (β, λ) . We then estimate a restricted version of Model A.3, with myopic players and sequential moves, which allows contemporaneous interactions but imposes $\lambda = 0$. In this specification, CL point identifies β , while FD additionally point identifies γ . Comparing these two specifications is informative because they differ primarily in the timing of strategic interactions: lagged in A.1 and contemporaneous in the restricted version of A.3. Finally, we consider Models B.2 and B.3, with myopic and forward-looking players, respectively. Both feature simultaneous moves and contemporaneous interactions, and we impose $\lambda = 0$ in each specification.

Conditional likelihood. Each estimation window has four observations, so $T = 3$. We enumerate the 256 possible joint histories, use $s(\mathbf{h}) = (y_0, y_T, N_1, N_2, N_{11})$ for Model A.1, use $s(\mathbf{h}) = (y_0, y_T, N_1, N_2, N_{11}, C_{12}, C_{21})$ for Model A.3, and maximize in β the corresponding conditional likelihood.

Sequential FD estimation. For each same-sign interaction regime we first estimate β from the restricted-history CL in Corollary 2, and then estimate γ using all history frequencies.

For these empirical estimates, the null spaces are recovered numerically from likelihood evaluations on a 17×17 Cartesian grid in $[-6, 6]^2$, retaining the generic dimension for each initial state and sign regime. Checks at independent fixed effects and a 21×21 grid in $[-7, 7]^2$ assess numerical stability. The polynomial coef-

ficient construction in Section 3 supplies the exact theoretical basis; the numerical implementation is an approximation to that space.

For an interior solution, let $H_{\beta\beta}$ be the negative first-stage score derivative and $H_{\gamma\gamma}, H_{\gamma\beta}$ the second-step derivative blocks. We use the actual curvature of the minimum-distance objective. The local influence functions have the form

$$\text{IF}_{\beta,m} = H_{\beta\beta}^{-1}s_{\beta,m}, \quad \text{IF}_{\gamma,m} = -H_{\gamma\gamma}^{-1}[\zeta_{\gamma,m} + H_{\gamma\beta}\text{IF}_{\beta,m}], \quad (53)$$

where $\zeta_{\gamma,m}$ includes the contribution of both history frequencies and initial-state weights. We aggregate these contributions within sampling clusters. Reported sequential standard errors include first-stage uncertainty; finite-difference steps .003 and .006 give similar results at the stable interior solutions. Boundary searches and separated first stages are reported as diagnostics, with conventional interaction standard errors suppressed.

Selection-robust bounds. We implement the signed SS construction in each of the four own-lag orthants, as well as (52) and its player-swapped counterpart. For the latter, define the count n_D of (00, 10, 00, 11) and

$$n_R = n(00, 00, 10, 00) + n(00, 00, 10, 01) + n(00, 00, 10, 11) - n(00, 10, 00, 00).$$

The sample analog is $e^{\beta_1}n_D \leq n_R$. If both counts are positive, $\ln(n_R/n_D)$ is a plug-in upper endpoint. Otherwise we retain the probability inequality and report the zero-cell or negative-numerator diagnostic. No pseudo-count is introduced. The restrictions apply to heterogeneous effective substitutes and permit forward-looking behavior, but maintain $\lambda = 0$. Sample inequalities and their intersections are not confidence regions: sampling error can make their intersection empty even under a valid population model. Appendix G reports all component counts and this distinction explicitly. Calibrated inference for the full family requires a moment-inequality procedure with the appropriate dependence structure; see, for example, Cox and Shi (2023) for inference in moment-inequality models.

5.1 Employment Decisions within Couples

5.1.1 Model

A household is a market, and the two agents are its female and male partners. Decisions are made monthly. We assign player 1 to the female partner and player 2 to the male partner. The action $y_{imt} = 1$ means that person i was employed and at work during the month t ; all other valid labor-force statuses are coded zero. In particular, employed individuals who were absent from work are coded zero. This distinction makes the outcome more reversible than an indicator of holding a job and aligns the definition with [Honoré et al. \(2023\)](#). A transition into zero therefore need not represent job destruction, and a subsequent return to one need not represent a new job match.

The incidental parameter α_{im} captures persistent differences in preferences, work incentives, labor-market opportunities, and household arrangements. A positive β_i means that working in the previous month raises the utility of working this month, holding fixed the partner's previous status and the individual effect. Potential mechanisms include employment continuity, adjustment costs, and the persistence of arrangements facilitating work. The coefficient does not separately identify entry and exit costs. The cross-lag coefficient λ_i captures how the partner's employment status in the previous month affects player i 's current incentive to work. It may reflect strategic substitutability or complementarity in employment through mechanisms that operate with a one-month lag (time-to-build).

The sequential specification replaces the lagged partner effect with a contemporaneous effect. A negative γ_i represents strategic substitutability in work, potentially consistent with income effects or household specialization. A positive coefficient represents complementarity in work incentives, potentially consistent with common scheduling or the complementarity of leisure.

The female-first and male-first versions are maintained within-month decision protocols. We treat each of these two sequential specifications as a robustness check.

5.1.2 Data

We construct the panel from the twelve public Basic Monthly CPS files for 2019, provided by the US Census Bureau ([US Census Bureau, 2019](#)). The sampling rules follow [Honoré et al. \(2023\)](#): one household reference person and one married or

unmarried partner of the opposite sex, with both partners aged 25–65 inclusive. The original study uses IPUMS data for 1982–2021; our sample is an independently constructed one-year public-file analogue and is not a replication of that original extract. The original paper is particularly relevant because it estimates a dynamic bivariate binary-response model with unobserved heterogeneity. Its contemporaneous association parameter belongs to a Schmidt–Strauss joint distribution and should not be identified with either directional payoff parameter in our sequential game. Related multivariate dynamic-logit methods are discussed by [Honoré and Kyriazidou \(2019\)](#).

Households in the CPS are interviewed for four consecutive months, leave the sample for eight months, and return for four additional months. We use only the first four interview months. Households are linked through the two public household identifiers, and the same partners are tracked using their person line numbers. We require stable sex and race, nondecreasing ages with changes of at most one year, the correct month-in-sample progression, and four consecutive calendar months. The resulting panel contains 16,155 couples. The first interview occurs between January and September 2019, so every retained four-month panel lies within the downloaded calendar year. We do not combine the first and second interview spells.

5.1.3 Descriptive Statistics

Table 5 presents descriptive statistics. Employment decisions exhibit substantial persistence: the frequency of no switching is 91.9% for males and 93.8% for females. In contrast, the pooled correlation between partners’ employment choices is relatively small, at 0.1529. These patterns may reflect structural state dependence and strategic interactions, but they may also arise from unobserved heterogeneity. Our objective is to disentangle these structural mechanisms from the dependence induced by unobserved heterogeneity.

In Appendix G, tables 20, 21, and 22 report the frequencies of individual histories, joint histories, and the transition probabilities, respectively. The most common joint history has both partners at work in all four months, (11, 11, 11, 11), and accounts for 47.77% of couples. A further 17.56% have a woman who is never at work and a man who is always at work, (01, 01, 01, 01). These observations help describe the distribution of employment but do not contribute to the conditional likelihood.

Table 5: Couples employment: sample and action statistics

Statistic	Value
Couples (households)	16,155
Four-observation histories	16,155
Observed types of joint histories (out of 256)	190
Informative histories	870
Female: fraction choosing 1 (working)	0.6786
Male: fraction choosing 1 (working)	0.8438
Female: standard deviation of action	0.4670
Male: standard deviation of action	0.3631
Male: frequency no-switching working status	0.9193
Female: frequency no-switching working status	0.9379
Pooled contemporaneous correlation	0.1529
Joint state 00 - No one working (percent)	7.61
Joint state 01 - Only Male (percent)	24.53
Joint state 10 - Only Female (percent)	8.01
Joint state 11 - Both working (percent)	59.85

Action means and joint-state frequencies pool all four observations in each window. Switching rates pool the three transitions.

The four individual histories highlighted by the identification analysis—0101, 1010, 1001, and 0110—occur 28, 42, 128, and 50 times for women, and 25, 34, 54, and 31 times for men. Their combined frequency is modest relative to the full sample.

Of the 256 logically possible joint histories, 190 occur in the CPS sample. More importantly, only 870 couples belong to conditional groups that are informative for the four-parameter baseline likelihood. Thus 5.39% of the sample contributes to the identification of the slopes through these comparisons. These numbers clarify why a substantially larger calendar span may improve precision, under the additional assumption that incidental parameters are constant across multiple years.

5.1.4 Estimation of Structural Parameters

(a.) CL estimates of β and λ The estimates of own-state dependence are positive for both partners: $\hat{\beta}_1 = 1.851$ for women and $\hat{\beta}_2 = 1.341$ for men. Their household-clustered standard errors are 0.136 and 0.154. The estimated cross-lag coefficients are negative: $\hat{\lambda}_1 = -0.355$ and $\hat{\lambda}_2 = -0.590$. This pattern is consistent with

Table 6: Couples Employment: Conditional Likelihood Estimates

Parameter	Model A.1 ($\gamma_1 = \gamma_2 = 0$)	Models A2-A4 ($\lambda_1 = \lambda_2 = 0$)
β_1	1.851*** (0.136)	1.884*** (0.153)
β_2	1.341*** (0.154)	1.468*** (0.184)
λ_1	-0.355* (0.200)	0 (-)
λ_2	-0.590*** (0.178)	0 (-)
Informative histories	870	561
Total histories	16,155	16,155

Standard errors in parentheses, with clustering as described in the text. Both columns condition out on $(y_0, y_T, N_1, N_2, N_11)$. The second column additionally conditions on the cross-lag statistics (C_{12}, C_{21}) and estimates only the two own-lag coefficients. No coefficient sign restriction is imposed.

lagged household substitution under the maintained model. Adding the two cross-lag statistics to the conditioning set yields $\hat{\beta}_1 = 1.884$ and $\hat{\beta}_2 = 1.468$. The relatively small changes in the own-state coefficients indicate that, within this class of models, the estimated degree of own-state dependence is robust to whether strategic interactions operate contemporaneously or with a one-month lag.

The complementary sequential specification produces $\hat{\beta}_1 = 2.049$, $\hat{\beta}_2 = 1.551$, $\hat{\gamma}_1 = 0.401$, and $\hat{\gamma}_2 = 1.199$. The corresponding clustered standard errors are 0.175, 0.240, 0.278, and 0.305. These are interior solutions, and changing the fixed-effect evaluation grid and expanding the interaction search range leaves the estimates unchanged to the reported precision. The first stage uses 444 informative histories. The estimate for the woman's contemporaneous interaction is considerably less precise than that for the man's; the point estimates should not be presented as equally strong evidence for two positive directional effects.

In the complementary case, reversing the move order gives the same probabilities after relabeling the coefficients. Indeed, $p_{10} = F_{10}(1 - F_{21})$, $p_{01} = (1 - F_{11})F_{20}$, and $p_{11} = F_{11}F_{21}$ under either order, and p_{00} is the remaining probability. Thus the identical estimates under the two orders are implied by this specification, rather than evidence that the data reveal who moves first.

The substitutes specification does not yield an interior interaction estimate in the initial search range. The follower's interaction reaches the upper numerical

Table 7: Couples employment: sequential-move estimates and diagnostics

Interaction	Order	β_1	β_2	γ_1	γ_2	First stage
Substitutes	1 first	1.110	1.516	-0.759	-0.050	167
				Boundary /	search-limited	
Substitutes	2 first	2.039	1.379	-0.050	-1.006	394
				Boundary /	search-limited	
Complements	1 first	2.049	1.551	0.401	1.199	444
		(0.175)	(0.240)	(0.278)	(0.305)	
Complements	2 first	2.049	1.551	0.401	1.199	444
		(0.175)	(0.240)	(0.278)	(0.305)	

All coefficients refer to the original player labels, including when player 2 moves first. The sequential model sets both cross-lag parameters to zero. The last column counts first-stage informative histories. For substitutes the search box for each interaction is $[-6, -0.05]$; for complements it is $[0.05, 6]$. Only the stable interior complementary estimates for CPS and gasoline receive standard errors, calculated from the two-step estimating equations with first-stage uncertainty and the appropriate clustering. Other displayed interaction values are constrained search outputs, not established interior structural estimates. A separated first stage invalidates the subsequent plug-in estimates, which are suppressed. Positive-interaction move orders are observationally equivalent in this normalized complete-information game.

boundary, -0.05 , under either move order. We report these outputs only as search diagnostics, without treating the boundary as an estimated strictly negative structural effect. The first-stage own-lag estimates remain finite, but differ across orders because the product-probability histories used for conditioning differ.

The interaction coefficients in the sequential model and the cross-lag coefficients in the baseline answer different questions. In particular, a negative estimate of λ_i and a positive estimate of γ_i are not contradictory: one concerns a partner's lagged action under the no-contemporaneous-interaction model, whereas the other concerns a partner's current action in a different restricted game. Neither model permits unrestricted common shocks or arbitrary time-varying household circumstances. Shared absences, vacations, and calendar effects are particularly relevant for the "at work" outcome. The empirical contribution is therefore to demonstrate feasible fixed-effect elimination and the additional information delivered by functional differencing, while making clear the economic restrictions under which these coefficients are structural.

Selection-robust evidence. The explicit FD upper-bound histories have no observations in either denominator cell. For player 1, the signed numerator count is two;

for player 2 it is minus one. Thus these particular inequalities produce no finite, usable plug-in upper endpoints. The signed-SS family also encounters zero cells. This is a support limitation of these sample comparisons, not a proof that employment state dependence is unidentified or a statistically calibrated rejection. The CL estimates use a different and more restrictive probability model; their precision does not eliminate the support problem for selection-robust inequalities.

5.2 Price Adjustment by Neighboring Gasoline Stations

5.2.1 Model and Data

The second application studies the timing of price adjustment by two nearby retail gasoline stations. A station is active in the binary sense when it changes its end-of-day posted diesel price relative to the preceding day. Thus $y_{imt} = 1$ denotes a price adjustment, not entry into the market, a temporary promotion, a price increase, or a high-price regime. Both upward and downward price changes count as adjustments. This coding produces a reversible action without imposing an arbitrary threshold on the price level.

The payoff index can be interpreted as the net benefit of adjusting the price today. A station-specific fixed effect absorbs persistent differences in its propensity to reprice, including aspects of pricing organization and location. The coefficient β_i describes how adjustment yesterday affects the current adjustment incentive. A negative value is consistent with a temporary reduction in the need to adjust after a recent repricing, whereas a positive value would indicate clustering of adjustments. This coefficient is not an entry or exit cost. The cross-lag coefficient λ_i measures a response to the rival's previous adjustment; γ_i in the sequential specification measures interaction in contemporaneous adjustment incentives. These coefficients concern the decision to adjust, not the level or magnitude of the new price.

A richer structural pricing model would include current own and rival prices, wholesale costs, demand conditions, and possibly the duration since the last adjustment. Our parsimonious binary specification leaves these states out. It is therefore useful for illustrating the identification procedures, but it is not a complete model of optimal price setting. The distinction is economically substantive: two stations can have the same binary adjustment history and different current price gaps. Price gaps can affect incentives even after persistent station heterogeneity has been eliminated.

The source is the French government’s public fuel-price reporting system, specifically its annual 2019 XML archive ([French fuel-price reporting system, 2019](#)). Its records identify stations, geographic coordinates, fuel types, price updates, reported shortages, and closures. The documentation explicitly excludes station names and brands. We study diesel (Gazole) and road stations, excluding motorway stations. The public archive is closely related to the data used by [Gautier and Saout \(2015\)](#), who study price adjustment and estimate a reduced-form model of state-dependent repricing using French daily station data. Their analysis makes the distinction between the adjustment decision and the price level particularly relevant here. Their sample and model differ from ours: we neither reproduce their estimates nor infer that a two-station binary game is their maintained specification.

We first retain 12,083 road stations with valid mainland coordinates and form disjoint mutual-nearest-neighbor pairs. The two stations must be at least 50 meters and at most five kilometers apart. For each member of a pair, the distance to the next station must exceed twice the within-pair distance. This rule is applied to the geographic station universe rather than to a sample selected for observed switching. It yields 1,444 candidate pairs. Requiring usable diesel-price histories produces 905 pairs in the main analysis sample. Station 1 is the station with the lower public identifier. This deterministic labeling is reproducible but has no intrinsic economic content.

For each station-day we retain the last reported diesel price and carry it forward to subsequent days. Repeated reports of the same price are not counted as adjustments. Quotes older than 30 days are excluded, as are reported closures and diesel stockouts. Coding a valid price-change indicator requires valid prices on both the current and preceding day. An intraday adjustment that is reversed before the end of the day does not generate $y = 1$. The archive’s price values are expressed in thousandths of euros; the binary comparison is invariant to this scale conversion.

January initializes the carried-forward prices. The estimation sample starts on February 1 and divides the remaining period into nonoverlapping four-day blocks, subject to complete valid coverage for both stations. This gives 61,512 histories. Each block conditions on its first observed binary action and contributes three modeled transitions. A second sample uses Monday–Thursday histories to examine sensitivity to the weekly calendar. Neither sample treats missing information as a zero once the coverage criteria fail.

The geographic construction is an operational definition of a pair, not proof of an isolated duopoly. A third station may matter beyond the distance cutoff, and two members may share ownership or centralized pricing. Since contemporaneous ownership is unavailable in this archive, the independent-agent interpretation remains a substantive limitation. The first-moving station is also not identified by the numeric station label. We examine alternative move orders as specification comparisons rather than assigning economic leadership to the lower identifier.

5.2.2 Descriptive Statistics

Table 8: Gasoline price adjustment: sample and action statistics

Statistic	Value
Four-observation histories	61,512
Distinct pairs / pair-matches	905
Primary sampling clusters	905
Observed joint histories (out of 256)	256
Informative baseline histories	14,323
Informative baseline pairs	856
Player 1: fraction choosing one	0.1720
Player 1: standard deviation of action	0.3774
Player 1: fraction switching per transition	0.2658
Player 2: fraction choosing one	0.1758
Player 2: standard deviation of action	0.3807
Player 2: fraction switching per transition	0.2709
Pooled contemporaneous correlation	0.2021
Joint state 00 (percent)	71.15
Joint state 01 (percent)	11.65
Joint state 10 (percent)	11.27
Joint state 11 (percent)	5.93
Median within-pair distance (km)	0.571
Mean histories per pair	67.97

Action means and joint-state frequencies pool all four observations in each window. Switching rates pool the three transitions. Primary clusters are households for CPS, station pairs for gasoline, and sessions for the experiment. Counts of histories are not counts of independent observations.

Stations 1 and 2 adjust on 17.20 and 17.58 percent of station-days represented in the analysis histories. The pooled correlation of their actions is 0.202. Both the frequency of adjustment and this positive correlation are descriptive properties of the chosen station pairs and coverage rule. They can reflect common input

costs, synchronized reporting, ownership, or strategic responses. None of these explanations is isolated by the unconditional correlation.

All 256 joint four-day histories occur. In contrast to employment, the all-zero history is particularly common: both stations make no end-of-day price adjustment throughout the block in 31.21 percent of histories. This represents inactivity in repricing, not station closure. Individual reversals are abundant. For station 1, the histories 0101, 1010, 1001, and 0110 occur 1,174, 1,207, 1,475, and 1,268 times, respectively. For station 2 the corresponding counts are 1,268, 1,294, 1,488, and 1,373. These counts suggest much richer support for short-history comparisons than in the other applications.

The joint transition table helps separate the persistence of inaction from the persistence of adjustment. Because the action itself is a change in a continuous variable, consecutive zeros need not indicate a stationary economic state, and consecutive ones need not mean that prices move in the same direction. The table should therefore not be interpreted as the transition matrix of a high-price/low-price regime model.

The full baseline conditional likelihood has 14,323 informative histories from 856 pairs, or 23.29 percent of all retained histories. The β -only conditional likelihood uses 3,595 informative histories. These observations are not independent replications of markets: the same 905 pairs contribute repeatedly during the year. The distinction is material for inference. Pair clustering allows arbitrary dependence among a pair's histories, while calendar-block clustering addresses shocks shared by pairs observed in the same four-day block. Neither correction removes omitted common shocks from the conditional mean.

5.2.3 Estimation of Structural Parameters

The baseline own-state estimates are $\hat{\beta}_1 = -0.217$ and $\hat{\beta}_2 = -0.267$. With two-way clustering by station pair and calendar block, their standard errors are 0.085 and 0.110. The corresponding conditional odds multipliers, approximately 0.805 and 0.766, describe a lower propensity to adjust immediately after a previous-day adjustment under the maintained model. They do not imply negative persistence in the level of prices. The estimated cross-lag effects are positive, at 0.329 and 0.394, consistent with a tendency to adjust after the neighboring station has done so.

Additional conditioning on the cross-lag statistics yields own-state estimates

Table 9: Gasoline price adjustment: no contemporaneous interactions

Parameter	Four-parameter CL	Additional conditioning
β_1	-0.217 (0.085)	-0.108 (0.111)
β_2	-0.267 (0.110)	-0.220 (0.139)
λ_1	0.329 (0.048)	Conditioned out
λ_2	0.394 (0.048)	Conditioned out
Informative histories	14,323	3,595
Total histories	61,512	61,512

Standard errors in parentheses, with clustering as described in the text. Both columns condition out the individual intercepts. The second column additionally conditions on the cross-lag statistics and estimates only the two own-lag coefficients. No coefficient sign restriction is imposed.

of -0.108 and -0.220 . Their uncertainty increases, particularly for the first station. This comparison illustrates the cost of relying on a smaller set of histories and shows that conclusions about the magnitude of short-run repricing persistence depend on how cross-lag heterogeneity is handled.

Table 10: Gasoline: sensitivity to the weekly calendar

Sample	histories	β_1	β_2	λ_1	λ_2
All calendar days	61,512	-0.217 (0.085)	-0.267 (0.110)	0.329 (0.048)	0.394 (0.048)
Monday–Thursday	34,785	-0.311 (0.075)	-0.419 (0.066)	0.268 (0.040)	0.325 (0.043)

Two-way pair/calendar-block clustered standard errors in parentheses. The weekday sample contains 904 distinct pairs. Both rows estimate the baseline model with unrestricted common cross-lag coefficients.

The Monday–Thursday sample contains 34,785 histories from 904 pairs and produces more negative own-state estimates, -0.311 and -0.419 . This difference is economically plausible in a setting with systematic variation in adjustment opportunities across the week. However, restricting the calendar is not equivalent to incorporating arbitrary day effects into the fixed-effect-free likelihood. A model with time-varying intercepts would generally require a new conditioning argument or a different set of moments.

Table 11: Gasoline price adjustment: sequential-move estimates and diagnostics

Interaction	Order	β_1	β_2	γ_1	γ_2	First stage
Substitutes	1 first	-0.125	-0.672	-0.050	-0.050	1,642
				Boundary /	search-limited	
Substitutes	2 first	-0.057	-0.195	-0.050	-0.050	1,831
				Boundary /	search-limited	
Complements	1 first	-0.464	-0.891	0.384	0.808	122
		(0.291)	(0.259)	(0.140)	(0.202)	
Complements	2 first	-0.464	-0.891	0.384	0.808	122
		(0.291)	(0.259)	(0.140)	(0.202)	

All coefficients refer to the original player labels, including when player 2 moves first. The sequential model sets both cross-lag parameters to zero. The last column counts first-stage informative histories. For substitutes the search box for each interaction is $[-6, -0.05]$; for complements it is $[0.05, 6]$. Only the stable interior complementary estimates for CPS and gasoline receive standard errors, calculated from the two-step estimating equations with first-stage uncertainty and the appropriate clustering. Other displayed interaction values are constrained search outputs, not established interior structural estimates. A separated first stage invalidates the subsequent plug-in estimates, which are suppressed. Positive-interaction move orders are observationally equivalent in this normalized complete-information game.

Under complementary contemporaneous interactions, the sequential model yields $(\hat{\beta}_1, \hat{\beta}_2) = (-0.464, -0.891)$ and $(\hat{\gamma}_1, \hat{\gamma}_2) = (0.384, 0.808)$. The respective two-way clustered standard errors, including first-stage uncertainty, are $(0.291, 0.259, 0.140, 0.202)$. The interaction estimates are interior and are stable under the denser fixed-effect grid and the expanded search box. As in the couples application, the two move orders imply the same complementary-game probabilities after preserving the original player labels.

Only 122 histories identify the own-lag coefficients in the first-stage complementary conditional likelihood. This is substantially less than the 14,323 informative histories in the baseline. The reduction occurs because histories containing a post-initial 00 outcome are excluded from the product-probability subset. Since no adjustment by either station is common, the restriction removes a large share of observed behavior. Functional differencing subsequently uses all history frequencies, but the two-step interaction estimates still inherit the uncertainty of this selective first stage.

In the substitutes specification, both interaction estimates approach the upper numerical boundary of the search range. The resulting constrained outputs should not be read as precise estimates of small negative interactions. Near zero interac-

tions, the generic likelihood rank becomes numerically delicate, providing a further reason to emphasize the well-separated interior complementary solution rather than infer economically meaningful negative parameters from the boundary search.

A positive contemporaneous interaction in adjustment does not establish that one station causally responds to the other's action. The same pattern could arise if both respond to unobserved wholesale-cost movements or if a common owner updates both prices. Moreover, the sequential specification removes the cross-lag coefficients that were positive in the baseline. The estimated γ 's therefore cannot be interpreted as effects holding unrestricted lagged competitive responses fixed. Before using this application for counterfactual pricing exercises, the state vector should be expanded and pair ownership verified. The present results establish the availability of useful histories and demonstrate implementation of the two identification methods under clearly stated restrictions.

Selection-robust evidence. The explicit FD inequalities have substantially better support here. Their denominator counts are 71 and 97 and their signed numerator counts are 735 and 523, giving plug-in upper endpoints of 2.337 and 1.685. These are loose restrictions, valid under a broader treatment of continuation values and selection but with zero cross-lag effects and effective substitutes. They are not confidence limits. Enumerating the entire specified signed-SS family at initial state 00 produces an empty sample intersection in each own-lag orthant. Unlike in the CPS, this is not caused by zero cells: every count in those comparisons is positive. The diagnostic highlights tensions among the maintained restrictions and the observed history frequencies, while a formal rejection would require sampling-error calibration. It would be misleading to present a single surviving inequality as a well-determined structural identified set. Appendix G provides the component counts and implementation details.

5.3 Cooperation in Repeated Prisoner's Dilemma Games

5.3.1 Model and Data

The third application uses the experiment of [DalBó and Fréchette \(2011\)](#). The binary action is cooperation, with defection coded zero. A market is a particular pair of subjects within one repeated match. A new match, even if it happens to involve the

same two subjects, is a separate interaction. We never join histories across a change of partner. Player 1 is the lower subject identifier; this labeling does not correspond to an experimental role or an observed first mover.

The experiment varies the payoff to mutual cooperation, $R \in \{32, 40, 48\}$, and the continuation probability, $\delta \in \{1/2, 3/4\}$, across six treatments. There are three sessions per treatment. The continuation rule randomly terminates matches, and subjects are rematched after a match ends. The actual stage game has simultaneous decisions. The data therefore offer a clean record of two-agent binary interaction, but do not implement our myopic sequential-move game. Estimates of that game below are an explicitly counterfactual timing benchmark rather than a recovery of the experiment's known institutional structure.

The material payoff matrix, in experimental points, is

$$\begin{array}{c|cc}
 & C & D \\
 \hline
 C & (R, R) & (12, 50) \\
 D & (50, 12) & (25, 25).
 \end{array} \tag{54}$$

Holding the other player's action fixed, the material payoff difference between cooperating and defecting is

$$-13 + (R - 37)Y_j. \tag{55}$$

It contains no own-lag or cross-lag term. Thus nonzero estimated β_i and λ_i in the behavioral benchmark cannot literally be read as lagged terms in the material stage payoff. They can summarize habit, reciprocity, beliefs, or continuation incentives under a reduced state representation. Likewise, the material contemporaneous difference coefficient equals -5 , 3 , or 11 points across treatments, not a common coefficient in normalized logit units.

There is an additional reason for care with the sequential comparison. Normalizing the payoff of action zero to zero by subtracting a function of the rival's action preserves simultaneous best responses to a fixed rival action. It need not preserve a leader's incentives when the rival's response changes with the leader's own action. Equation (??) is therefore a maintained behavioral payoff specification, not an innocuous sequential re-expression of the experimental payoff matrix. This distinction prevents a mechanical comparison between our estimated γ 's and

$R = 37$.

The original data and dictionary are publicly available from the authors and in the journal’s replication archive. The file contains 37,042 subject-round observations, 266 subjects, 18 sessions, and 6,944 pair-matches. Our main four-observation sample retains the first four rounds of matches lasting at least four rounds, giving 1,469 histories. Round 1 serves as the initial condition. The random stopping rule supports selection on match length independently of choices, but does not remove learning, history dependence beyond one lag, or differences across treatments.

The original study estimates mixtures of repeated-game strategies and a belief-based learning model. Those approaches provide a useful contrast to the fixed-effect binary transition model studied here: they attach structure to strategy choice and learning, whereas our baseline absorbs persistent propensities through unrestricted intercepts and restricts the dependence on observed histories. We do not interpret our parameters as the original study’s strategy probabilities or decision-noise parameter. The distinction also matters for pooling: unrestricted intercepts can absorb level differences across matches and treatments, but do not allow the structural slopes to differ unless the sample is explicitly split.

Because subjects appear in multiple matches, pair-matches are not independent markets for inference. We cluster at the session level throughout. Eighteen sessions provide much less independent information than the count of 1,469 eligible histories might suggest; within an individual treatment, there are only three clusters. Consequently, the experiment is most useful as a transparent demonstration of the methods and their finite-sample limitations, rather than as the strongest large- M application among the three datasets.

5.3.2 Descriptive Statistics

Cooperation rates in the retained four-round histories are 42.73 percent for player 1 and 42.27 percent for player 2. Their pooled contemporaneous correlation is 0.589, substantially larger than in the other applications. The correlation combines between-match heterogeneity, common treatment conditions, and within-match strategic dependence. Since subject labels are arbitrary, the small difference in marginal rates should not be interpreted as an economically meaningful asymmetry between player types.

Both subjects defect in all four rounds in 28.05 percent of eligible matches, and

Table 12: Experimental cooperation: sample and action statistics

Statistic	Value
Four-observation histories	1,469
Distinct pairs / pair-matches	1,469
Primary sampling clusters	18
Observed joint histories (out of 256)	131
Informative baseline histories	275
Informative baseline pairs	275
Player 1: fraction choosing one	0.4273
Player 1: standard deviation of action	0.4947
Player 1: fraction switching per transition	0.1582
Player 2: fraction choosing one	0.4227
Player 2: standard deviation of action	0.4940
Player 2: fraction switching per transition	0.1439
Pooled contemporaneous correlation	0.5888
Joint state 00 (percent)	47.45
Joint state 01 (percent)	9.82
Joint state 10 (percent)	10.28
Joint state 11 (percent)	32.45

Action means and joint-state frequencies pool all four observations in each window. Switching rates pool the three transitions. Primary clusters are households for CPS, station pairs for gasoline, and sessions for the experiment. Counts of histories are not counts of independent observations.

Table 13: Experiment: sample composition by treatment

R	δ	Eligible matches	Sessions	Player 1 (%)	Player 2 (%)
32	0.5	140	3	9.46	10.00
32	0.75	323	3	21.13	20.98
40	0.5	177	3	16.38	15.82
40	0.75	266	3	60.81	63.16
48	0.5	189	3	32.14	33.07
48	0.75	374	3	78.81	75.07

Cooperation percentages pool the first four rounds of matches lasting at least four rounds. The number of sessions is three in every treatment. The pooled analysis imposes common slope coefficients across treatments.

both cooperate in all four rounds in 26.48 percent. These two histories together account for more than half of the sample. Individual reversals nevertheless occur: the four histories 0101, 1010, 1001, and 0110 appear 19, 27, 18, and 22 times for player 1 and 23, 15, 11, and 22 times for player 2. The sample contains 131 of the 256 possible joint histories, leaving many cells empty.

The baseline conditional likelihood is informative for 275 matches, or 18.72 percent of the eligible sample. The β -only version uses 93. The sequential conditional likelihood can be more selective still because it restricts the set of outcomes within the history. In the complementary-interaction specification, only 12 first-four-round histories contribute to the first-stage estimation of the β 's. This is a direct warning about precision: the existence of identifying histories in the population does not ensure that they occur often in a particular experiment.

As an additional support check, we also construct all consecutive four-round histories lying entirely within a match. These histories overlap, retain a fixed pair, and continue to be clustered by session. They increase the number of observed histories without creating additional independent sessions. We present this extension separately so that richer use of long matches is not confused with a larger number of independent experimental markets.

5.3.3 Estimation of Structural Parameters

Table 14: Experimental cooperation: no contemporaneous interactions

Parameter	Four-parameter CL	Additional conditioning
β_1	1.472 (0.359)	2.094 (0.583)
β_2	1.956 (0.392)	2.144 (0.344)
λ_1	2.881 (0.332)	Conditioned out
λ_2	2.077 (0.337)	Conditioned out
Informative histories	275	93
Total histories	1,469	1,469

Standard errors in parentheses, with clustering as described in the text. Both columns condition out the individual intercepts. The second column additionally conditions on the cross-lag statistics and estimates only the two own-lag coefficients. No coefficient sign restriction is imposed.

The baseline estimates are $\hat{\beta}_1 = 1.472$ and $\hat{\beta}_2 = 1.956$, with session-clustered standard errors 0.359 and 0.392. The cross-lag estimates are also positive and relatively large: $\hat{\lambda}_1 = 2.881$ and $\hat{\lambda}_2 = 2.077$. Within the restricted transition model, this combination describes persistence in own cooperation and a positive response to the opponent's previous cooperation. It is consistent with reciprocal behavior, but does not distinguish tit-for-tat from other strategies, learning, or forward-looking incentives. Conditioning out the cross-lag coefficients yields $\hat{\beta}_1 = 2.094$ and $\hat{\beta}_2 = 2.144$ with less precision.

Table 15: Experiment: treatment-specific baseline estimates

R	δ	Informative	β_1	β_2	λ_1	λ_2	Joint MLE
32	0.5	20	—	—	—	—	Separated
32	0.75	77	1.461	2.585	2.772	2.588	Finite
40	0.5	34	—	—	—	—	Separated
40	0.75	48	2.234	2.712	4.399	2.623	Finite
48	0.5	31	0.941	0.794	0.261	1.535	Finite
48	0.75	65	1.723	1.239	3.077	1.074	Finite

All four slopes are estimated separately by treatment. A dash denotes absence of a finite joint conditional-likelihood maximum, not a zero coefficient. The separated directions concern λ_1 for $(R, \delta) = (32, 0.5)$ and λ_2 for $(40, 0.5)$. There are only three session clusters per row; treatment-specific significance comparisons are not emphasized.

Pooling the six treatments imposes common slope coefficients. The treatment-specific fits reveal that this restriction should not be interpreted as innocuous. Two four-parameter baseline specifications have no finite joint conditional-likelihood maximum: the $R = 32$, $\delta = 1/2$ treatment separates in λ_1 , and the $R = 40$, $\delta = 1/2$ treatment separates in λ_2 . We suppress their coefficient rows rather than replace infinite estimates with numerical stopping values. The remaining treatment estimates are also based on few informative observations and only three session clusters each. They are reported as descriptive sensitivity results, not as precise estimates of treatment heterogeneity.

The first-four-round sequential estimates do not support the same degree of precision as in the other applications. Under complementary interactions, the first-stage estimates are $\hat{\beta}_1 = 1.415$ and $\hat{\beta}_2 = 0.844$, based on only 12 informative histories. In the initial interaction search box, the minimum occurs near $(\gamma_1, \gamma_2) = (4.355, 6.000)$, with the second coefficient at its upper bound. Expanding the search range to $[0.02, 12]^2$ moves the numerical solution to approximately $(5.65, 8.90)$. We

Table 16: Experimental cooperation: sequential-move estimates and diagnostics

Interaction	Order	β_1	β_2	γ_1	γ_2	First stage
Substitutes	1 first	1.917	0.693	-0.050	-6.000	45
				Boundary /	search-limited	
Substitutes	2 first	—	—	—	—	39
		Separated	Separated	Not reported	Not reported	
Complements	1 first	1.415	0.844	4.355	6.000	12
				Boundary /	search-limited	
Complements	2 first	1.415	0.844	4.354	6.000	12
				Boundary /	search-limited	

All coefficients refer to the original player labels, including when player 2 moves first. The sequential model sets both cross-lag parameters to zero. The last column counts first-stage informative histories. For substitutes the search box for each interaction is $[-6, -0.05]$; for complements it is $[0.05, 6]$. Only the stable interior complementary estimates for CPS and gasoline receive standard errors, calculated from the two-step estimating equations with first-stage uncertainty and the appropriate clustering. Other displayed interaction values are constrained search outputs, not established interior structural estimates. A separated first stage invalidates the subsequent plug-in estimates, which are suppressed. Positive-interaction move orders are observationally equivalent in this normalized complete-information game.

report the latter as a provisional minimum-distance output, not as a well-established finite structural point estimate: the likelihood span is poorly conditioned in this region, and ordinary local standard errors would conceal the sensitivity to the search range and the weak first-stage support.

The all-window extension contains 5,316 overlapping four-round histories from the same 1,469 pair-matches and the same 18 sessions. It raises the number of first-stage informative histories for the complementary specification from 12 to 34. The corresponding first-stage estimates are (1.460, 0.879). Nevertheless, the interaction search remains sensitive: the initial-box output is approximately (4.097, 6.000), while the wider search gives approximately (5.38, 8.96) without meeting the optimizer's convergence tolerance. This extension therefore increases observed history support without resolving the instability of the contemporaneous estimates.

The substitutes specification is also problematic. With player 1 moving first, the initial search reaches interaction boundaries. With player 2 moving first, the first-stage conditional likelihood separates in an own-lag direction. Its numerical coefficient of roughly 31 is an optimizer stopping value, not a finite estimate. We suppress the resulting interaction estimates because they would be based on an invalid finite first-stage solution. No penalty or artificial pseudo-count is used to

conceal the separation.

The difficulty of estimating the contemporaneous coefficients in this sample is itself informative for empirical implementation. Many observations are concentrated in constant cooperation or constant defection histories, while the sequential conditional comparisons use a small fraction of the sample. In addition, the actual protocol is simultaneous and the experimental continuation incentives are explicit. A boundary estimate under a myopic sequential benchmark can reflect weak finite-sample information, misspecification, or both. It is not evidence that the population parameter is infinite, that the theoretical identification result fails, or that the original experiment lacks strategic interaction.

Selection-robust evidence. Both explicit FD denominator histories have zero observations, and each signed numerator count is minus two. The literal sample inequalities are therefore violated independently of β , but each violation involves only two net observations in a sample with 18 independent experimental sessions. We do not interpret this as a population impossibility or a formal specification rejection. The correct empirical conclusion is that these particular bounds are not estimated precisely enough to report finite structural endpoints. This accords with the sparse first-stage support in the sequential exercise, although the two methods use different restrictions and comparisons.

Taken together, the applications demonstrate three different empirical situations. Couples provide many distinct pairs but relatively sparse identifying reversals. Gasoline prices provide dense support for the short-history comparisons but require substantial additional structure for a pricing interpretation. The experiment offers precisely observed choices and known incentives, but its small number of independent sessions and concentration in constant histories limit the empirical content of the most restrictive specifications. These differences underscore the value of reporting the histories and conditional comparisons underlying each estimate, rather than only the total number of panel observations.

6 Conclusion

Short choice histories can distinguish state dependence from persistent heterogeneity because their ordering contains information beyond their marginal frequencies.

In a complete model, exact history comparisons can remove fixed effects and deliver explicit probability-ratio formulas. In a sequential game, functional differencing also uses histories omitted by a convenient conditional likelihood and can identify contemporaneous interactions under a separate global rank condition. These are distinct sources of identifying information, and neither the number of moments nor the numerical rank of a likelihood matrix should be substituted for an identification argument.

With multiple equilibria and unrestricted selection, the same organizing idea survives in a different form. incidental components can cancel between an upper bound for one history and a lower bound for another. Uniform signs can replace exact matching for selected continuation-value terms. Nonnegative combinations of bounds add further inequalities and substantially tighten the own-state projections in the myopic numerical example. These methods eliminate incidental continuation values from valid restrictions without requiring the researcher to solve the game during estimation. Their conclusions concern outer regions unless the remaining mixture and equilibrium-feasibility restrictions are also characterized.

The applications show why empirical implementation must report identifying histories, not just sample size. Useful reversals are common in gasoline adjustment but much less frequent in employment and experimental cooperation. Even with many observations, a restrictive first-stage comparison can be weak, and a selection-robust plug-in region can be empty. The resulting empirical task is to combine transparent probability restrictions with appropriate inference and a defensible state, timing, and shock specification. Extending the sharp feasibility analysis, establishing broader primitive conditions for effective-interaction signs, and developing the unresolved global short-panel results remain useful directions for further work.

References

- Aguirregabiria, V. and Carro, J. (2026). Identification of average marginal effects in fixed effects dynamic discrete choice models. *Review of Economics and Statistics*, Forthcoming. Available at https://doi.org/10.1162/rest_a_01521.
- Aguirregabiria, V., Gu, J., and Luo, Y. (2021). Sufficient statistics for unobserved heterogeneity in dynamic structural logit models. *Journal of Econometrics*, 223(2):280–311.
- Aguirregabiria, V. and Mira, P. (2007). Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53. Available at <https://doi.org/10.1111/j.1468-0262.2007.00731.x>.
- Aguirregabiria, V. and Mira, P. (2019). Identification of games of incomplete information with multiple equilibria and unobserved heterogeneity. *Quantitative Economics*, 10(4):1659–1701.
- Arcidiacono, P. and Miller, R. A. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79(6):1823–1867. Available at <https://doi.org/10.3982/ECTA7743>.
- Besanko, D., Doraszelski, U., Kryukov, Y., and Satterthwaite, M. (2010). Learning-by-doing, organizational forgetting, and industry dynamics. *Econometrica*, 78(2):453–508.
- Blevins, J. R. (2015). Structural estimation of sequential games of complete information. *Economic Inquiry*, 53(2):791–811.
- Bonhomme, S. (2012). Functional differencing. *Econometrica*, 80(4):1337–1385. Available at <https://doi.org/10.3982/ECTA9311>.
- Bresnahan, T. F. and Reiss, P. C. (1991). Empirical models of discrete games. *Journal of Econometrics*, 48(1–2):57–81.
- Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies*, 47(1):225–238.

- Chamberlain, G. (1985). Heterogeneity, omitted variable bias, and duration dependence. In *Longitudinal Analysis of Labor Market Data*. Cambridge University Press.
- Chamberlain, G. (2010). Binary response models for panel data: Identification and information. *Econometrica*, 78(1):159–168.
- Cox, G. and Shi, X. (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. Available at <https://doi.org/10.1093/restud/rdac015>.
- DalBó, P. and Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, 101(1):411–429. Available at <https://doi.org/10.1257/aer.101.1.411>. Data https://gfrechette.com/data/Dal_Bo_2011a_data.txt.
- Dobronyi, C., Gu, J., Kim, K., and Russell, T. M. (2026). Identification of dynamic panel logit models with fixed effects. *Review of Economic Studies*, Forthcoming. Available at <https://arxiv.org/abs/2104.04590>.
- French fuel-price reporting system (2019). Annual open-data archive. Data available at <https://donnees.roulez-eco.fr/opendata/annee/2019> Documentation <https://www.prix-carburants.gouv.fr/index.php/rubrique/opendata/>.
- Gallant, A. R., Hong, H., and Khwaja, A. (2018). The dynamic spillovers of entry: An application to the generic drug industry. *Management Science*, 64(3):1189–1211.
- Gautier, E. and Saout, R. L. (2015). The dynamics of gasoline prices: Evidence from daily french micro data. *Journal of Money, Credit and Banking*, 47(6):1063–1089. Available at <https://doi.org/10.1111/jmcb.12237>.
- Honoré, B., Hu, L., Kyriazidou, E., and Weidner, M. (2023). Simultaneity in binary outcome models with an application to employment for couples. *Empirical Economics*, 64:3197–3233. Available at <https://doi.org/10.1007/s00181-023-02417-7>.
- Honoré, B. and Kyriazidou, E. (2000). Panel data discrete choice models with lagged dependent variables. *Econometrica*, 68(4):839–874. Available at <https://doi.org/10.1111/1468-0262.00139>.

- Honoré, B. and Kyriazidou, E. (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 Weidner, M. (2025). Moment conditions for dynamic panel logit models with fixed effects. *Review of Economic Studies*, 92(5):3112–3137. Available at <https://doi.org/10.1093/restud/rdae097>.
- Igami, M. (2017). Estimating the innovator’s dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998. *Journal of Political Economy*, 125(3):798–847.
- Igami, M. and Uetake, K. (2020). Mergers, innovation, and entry-exit dynamics: consolidation of the hard disk drive industry, 1996–2016. *The Review of Economic Studies*, 87(6):2672–2020.
- Iskhakov, F., Rust, J., and Schjerning, B. (2016). Recursive lexicographical search: Finding all Markov perfect equilibria of finite state directional dynamic games. *Review of Economic Studies*, 83(2):658–703.
- Kasahara, H. and Shimotsu, K. (2009). Nonparametric identification of finite mixture models of dynamic discrete choices. *Econometrica*, 77(1):135–175. Available at <https://doi.org/10.3982/ECTA6763>.
- Kooreman, P. (1994). Estimation of econometric models of some discrete games. *Journal of Applied Econometrics*, 9(3):255–268.
- Maruyama, S. (2014). Estimation of finite sequential games. *Journal of Econometrics*, 178(2):716–726.
- Narendranathan, W., Nickell, S., and Metcalf, D. (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. Available at <https://doi.org/10.2307/2981970>.
- Pakel, C. and Weidner, M. (2026). Bounds on average effects in discrete choice panel data models. Technical report, Working paper. Available at <https://arxiv.org/abs/2309.09299>.

- Schmidt-Dengler, P. (2025). The timing of new technology adoption: The case of MRI. *RAND Journal of Economics*, 56(4):511–527.
- Soetevent, A. R. and Kooreman, P. (2007). A discrete-choice model with social interactions: With an application to high school teen behavior. *Journal of Applied Econometrics*, 22(3):599–624.
- Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies*, 70(1):147–165. Available at <https://doi.org/10.1111/1467-937X.00240>.
- US Census Bureau (2019). Basic monthly current population survey: 2019 public-use files. Available at <https://www.census.gov/data/datasets/2019/demo/cps/cps-basic-2019.html>.

A Model A.1. Proof of Corollary 1

We abbreviate $(0,0)$ by 00 , and similarly for the other joint outcomes. For model A.1, define

$$d_i(u, v) = \ln[1 + \exp\{\alpha_i + \beta_i u + \lambda_i v\}].$$

Conditional independence of the current shocks gives

$$\ln p_h = \alpha_1 N_1 + \alpha_2 N_2 + \beta_1 C_{11} + \beta_2 C_{22} + \lambda_1 C_{12} + \lambda_2 C_{21} - \sum_{i=1}^2 \sum_{u,v=0}^1 n_i^{uv}(\mathbf{h}) d_i(u, v), \quad (56)$$

where n_i^{uv} counts dates with $(y_{i,t-1}, y_{j,t-1}) = (u, v)$. The four lagged-state counts are determined by T, y_0, y_T, N_1, N_2 , and N_{11} . For example,

$$n^{11} = N_{11} + y_{10}y_{20} - y_{1T}y_{2T}, \quad n^{10} = N_1 + y_{10} - y_{1T} - n^{11},$$

with the analogous expression for n^{01} and $n^{00} = T - n^{11} - n^{10} - n^{01}$. Hence the vector $s = (y_0, y_T, N_1, N_2, N_{11})$ fixes every incidental term in (56).

Table 17: Benchmark comparisons at $T = 3$

History A	History B	$\ln(P_A/P_B)$
(00, 00, 10, 10)	(00, 10, 00, 10)	β_1
(00, 00, 01, 01)	(00, 01, 00, 01)	β_2
(00, 00, 01, 10)	(00, 01, 00, 10)	λ_1
(00, 00, 10, 01)	(00, 10, 00, 01)	λ_2

Each row in Table 17 has a common initial state, common terminal state, and identical (N_1, N_2, N_{11}) . The regressor differences are the four coordinate vectors in the order $(\beta_1, \beta_2, \lambda_1, \lambda_2)$. This proves the benchmark rank claim when initial state 00 has positive probability.

The first two rows also have the same (C_{12}, C_{21}) within each pair. Their ratios therefore continue to identify β if the cross-lag coefficients are included in the time-invariant incidental vector.

For longer panels, one may use the marginal distribution of an initial four-observation subpanel.

B Models A.2-A.4: Proof of Corollary 2

Fix a market and, for that market, write $\gamma_i \equiv \kappa_{im}$ for player i 's action-interaction. Fix the lagged state and write $e_i^z = \alpha_i + \beta_i s_i + \lambda_i s_j + \gamma_i z$. Player 2's response to action x is $r(x) = 1\{\varepsilon_2 \leq e_2^x\}$. Because player 1's deterministic action-zero payoff is zero regardless of $r(0)$, her comparison is

$$\alpha_1 + \beta_1 s_1 + \lambda_1 s_2 + \gamma_1 r(1) - \varepsilon_1 \geq 0.$$

Thus her threshold depends on the follower's response to *action one*. This is the reason the normalized level payoff matters, and, as part (ii) below shows, the reason λ_2 – which enters $e_2^{r(1)}$ – is what the leader's comparison cannot be purged of.

If $\gamma_2 < 0$, the follower always chooses one when $\varepsilon_2 \leq e_2^1$, always chooses zero when $\varepsilon_2 > e_2^0$, and responds with $r(0) = 1, r(1) = 0$ in between. Integrating over these three intervals gives

$$p_{00} = (1 - F_1^0)(1 - F_2^0), \quad p_{10} = F_1^0(1 - F_2^1), \quad p_{11} = F_1^1 F_2^1,$$

$$p_{01} = (1 - F_1^1)F_2^1 + (1 - F_1^0)(F_2^0 - F_2^1).$$

These probabilities are nonnegative and sum to one. If $\gamma_2 > 0$, the middle interval instead has $r(0) = 0, r(1) = 1$ and yields the positive-interaction expressions in (17). The formulas also have well-defined limits when an interaction is zero. Note that $F_i^z \equiv F_i(e_i^z)$ depends on *both* states, s_i and s_j , through $\beta_i s_i + \lambda_i s_j$; the two parts of the proof differ only in whether the restricted-history conditioning set controls that second dependence.

Part (i): $\lambda_1 = \lambda_2 = 0$. When both cross-state terms vanish, F_i^z depends on the rival only through z (i.e. through the current action/response), never through the rival's *lagged* state. On the product-outcome history set, the denominator for player i depends only on $(y_{i,t-1}, y_{it})$. Its four counts are fixed by $T, N_j, N_i + y_{i0} - y_{iT}$, and C_{ji} . Hence (19) cancels them. Both intercept terms and the interaction numerator $(\gamma_1 + \gamma_2)N_{11}$ are fixed. For substitutes, the two identifying comparisons can be chosen as

$$\begin{aligned} (00, 00, 10, 10) & \text{ versus } (00, 10, 00, 10) \quad \text{for } \beta_1, \\ (10, 10, 11, 11) & \text{ versus } (10, 11, 10, 11) \quad \text{for } \beta_2. \end{aligned}$$

For complements use (01, 01, 11, 11) versus (01, 11, 01, 11) for β_1 , and the second displayed comparison for β_2 . The own-lag statistic is one in the first history and zero in the second for the player whose coefficient is being identified. All post-initial outcomes lie in the required product set. This proves part (i) of Corollary 2 under the stated initial-state support condition.

Part (ii): $\lambda_1 \neq 0$ or $\lambda_2 \neq 0$. The cancellation in part (i) relied on $(y_{i,t-1}, y_{jt})$ being a sufficient pair of arguments for player i 's period- t kernel. Once $\lambda_i \neq 0$ for some i , the kernel for player i depends on the *triple* $(y_{i,t-1}, y_{j,t-1}, y_{jt})$, since e_i^z loads on $s_j = y_{j,t-1}$ directly. The counts $T, N_j, N_i + y_{i0} - y_{iT}, C_{ji}$ fix only the joint distribution of $(y_{i,t-1}, y_{jt})$ across the two compared histories; they place no restriction on how $y_{j,t-1}$ lines up with those pairs period by period. Consequently two histories with identical $(y_{i,t-1}, y_{jt})$ counts can differ in their $y_{j,t-1}$ -path, and the conditional probability of either history given the sufficient statistic is no longer invariant to that path once $\lambda_j \neq 0$.

Concretely, take the β_1 comparison (00, 00, 10, 10) versus (00, 10, 00, 10), read as the sequence of $(y_{1,t-1}, y_{2t})$ pairs. Both histories hold $y_{2t} = 0$ for $t = 1, \dots, 4$, so $s_2 = y_{2,t-1} = 0$ throughout and $\lambda_1 s_2 \equiv 0$ in both: λ_1 alone does not disturb this comparison. But the leader's threshold uses $r(1)$, whose distribution is governed by $F_2^1 = F_2(\alpha_2 + \beta_2 \cdot 0 + \lambda_2 y_{1,t-1} + \gamma_2)$ – a function of $y_{1,t-1}$ whenever $\lambda_2 \neq 0$. The two histories assign the value $y_{1,t-1} = 1$ to different periods (holding its total count fixed), so F_2^1 is no longer constant across periods common to both histories, and the period-by-period factorization that isolated β_1 in part (i) no longer goes through: the ratio of the two histories' probabilities now carries λ_2 (and its interaction with γ_1, γ_2) in a form that does not cancel against the conditioning statistic. The symmetric argument applies to the β_2 comparison (10, 10, 11, 11) versus (10, 11, 10, 11), read as $(y_{2,t-1}, y_{1t})$: there $s_2 = y_{2,t-1} = 1$ is pinned, so λ_1 's contribution is the constant λ_1 and is absorbed into α_1 without consequence, but player 2's *own* threshold loads λ_2 on $y_{1,t-1}$, which again takes the value one in different periods across the two histories and is not controlled by $T, N_1, N_2 + y_{20} - y_{2T}, C_{12}$.

In either case, once $\lambda_1 \neq 0$ or $\lambda_2 \neq 0$, the conditional probability of the restricted history given the sufficient statistic is a function of $(\alpha, \beta, \gamma, \lambda)$ jointly rather than of β_i alone, for both of the comparisons on which CL-SS relies. Since $\mathcal{H}_Q$ identifies (β_1, β_2) *only* through these two comparisons, CL-SS on $\mathcal{H}_Q$ fails to point identify (β_1, β_2) , and since no other statistic in the restricted-history construction isolates

$\alpha_i, \gamma_i, \lambda_i$ either, no parameter of the model is identified by this approach. This proves part (ii).

With positive interactions for both players, reversing the within-period move order under the same normalized payoffs gives the same transition probabilities. In the multiplicity rectangle the normalization selects 11 under either order. Thus these positive-interaction probabilities do not identify move order. This equivalence is specific to this payoff and information structure; it is not a general property of sequential games.

C Exact construction of polynomial-weight moments

For fixed β and a sign regime, express every transition probability as a rational function of (A_1, A_2, G_1, G_2) using (17). Every denominator is a product of factors $1 + A_i B_i^u G_i^v$, where $B_i = e^{\beta_i}$ and $u, v \in \{0, 1\}$. One may use their least common multiple across histories, or any common positive multiple. Multiplying the trial identity (32) by this denominator yields a polynomial. Each monomial coefficient is linear in the unknown entries of K . The kernel of this finite coefficient matrix constructs exact valid identities, including at configurations where a particular rational basis formula has a zero denominator. At such configurations the rank and basis must be recomputed.

The explicit moment (30) can be checked by inserting the four transition formulas above into the six history probabilities. Clearing the positive factors $1 + A_i B_i^u G_i^v$ makes the numerator identically zero. The replication code records this symbolic calculation with A_i, B_i, G_i left as indeterminates; it is therefore an identity for all positive A_i, B_i, G_i within the negative-interaction kernel, not a check at a finite grid.

For the numerical example, exact polynomial cancellation verifies 21 rational coefficient arrays, and exact rational population probabilities give rank three and normalized lift $(1, 2/3, 1/2, 1/3)$. The four lifted monomials are $1, G_1, G_2, G_1 G_2$, and $B = (3/2, 4/5)$. The supplied coefficient arrays and population probabilities allow direct reproduction of the rank calculation. Numerical null-space calculations are a useful way to find candidate identities. They are distinct from their exact certification by polynomial coefficient cancellation. Proposition 3 requires the latter validity and the observable rank condition; it makes no blanket claim that the condition holds at every distribution of fixed effects.

D Recursive Models and Additional CL Restrictions

Table 18: Recursive extensions: methods and scope

Model	CL-SS	FD
Myopic recursive; separate effects; (57)	Point(β), $T \geq 3$; additional zero restrictions identify γ_2 or λ_1	? Additional targets with unrestricted separate effects
Myopic recursive; common effect; initial state 00	Not used for the $T = 2$ result	Four generic moments at $T = 2$; local identification under Jacobian rank; ? general global uniqueness
Forward; autonomous player 1; $\lambda_2 = 0$; (58)	Point(β), $T \geq 3$	? Primitive interaction with heterogeneous continuation values

D.1 Myopic recursive model

Consider

$$\begin{aligned} y_{1t} &= \mathbf{1}\{\alpha_1 + \beta_1 y_{1,t-1} + \lambda_1 y_{2,t-1} - \varepsilon_{1t} \geq 0\}, \\ y_{2t} &= \mathbf{1}\{\alpha_2 + \beta_2 y_{2,t-1} + \gamma_2 y_{1t} - \varepsilon_{2t} \geq 0\}. \end{aligned} \quad (57)$$

Here $\gamma_1 = \lambda_2 = 0$. Player 1's contemporaneous action is determined independently of player 2's current action. The system is complete and its probability is a product of the two appropriate logit factors. This is a recursive simultaneous best-response system; its completeness does not require the sequential payoff normalization used for a leader with a nonzero contemporaneous interaction.

Corollary 4 (Recursive CL results). *In (57), $T \geq 3$ and sufficient initial-state support give the following conclusions.*

- (i) *Conditioning on $(y_0, y_T, N_1, N_2, N_{11}, C_{12})$ identifies β_1, β_2 .*
- (ii) *If $\beta_1 = 0$ or $\lambda_1 = 0$, a coarser conditioning set also identifies γ_2 .*
- (iii) *If $\gamma_2 = 0$ or $\beta_2 = 0$, a coarser conditioning set also identifies λ_1 .*

The additional zero restrictions are sufficient conditions for the specified CL constructions, not necessary conditions for model-level identification.

Proof. The first player's denominator depends on the lagged joint state, whose counts are fixed by $(y_0, y_T, N_1, N_2, N_{11})$. The second player's denominator depends on $(y_{2,t-1}, y_{1t})$, whose counts are fixed by $(y_0, y_T, N_1, N_2, C_{12})$. The numerator is $\alpha_1 N_1 + \alpha_2 N_2 + \beta_1 C_{11} + \beta_2 C_{22} + \lambda_1 C_{12} + \gamma_2 N_{11}$. Thus the stated conditioning set leaves only $\beta_1 C_{11} + \beta_2 C_{22}$ variable. The first two comparisons in Table 17 establish the rank condition.

If $\beta_1 = 0$ or $\lambda_1 = 0$, the first denominator depends on at most one lagged action. Conditioning on N_{11} is unnecessary. Keep $(y_0, y_T, N_1, N_2, C_{12})$ and compare $(00, 00, 11, 00)$ with $(00, 10, 01, 00)$. The differences in C_{11}, C_{22}, C_{12} are zero and the difference in N_{11} is one, giving $\ln(P_A/P_B) = \gamma_2$. If $\gamma_2 = 0$ or $\beta_2 = 0$, the second denominator depends on at most one of its two arguments. Conditioning on C_{12} is unnecessary. Keep $(y_0, y_T, N_1, N_2, N_{11})$ and use the λ_1 comparison in Table 17. Its only nonzero numerator difference is $C_{12}(A) - C_{12}(B) = 1$. Together with the relevant β comparisons, these are full-rank systems for the coefficients not fixed at zero. $\square$

D.2 Forward-looking model with an autonomous player

Impose the level payoff $u_1 = y_1(\alpha_1 + \beta_1 s_1)$, so $\lambda_1 = \gamma_1 = 0$, and maintain the stationary single-agent continuation solution for player 1. Her state and rewards then depend only on her own past action. Her effective interaction is zero. For the result here we additionally impose $\lambda_2 = 0$. The effective system is

$$y_{1t} = \mathbf{1}\{a_1 + \beta_1 y_{1,t-1} - \varepsilon_{1t} \geq 0\}, \quad y_{2t} = \mathbf{1}\{a_2 + \beta_2 y_{2,t-1} + \kappa_2 y_{1t} - \varepsilon_{2t} \geq 0\}. \quad (58)$$

Here (a_1, a_2, κ_2) can vary arbitrarily across markets in the encompassing model. Conditioning on $(y_0, y_T, N_1, N_2, N_{11}, C_{12})$ cancels the incidental terms, leaving the same index $\beta_1 C_{11} + \beta_2 C_{22}$. The first two benchmark comparisons again identify β at $T \geq 3$. The conditional likelihood has $\exp\{\beta_1 C_{11}(\mathbf{h}) + \beta_2 C_{22}(\mathbf{h})\}$ in its numerator and a sum of those expressions over histories with the same sufficient statistic in its denominator. The sufficient statistic and the regressor statistic have different roles.

Keeping a nonzero $\lambda_2 y_{1,t-1}$ would change the denominator counts and requires a different argument. We do not silently omit that term. The result above recovers both own-state coefficients without identifying the heterogeneous effective

interaction or the primitive γ_2 from this conditional likelihood.

E Probability Bounds and Their Construction

E.1 One-period rectangles and history bounds

For substitutes, outcome 10 is a Nash equilibrium whenever $\varepsilon_1 \leq e_1^0$ and $\varepsilon_2 > e_2^1$. This rectangle gives its upper probability U_{10} . The multiplicity rectangle has probability D , so $U_{10} - D$ is its sharp individual lower bound. Outcome 01 follows symmetrically. The regions for 00 and 11 are unique. This proves (??).

For a convenient lower bound on 10, either of the rectangles

$$\{\varepsilon_1 \leq e_1^0, \varepsilon_2 > e_2^0\}, \quad \{\varepsilon_1 \leq e_1^1, \varepsilon_2 > e_2^1\}$$

contains only that outcome. Their probabilities correspond to $b = (0,0)$ and $b = (1,1)$ in (36). The analogous rectangles give lower bounds on 01. For 00 and 11 use their exact product probabilities. The union of the two rectangles gives the sharper lower bound, but its probability is a sum of products. Keeping an individual rectangle preserves the log-linear decomposition needed for SS-Bounds.

Let q_t be the conditional probability of an outcome at date t given the preceding history, incidental quantities, and the maintained continuation profile. Conditional shock independence and equilibrium behavior imply $L_t \leq q_t \leq U_t$. By the chain rule,

$$\prod_{t=1}^T L_t \leq p_h(\boldsymbol{\eta}, \boldsymbol{\theta}) \leq \prod_{t=1}^T U_t.$$

This argument permits history-dependent selection: it does not multiply unconditional probabilities or assume independent selection shocks. The bounds need not be jointly attainable for several histories at once.

E.2 Signed denominator terms

Write $\sigma(u, v) = -\ln(1 + e^{a+\beta u+\kappa v})$. Its saturated expansion is

$$\sigma(u, v) = \sigma(0, 0) + du + ev + fuv.$$

The first differences satisfy $\text{sign}(d) = -\text{sign}(\beta)$ and $e \geq 0$ when $\kappa \leq 0$. The cross difference is

$$f = - \int_0^\beta \int_0^\kappa \Lambda'(a+x+z) dz dx, \quad (59)$$

so $\text{sign}(f) = \text{sign}(\beta)$ when $\kappa < 0$, with the appropriate zero limits. Summing over dates and using $\sum_t y_{i,t-1} = N_i + y_{i0} - y_{iT}$ gives (43). Exact matching of N_i eliminates the unrestricted coefficient $a_i + d_i$. The four remaining sign conditions in (44) make the incidental difference weakly negative. They apply player by player and therefore permit different known signs of β_1 and β_2 .

A transparent implementation enumerates each initial state, its 4^T histories, the unique product upper bound, and every allowed product lower bound. For myopic players it retains pairs with equal (N_i, D_i^{uv}) and records their differences in (C_{ii}, K_i) . For forward-looking players it retains equal N_i and checks the signed conditions instead. Duplicate inequalities can be removed. This procedure includes the distinct statistic-matching cases in the original bounds argument without repeating their algebra. The program returns the history pair and bound choice for every retained row, so each inequality is auditable.

E.3 Polynomial cones and an explicit FD certificate

For $\lambda = 0$ and a fixed sign regime, both the sharp one-period bounds and their history products are rational functions of $A_i = e^{a_i}$. To generate equalities between nonnegative weighted bounds, clear *one common positive denominator* and expand the lower- and upper-bound numerator vectors in a common monomial basis. The equation $G'_L \mu - G'_U \zeta = 0$ with $\mu, \zeta \geq 0$ defines the cone in (50). Its nonnegativity restrictions are essential. Arbitrary signed vectors in a null space do not preserve an inequality when probabilities replace bounds.

To check the explicit certificate in (51), insert

$$F_i^z(s) = \frac{A_i B_i^{s_i} G_i^z}{1 + A_i B_i^{s_i} G_i^z}$$

into (??), multiply the appropriate three transition bounds, and collect terms. The difference between the two displayed sides has zero numerator after clearing the positive denominators. The supplied symbolic verification keeps

$A_1, A_2, B_1, B_2, G_1, G_2$ as indeterminates. Its coefficients are nonnegative on $0 < G_i \leq 1$, which is the sign region required for converting the identity into the observable inequality.

For the heterogeneous-interaction inequality, divide the pointwise observable expression by $G_1 > 0$. The coefficients on the first and third upper-bound terms become $(1 - G_1)/G_1 \geq 0$ and $G_1 \geq 0$. Dropping these nonnegative probability terms yields (52), whose remaining coefficient B_1 is common across markets. This derivation is pointwise, so heterogeneous G_i can then be integrated out. Replacing a heterogeneous coefficient by its mean before integrating would not be valid.

E.4 Population numerical experiments

All numerical probability calculations use the exact finite-state transition kernel and enumeration of the 4^3 histories; there is no simulation noise. In the myopic simultaneous experiment, equilibrium 10 is selected with probability one half in the multiplicity rectangle, so

$$q_{10} = U_{10} - D/2, \quad q_{01} = U_{01} - D/2.$$

The population law averages over the two stated fixed-effect types. The SS region is an intersection of linear inequalities in (β, γ) , and its coordinate extrema are computed by linear programming. Intersecting it with the explicit FD inequality and its player-swapped counterpart produces the refined region. Its reported extrema use constrained nonlinear optimization from multiple starts. Their agreement and feasibility checks are numerical diagnostics, not a proof that every nonlinear optimization problem has a unique global solution.

For the forward-looking experiment, the same primitive parameters and selection probability are used with $\delta = .8$. For each type we solve the eight Bellman equations for the two players at four states. A convenient normalization replaces the action shocks by $(0, -\varepsilon_i)$, where ε_i is logistic. Relative to the independent extreme-value specification, this subtracts an action-independent shock with a constant unconditional mean. It shifts values by a constant and leaves value differences and equilibrium behavior unchanged.

Let $M(v) = v\Lambda(v) - \ln(1 + e^v) = E[\varepsilon_i \mathbf{1}\{\varepsilon_i \leq v\}]$, $d_i(s) = \alpha_i + \beta_i s_i$, and $e_i^z = d_i + \gamma_i z + \tilde{V}_i(z)$. Under the half-half selection rule, expected normalized flow payoff

for player i is

$$r_i(s) = [d_i F_i^0 - M(e_i^0)](1 - F_j^1) + [(d_i + \gamma_i) F_i^1 - M(e_i^1)] F_j^1 - \frac{1}{2} \{d_i (F_i^0 - F_i^1) - [M(e_i^0) - M(e_i^1)]\} (F_j^0 - F_j^1). \quad (60)$$

The Bellman equations are $\bar{V}_i(s) = r_i(s) + \delta \sum_y q_y(s) \bar{V}_i(y)$. The computed solution has negative effective interactions at both types and maximum residual below 3×10^{-11} . Existence of this numerical solution does not establish uniqueness of the dynamic equilibrium. The experiment conditions on it, as does its probability calculation.

The signed-SS enumeration uses the known orthant $\beta_1 \leq 0, \beta_2 \geq 0$, rather than imposing a common positive sign. For the expanded comparison family, both own-lag coefficients have finite SS intervals in Table 4. This table therefore differs from a calculation restricted to the slides' displayed lower-bound comparisons. The additional uniform FD inequality improves the second coefficient's upper bound in this experiment; it need not improve every coordinate in every population.

F Outer Regions and Sharp Identification

For a complete static or effective-coefficient model, let $\mathcal{P}_s(\theta) = \{p_s(\eta, \theta) : \eta \text{ admissible}\}$. Conditional on initial state s , an unrestricted mixing distribution generates a barycenter of this set. Exact compatibility requires an admissible probability measure F_s such that

$$P_s = \int p_s(\eta, \theta) dF_s(\eta). \quad (61)$$

The closure of the attainable mixture set is the closed convex hull of $\mathcal{P}_s(\theta)$. Supporting affine inequalities characterize that closure. Under compactness and continuity, the distinction between the attainable set and its closure disappears; with unbounded fixed effects, limiting types may lie only in the closure. This distinction should be stated whenever sharpness is claimed.

Linear FD equalities describe an affine subspace containing the mixture set. They generally do not impose its nonnegativity or moment-feasibility restrictions. The common-effect example makes the distinction concrete. A finite grid of fixed

effects, in turn, replaces the unrestricted mixture set by the convex hull of finitely many probability vectors. It is a restriction on the mixing distributions. Rejection of feasibility on that grid is not rejection under unrestricted support, and a grid calculation is not automatically a sharp identified-set computation.

For an incomplete model, let $\mathcal{P}_s(\boldsymbol{\eta}, \boldsymbol{\theta})$ contain all history-probability vectors compatible with the admissible selection rules. The exact mixture feasibility problem must integrate vectors from this set, maintaining any cross-period consistency of selection. Separate scalar history bounds generally relax this joint compatibility problem. The SS-Bounds and FD-Bounds regions are therefore outer regions. The complete cone of valid combinations of the chosen scalar bounds need not exhaust all supporting inequalities of the true history-probability correspondence.

For a structural forward-looking game, effective intercepts and interactions must additionally be generated by primitives, value functions, and equilibrium behavior satisfying the Bellman equations. Treating them as arbitrary incidental quantities relaxes these restrictions. Thus even a sharp solution of the enlarged effective-coefficient problem would not establish structural sharpness. Developing a full characterization of this additional feasibility problem remains an extension. The main identification and validity results do not rely on claiming that it has been solved.

G Empirical History Support and Implementation

Tables 20–28 give the additional individual histories, joint histories, and transition matrices. Counts refer to the samples defined in Section 5. Individual histories are descriptive measures of reversibility. A conditioning class, rather than an individual sequence in isolation, determines whether an observation informs a particular CL coefficient.

For signed SS-Bounds, the algorithm in Appendix E generates respectively 7, 12, 12, and 17 nontrivial comparisons at initial state 00 in the orthants $(-, -)$, $(-, +)$, $(+, -)$, and $(+, +)$. In gasoline all probabilities used by these comparisons have positive empirical counts, yet linear-programming feasibility fails in every orthant. For couples and the experiment some comparisons involve zero counts. In implementation, $n_A > 0, n_B = 0$ violates $n_A \leq e^{d'\beta} n_B$ at every finite parameter; $n_A = 0, n_B > 0$ imposes no empirical restriction; and two zero counts give a

Table 19: Components of the explicit selection-robust FD inequalities

Sample	Player	n_D	n_1	n_2	n_3	n_4	n_R	Plug-in upper
Couples	1	0	2	1	4	5	2	No finite endpoint
	2	0	4	0	3	8	-1	Sample violation
Gasoline	1	71	2,284	479	98	2,126	735	2.337
	2	97	2,231	452	110	2,270	523	1.685
Experiment	1	0	9	3	1	15	-2	Sample violation
	2	0	8	0	1	11	-2	Sample violation

For player 1, $n_D = n(00,10,00,11)$ and (n_1, n_2, n_3, n_4) count $(00,00,10,00)$, $(00,00,10,01)$, $(00,00,10,11)$, and $(00,10,00,00)$. For player 2 interchange players in every history. Here $n_R = n_1 + n_2 + n_3 - n_4$. The upper endpoint is $\ln(n_R/n_D)$ when both arguments are positive. These are empirical moment diagnostics, not confidence intervals. A negative numerator can reflect sampling error or failure of the maintained model. All histories include their common initial state.

tautology. We preserve these distinctions instead of taking undefined logarithms or adding pseudo-counts. An empty plug-in region is not a confidence statement and does not justify reporting sampling uncertainty as zero.

The empirical source files and scripts record the same player labels, window selection, and clustering conventions used in the tables. Baseline conditional-likelihood results are inherited from the supplied empirical draft and checked against its saved numerical outputs. The sequential estimates use the full transition kernel, with actual-Hessian two-step covariance calculations for the stable interior solutions. The new bounds calculations are recomputed directly from the saved four-observation panels. The data construction and economic restrictions are described in the application text; the estimation histories in the replication package permit independent reproduction of all reported history counts.

Table 20: Couples employment: distribution of individual histories

History	Player 1: count	Percent	Player 2: count	Percent
0000	4,076	25.23	1,782	11.03
0001	153	0.95	100	0.62
0010	59	0.37	14	0.09
0011	199	1.23	110	0.68
0100	76	0.47	31	0.19
0101	28	0.17	25	0.15
0110	50	0.31	31	0.19
0111	474	2.93	420	2.60
1000	213	1.32	103	0.64
1001	128	0.79	54	0.33
1010	42	0.26	34	0.21
1011	365	2.26	352	2.18
1100	223	1.38	115	0.71
1101	298	1.84	309	1.91
1110	485	3.00	403	2.49
1111	9,286	57.48	12,272	75.96

Each history includes the initial observation. Percentages use the full number of four-observation histories for the relevant application. The two marginal counts do not determine the frequencies of the joint histories required by the conditional comparisons.

Table 21: Couples employment: eight most frequent joint histories

Player 1 history	Player 2 history	Count	Percent
1111	1111	7,717	47.77
0000	1111	2,837	17.56
1111	0000	795	4.92
0000	0000	752	4.65
0111	1111	297	1.84
1110	1111	281	1.74
1011	1111	190	1.18
1111	0111	169	1.05

Histories are ordered by their observed joint frequency. Percentages are relative to all retained histories; the listed rows need not exhaust the sample.

Table 22: Couples employment: joint-state transition frequencies

Previous state	00	01	10	11	Transitions
00	72.09	11.73	5.32	10.86	3,665
01	3.46	85.54	0.44	10.56	11,880
10	5.67	1.12	77.04	16.16	3,916
11	1.58	4.43	2.02	91.97	29,004

Columns are current states. Entries are row percentages pooling the three within-window transitions. These are unconditional descriptive frequencies among observed histories, not structural transition probabilities evaluated at common fixed effects.

Table 23: Gasoline price adjustment: distribution of individual histories

History	Player 1: count	Percent	Player 2: count	Percent
0000	31,414	51.07	31,112	50.58
0001	4,770	7.75	4,730	7.69
0010	5,391	8.76	5,411	8.80
0011	1,014	1.65	1,073	1.74
0100	5,423	8.82	5,516	8.97
0101	1,174	1.91	1,268	2.06
0110	1,268	2.06	1,373	2.23
0111	446	0.73	497	0.81
1000	4,934	8.02	4,665	7.58
1001	1,475	2.40	1,488	2.42
1010	1,207	1.96	1,294	2.10
1011	437	0.71	506	0.82
1100	1,121	1.82	1,135	1.85
1101	500	0.81	503	0.82
1110	625	1.02	610	0.99
1111	313	0.51	331	0.54

Each history includes the initial observation. Percentages use the full number of four-observation histories for the relevant application. The two marginal counts do not determine the frequencies of the joint histories required by the conditional comparisons.

Table 24: Gasoline price adjustment: eight most frequent joint histories

Player 1 history	Player 2 history	Count	Percent
0000	0000	19,198	31.21
0010	0000	2,284	3.71
0000	0100	2,270	3.69
0000	0010	2,231	3.63
1000	0000	2,227	3.62
0000	0001	2,156	3.51
0100	0000	2,126	3.46
0001	0000	2,107	3.43

Histories are ordered by their observed joint frequency. Percentages are relative to all retained histories; the listed rows need not exhaust the sample.

Table 25: Gasoline price adjustment: joint-state transition frequencies

Previous state	00	01	10	11	Transitions
00	74.00	10.63	10.17	5.20	130,740
01	64.63	16.02	14.03	5.32	21,613
10	65.02	14.38	14.77	5.83	20,936
11	60.29	12.08	11.39	16.24	11,247

Columns are current states. Entries are row percentages pooling the three within-window transitions. These are unconditional descriptive frequencies among observed histories, not structural transition probabilities evaluated at common fixed effects.

Table 26: Experimental cooperation: distribution of individual histories

History	Player 1: count	Percent	Player 2: count	Percent
0000	581	39.55	605	41.18
0001	24	1.63	31	2.11
0010	26	1.77	24	1.63
0011	13	0.88	20	1.36
0100	37	2.52	32	2.18
0101	19	1.29	23	1.57
0110	22	1.50	22	1.50
0111	41	2.79	43	2.93
1000	126	8.58	105	7.15
1001	18	1.23	11	0.75
1010	27	1.84	15	1.02
1011	28	1.91	16	1.09
1100	51	3.47	57	3.88
1101	9	0.61	13	0.88
1110	24	1.63	28	1.91
1111	423	28.80	424	28.86

Each history includes the initial observation. Percentages use the full number of four-observation histories for the relevant application. The two marginal counts do not determine the frequencies of the joint histories required by the conditional comparisons.

Table 27: Experimental cooperation: eight most frequent joint histories

Player 1 history	Player 2 history	Count	Percent
0000	0000	412	28.05
1111	1111	389	26.48
1000	0000	96	6.54
0000	1000	76	5.17
0000	1100	32	2.18
1100	0000	30	2.04
0111	1111	21	1.43
1011	0111	21	1.43

Histories are ordered by their observed joint frequency. Percentages are relative to all retained histories; the listed rows need not exhaust the sample.

Table 28: Experimental cooperation: joint-state transition frequencies

Previous state	00	01	10	11	Transitions
00	90.23	4.33	4.99	0.45	1,985
01	40.12	23.87	21.40	14.61	486
10	48.94	24.86	16.76	9.44	519
11	0.35	1.83	4.02	93.79	1,417

Columns are current states. Entries are row percentages pooling the three within-window transitions. These are unconditional descriptive frequencies among observed histories, not structural transition probabilities evaluated at common fixed effects.