STRUCTURAL ESTIMATION OF DIRECTIONAL DYNAMIC GAMES WITH MULTIPLE EQUILIBRIA*

FEDOR ISKHAKOV †
AUSTRALIAN NATIONAL UNIVERSITY

DENNIS KRISTENSEN ‡
UNIVERSITY COLLEGE LONDON

JOHN RUST §
GEORGETOWN UNIVERSITY

BERTEL SCHJERNING ¶
UNIVERSITY OF COPENHAGEN

July, 2025

PRELIMINARY AND INCOMPLETE

Abstract

We develop a robust algorithm for computing the nested full solution maximum likelihood estimator for a class of directional dynamic stochastic games with multiple equilibria. We show how the computational burden of the full solution approach can be substantially reduced in large datasets, making it computationally feasible. The proposed estimator is remarkably robust to multiplicity of equilibria in the theoretical model, and reliably delivers efficient maximum likelihood estimates of the structural parameters while identifying the equilibria played in the data. Using the dynamic model of Bertrand competition with cost-reducing investments, we run a series of Monte Carlo experiments to explore the performance of our estimator in comparison to the battery of existing estimators for dynamic games. We find that our estimator outperforms all of the existing estimators in terms of accuracy and reliability.

KEYWORDS: Structural estimation, dynamic discrete games, multiple equilibria, directional dynamic games, Markov perfect equilibrium, recursive lexicographic search algorithm, MPEC, NFXP, nested recursive lexicographic search algorithm (NRLS)

*The authors would like to acknowledge the funding received from the Danish Council for Independent Research and Australian Research Council Future Fellowship FT180100632.

†Research School of Economics, the Australian National University, fedor.iskhakov@anu.edu.au

‡University College London, Department of Economics, d.kristensen@ucl.ac.uk

§ Corresponding Author: Georgetown University, Department of Economics, jr1393@georgetown.edu

¶University of Copenhagen, Department of Economics and Centre for Computational Economics (CCE), bertel.schjerning@econ.ku.dk

Contents

1 Introduction **1**

2 Directional dynamic games **4**
2.1 Discrete dynamic games of incomplete information . . . 4
2.2 Directionality in dynamic games . . . 6
2.3 Example: the leapfrogging game . . . 7

3 Full solution maximum likelihood estimator **11**
3.1 Nested RLS estimator (NRLS) . . . 11
3.2 Recursive lexicographic search algorithm (RLS) . . . 16
3.3 Inner loop of the NRLS estimator: branch and bound approach . . . 17
3.4 Statistical properties of the NRLS estimator . . . 17
General framework . . . 17
Same equilibrium across all markets . . . 18
Inference on equilbrium . . . 21
Multiple equilibria played . . . 22
Treating model with absorbing state as an approximation . . . 23
Partial identification inference . . . 24
Treat equilibrium selection as mixture model . . . 24

4 Monte Carlo experiments **25**
4.1 Existing estimators . . . 25
2-step PML and NPL estimator . . . 25
EPL estimator . . . 25
MPEC estimators . . . 25
4.2 Experiment A: Unique equilibrium . . . 28
4.3 Experiment B: Small multiplicity with stable equilibrium in the data . . . 29
4.4 Experiment C: Small multiplicity with *unstable* equilibrium in the data . . . 30
4.5 Experiment D: Discontinuous likelihood function . . . 32
4.6 Experiment E: Massive multiplicity . . . 33
4.7 Experiment F: Convergence properties of NRLS . . . 34
4.8 Experiment G: Multiple equilibria in the data . . . 34

5 Discussion and conclusions **34**

1 Introduction

Dynamic games are numerically difficult to solve and estimate. The main reason is the existence of multiple equilibria, particularly in discrete games, which presents methodological challenges for the estimation of these models. Existing applied work typically rely on solution algorithms and estimation methods that do not fully explore, or even assumes away, this multiplicity, thereby inadvertently operating as an equilibrium selection mechanism. However, rather than haphazardly selecting an equilibrium depending on for example how the solution algorithm is initialized, a robust estimator *should compute all equilibria and select one, or several of them on a priori economic grounds or based on what we can learn from the data.*

In this paper, we develop a robust algorithm for computing the "full solution" maximum-likelihood estimator of a particular class of dynamic stochastic games with multiple equilibria, namely directional dynamic games. Our method allows for multiple equilibria having been played in data without making any assumptions on the equilbirium selection rule. It is based on Rust's 1987 Nested Fixed Point (NFXP) maximum likelihood estimator, but uses the Recursive Lexicographic Search (RLS) algorithm IRS2016 within the "the inner loop" of NFXP to solve for *all* Markov perfect equilibria (MPE) of the game at each evaluation of the likelihood function. Until recently, the NFXP approach have generally not been possible to implement for dynamic games with multiple equilibria since no algorithm was guaranteed to find all MPE. RLS provides this capability for directional dynamic games.

In models with numerous equilibria, the full solution method that combines NFXP and RLS (NRLS), even though theoretically ideal, may be impractical because of the computational burden of finding all equilibria. Utilizing the special structure of equilibria in directional dynamic games, we develop a specialized combinatorial optimization algorithm that can be fine tuned to balance the computational cost and the statistical efficiency of the estimator.

Dynamic directional games are a subclass of stochastic games with finite state space where under all feasible Markovian strategies the game transitions through the points of the state space in a directed fashion from a subset of initial states to a subset of absorbing states. The *state recursion algorithm* developed in IRS2016 solves these games by backward induction on the state space, sequentially computing the solutions to the system of Bellman equations for particular points, and using the solutions already computed for the points downstream. In this manner, the large system of equations characterizing the equilibria of the whole game is decomposed into a number of smaller computational tasks (fixed point problems for the infinite horizon games or non-linear systems of equations otherwise).

In cases when the game has multiple equilibria, the state recursion algorithm itself is invoked by RLS sequentially and recursively for particular subsets of the state points. The RLS algorithm can be thought of as a tree traversal algorithm that simultaneously discovers the structure of the tree and computes the required statistics for each branch. The tree is composed of all MPE of the game: each branching corresponds to a point in the state space where the system of Bellman equations of the players for that point has multiple solutions. Each MPE in the whole game can be mapped into a path from the origin of the tree to each individual leaf.

Ideally, the discrete optimization problem of finding the equilibrium with the maximum likelihood for a given candidate value of the structural parameters has to be solved exactly. To solve this problem we employ the *branch and bound* optimization algorithm Land and Doig (1960) which is an ideal computational tool for the task due to the tree structure of the equilibria and the properties of the likelihood function. To employ the *branch and bound* algorithm we need to specify a monotone bounding function representing the best attainable objective on a given subset of equilibria (branch). In our setting, computing the bounding function amounts to recursive calculation of the partial likelihood over paths starting at the root of the RLS tree and ending at the leafs. This bounding function monotonically declines as data for more and more states is added and eventually equals to the full log-likelihood at the leafs of RLS tree.

We show that the use of this algorithm in its original form significantly reduces the computational burden compared to the MLE with full enumeration of equilibria, since it allows us to dismiss the branches of the RLS tree where the partial likelihood is below the current best attained value of the likelihood. The computational efficiency of the branch and bound algorithm in computing the maximum likelihood over the discrete set of equilibria, depends on how informative the data is in every points of the state space; especially in points close to the trunk of the tree that correspond to the later stages of the game.

To enhance the overall computational performance of the estimator in small samples, we therefore develop a refinement of the branch and bound algorithm for our problem which further sharpens the bounds using a statistical criterion that allows us to use the data to reject branches that appear particularly unlikely. The refined version of the algorithm computes an approximation to the true maximum of equilibrium specific likelihoods, and therefore trades off statistical accuracy for computational feasibility. However, as the sample size increases, the exact solution by the original branch and bound algorithm proves to be computationally feasible, resulting in fully efficient MLE estimator.

Both the basic and the refined versions of the branch and bound algorithm is particularly useful in an estimation context where equilibrium selection is determined in conjunction with estimation of a set of structural parameters. When estimating the structural parameters, we can exploit that the branch and bound algorithm is nested in the loop that searches over the a set of structural parameters. The nested stricture allows us to further sharpen the bounding function for each evaluation of the likelihood function by recursively updating the bounding function (i.e. the partial likelihood) during the course of estimation. Together, this significantly reduces the computational burden by only examining branches of the solution tree that will potentially result in an increase of the likelihood function.

There has been considerable progress in the development of algorithms for computing Markov perfect equilibria, starting with the pioneering work by Pakes and McGuire (1994) and recent progress on homotopy methods for finding multiple equilibria of both static and dynamic contexts Borkovsky, Doraszelski and Kryukov (2010); Besanko, Doraszelski, Kryukov and Satterthwaite (2010). Still, it is an extremely challenging problem to find even a single MPE of a dynamic game, much less all of them. The computational complexity has continuously led researchers to propose estimation methods that do not require to repeatedly compute the solution of the dynamic game during the course of estimation. Various two-step estimators have been proposed in the literature[1] (see references

[1] see for example Bajari, Benkard and Levin (2007), Pesendorfer and Schmidt-Dengler (2010), and Pakes, Ostrovsky and Berry (2007)

in Egesdal, Lai and Su (2015)), but these are known to suffer from potential large biases in finite samples. In an effort to reduce the finite-sample biases associated with the two stage pseudo maximum likelihood (2S-PML) estimator, AM2007 propose an algorithm to "swap the nesting" in the NFXP algorithm, and suggested *nested pseudo likelihood* (NPL) algorithm that should be robust to multiple equilibria as long as only *one* is played in the data. However, Pesendorfer and Schmidt-Dengler (2010) show that NPL can fail to produce consistent estimates such that we cannot expect to uncover structural parameters using NPL in even very large samples. Kasahara and Shimotsu (2012) propose a two-step modified version of NPL (NPL-Λ) that perform better. However, Egesdal, Lai and Su (2015) demonstrated that the lack of convergence of NPL and $NPL-\Lambda$ is not a trivial issue in practice.

To circumvent both the small-sample bias and the repeated solution of all equilibria, Su and Judd (2012) advocated for a constrained optimization approach to structural estimation which they called *mathematical programming with equilibrium constraints*. Egesdal, Lai and Su (2015) implement MPEC for the entry/exit dynamic game of AM2007 and argue MPEC overcomes two computational challenges: it does not need to solve for all Markov perfect equilibria and even with multiple equilibria the constrained optimization problem is smooth. They conduct Monte Carlo experiments to investigate the numerical performance and finite-sample properties of their constrained optimization approach and argue that the MPEC approach is a favorable method for estimating dynamic games.

We show that NRLS is computationally efficient and asymptotically equivalent to MLE and present Monte Carlo evidence to compare it's performance to a variety of existing estimators including MPEC. Our Monte Carlo evidence is based on simulations from equilibria in the dynamic Bertrand investment game in **?**. The models is a simple dynamic discrete choice extension of the classic static model of Bertrand price competition where competing duopolists are allowed to undertake cost-reducing investments in an attempt to "leapfrog" their rival to attain temporary low-cost leadership. A characteristic aspect of the findings in this paper is a possibility of plethora of equilibria: even a simple finite state, dynamic extension of the standard static textbook model of Bertrand price competition may result in *hundreds of millions of Markov Perfect Equilibria* with a big variety of different investment dynamics.

RLS gives us the ability to fully solve the Bertrand investment game and simulate data from any equilibrium in this model, and therefore provide a perfect testbed for the performance of estimators in case of multiple equilibria. We study the behavior of existing estimation methods in situations that have generally not been possible previously. Specifically, we simulate data from the Bertrand investment game where we vary the number of equilibria by choosing the appropriate model parameters. By mixing data from several of these equilibria we investigate the performance of existing methods when multiple equilibria are played in the data. We also analyze the consequences of having a vast multiplicity of equilibria in the theoretical model when only a single one is played in the data.

We find NRLS to be remarkably robust, computational fast and able to both obtain efficient MLE of the structural parameters and at the same time identify the equilibrium selection played in the data (out of millions of potential MPEs). Moreover, NRLS allows us to relax assumptions for the equilibrium selection rules to allow for different equilibria to be played at different markets without much additional computational cost.

2 Directional dynamic games

Our basic framework is a discrete version of stochastic game due to Shapley (1953) where some of the state variables constitute private information of the players as in AM2007. The Markov perfect equilibria of these games of incomplete information are characterized by a collection of choice probabilities defined over the common state variables, which allows construction of the likelihood function of the observed sample of actions and states. We introduce additional assumptions to give new characterization of the subclass of directional dynamic games (DDG) considered in (IRS2016). The dynamic Bertrand pricing and investment from (IRS2018) is a running example that we return to throughout the paper.

2.1 Discrete dynamic games of incomplete information

We follow AM2007 to set up the canonical dynamic discrete game of incomplete information. Consider a stochastic game of $N < \infty$ players who in each time period choose an action $a_i \in A$ from a finite set of actions $A = \{0, \ldots, J\}$. Vector $a = (a_1, \ldots, a_N)$ is the action profile of all players in the current time period. The state of the game is given by a finite set of variables x_i which are known to all players, and a set of private ε_i known only to player i. The vectors $x = (x_1, \ldots, x_N)$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)$ are the common knowledge and private information, respectively, where $x \in X$ has a discrete and finite support X, $|X| < \infty$. We assume that (x, ε) follow a controlled Markov process with transition probability $p(x', \varepsilon' | a, x, \varepsilon)$, where (x', ε') refer to the next period values. The instantaneous payoff of player i is given by $\tilde{\Pi}_i(a, x, \varepsilon_i)$. We further impose the following standard assumptions.

Assumptions. *(AS) Additive separability: the per-period payoff of player i is additively separable in x and ε_i, i.e. $\tilde{\Pi}_i(a, x, \varepsilon_i) = \Pi_i(a, x) + \eta \varepsilon_i[a_i]$, where $\varepsilon_i[a_i]$ is the $(a_i + 1)$-th component of vector $\varepsilon_i \in \mathbb{R}^{J+1}$ and η is scaling parameter; (CI) Conditional independence: The state (x, ε) evolves as a first-order controlled Markov process with transition density $p(x', \varepsilon' | a, x, \varepsilon) = f(x' | a, x) \cdot p_\varepsilon(\varepsilon')$, where f governs the transition of the common knowledge state and p_ε is the time-invariant distribution of private shocks; (IPV) Independent private values: Private shocks are independently distributed across players, i.e. $p_\varepsilon(\varepsilon) = \prod_{i=1}^{N} g_i(\varepsilon_i)$ where g_i is the probability density function of ε_i; (EV) Extreme value[2]: Each private shock ε_i is distributed according to the extreme value type I (Gumbel) distribution, i.e. $g_i(\varepsilon_i) = \exp(-\varepsilon_i - \exp(-\varepsilon_i))$.*

Let $\sigma = (\sigma_1, \ldots, \sigma_N)$ denote a Markovian strategy profile, where $\sigma_i : X \times \mathbb{R}^{J+1} \to A$ is a decision rule mapping of each player. The behavior of each player under strategy profile σ is perceived by the other the players through the choice probabilities denoted $P_i^\sigma(a_i | x) \equiv Pr\{\sigma_i(x, \varepsilon_i) = a_i | x\}$. Denote $a_{-i} \in A^{N-1}$ the vector of all possible actions of the players other than player i. Under a strategy profile σ the deterministic part of i's instantaneous payoff as well as the transition probability for the common knowledge states seen from i's point of view can be expressed as a function of their action

[2]We introduce the EV assumption right away to simplify exposition; all theoretical results and representations follow through with any g_i that are absolutely continuous with respect to Lebesgue measure (Hotz and Miller (1993), AM2007).

a_i as

$$\begin{aligned} \pi_i^\sigma(a_i,x) &= \sum_{a_{-i}\in A^{N-1}} \left(\prod_{j\neq i} P_j^\sigma\left(a_{-i}[j]|x\right)\right) \Pi_i(a,x), \\ f_i^\sigma(x'|a_i,x) &= \sum_{a_{-i}\in A^{N-1}} \left(\prod_{j\neq i} P_j^\sigma\left(a_{-i}[j]|x\right)\right) f(x'|a,x), \end{aligned} \tag{1}$$

where $a_{-i}[j]$ is the j-th element of a_{-i} and the full vector a is formed from a_i and a_{-i}. The dynamic behavior of player i is then described by the *integrated* Bellman equation

$$V_i^\sigma(x) = \int \max_{a_i\in A} \{v_i^\sigma(a_i,x) + \eta\varepsilon_i[a_i]\}\, g_i(d\varepsilon_i) = \eta \log \sum_{a_i\in A} \exp\left(\frac{v_i^\sigma(a_i,x)}{\eta}\right) + \eta\gamma, \tag{2}$$

where the last equality is due to (EV), $\gamma \approx 0.52277$ is the Euler-Mascheroni constant, and $v_i^\sigma(a_i,x)$ denotes the a_i-choice specific value function of player i. With the common discount factor β it is given by

$$v_i^\sigma(a_i,x) = \pi_i^\sigma(a_i,x) + \beta \sum_{x'\in X} V_i^\sigma(x') f_i^\sigma(x'|a_i,x). \tag{3}$$

Conditional on a fixed strategy profile σ, equations (2) and (3) define a Bellman operator which is a contraction mapping in the space of value functions $V_i^\sigma(x)$ (Aguirregabiria and Mira, 2002; Ma and Stachurski, 2021).

Definition 1. *A strategy profile $\sigma^\star$ constitutes a stationary Markov perfect equilibrium (MPE) of the game if for any firm i and for any state $(x,\varepsilon_i) \in X \times \mathbb{R}^{J+1}$ it holds*

$$\sigma_i^\star(x,\varepsilon_i) = \arg\max_{a_i\in A} \left\{ v_i^{\sigma^\star}(a_i,x) + \eta\varepsilon_i[a_i] \right\}, \tag{4}$$

that is the strategy profile $\sigma^\star$ prescribes mutual best responses that solve Bellman equation (2) for $i \in \{1,\ldots,N\}$.

The MPE equilibrium can equivalently be characterized in probability space. First, note that all payoff relevant information, namely π_i^σ, f_i^σ, V_i^σ and v_i^σ depend on players' strategies only through the choice probabilities $P_i^\sigma(a_i|x)$ associated with σ. Given (EV), the choice probabilities take logit form

$$P_i^\sigma(a_i|x) = Pr\{\sigma_i(x,\varepsilon_i) = a_i|x\} = \frac{\exp(v_i^\sigma(a_i,x)/\eta)}{\sum_{a_i'\in A}\exp(v_i^\sigma(a_i',x)/\eta)}, \; i \in \{1,\ldots,N\}. \tag{5}$$

Denote $\Lambda\colon P \mapsto P$ a mapping from the space of choice probabilities defined for all (a,x) for all players $i \in \{1,\ldots,N\}$, to itself, according to equation (5) where $v_i^\sigma(a_i,x)$ are computed using arguments P in place of $P_i^\sigma(a_i|x)$. The MPE equilibrium constitutes a fixed point of Λ, i.e. $P^\star = \Lambda(P^\star)$, existence of which is guaranteed by the Brouwer theorem. However, unlike the Bellman operator above, the *best response probability operator* Λ is not a contraction and may have multiple fixed points. Effectively, the equilibrium choice probabilities $P^\star$ solve N Bellman equations simultaneously.

For completeness, we present an alternative representation of the best response mapping that facilitates the definition of the directional dynamic games in the next section. Applying the law of

iterated expectations to the equation (2) we can rewrite it as

$$V_i^\sigma(x) = \sum_{a_i \in A} P_i^\sigma(a_i|x) \left[\pi_i^\sigma(a_i,x) + e_i^\sigma(a_i,x) + \beta \sum_{x' \in X} V_i^\sigma(x') f_i^\sigma(x'|a_i,x) \right], \tag{6}$$

where the term $e_i^\sigma(a_i,x) \equiv E(\varepsilon_i[a_i]|x, \sigma(x,\varepsilon_i) = a_i)$ is the expectation of the private shock ε_i conditional on action a_i being assigned by strategy σ_i in state x. Hotz and Miller (1993) show that this term is a function only of the choice probabilities $P_i^\star(a_i|x)$ and the private shock distribution g_i. Under (EV), it takes the form $e_i^\sigma(a_i,x) = \eta\gamma - \eta\log(P_i^\sigma(a_i|x))$.

Recognizing that for every $i \in \{1,\ldots,N\}$ the equation (6) represents a linear system of equations in the unknowns $V_i^\sigma(x)$, we can rewrite and solve it in matrix form. Let V_i^σ stack all values of $V_i^\sigma(x)$ over the state space to form a $|X| \times 1$ column vector, and define $\pi_i^\sigma(a_i)$, $e_i^\sigma(a_i)$ and $P_i^\sigma(a_i)$ similarly. The solution of the linear system (6) is given by

$$V_i^\sigma = (I - \beta F)^{-1} \sum_{a_i \in A} P_i^\sigma(a_i) * [\pi_i^\sigma(a_i) + e_i^\sigma(a_i)], \tag{7}$$

where $*$ denotes the element-wise Hadamard product, I is the identity matrix and F is defined by the following.

Definition 2 (Transition probablity matrices)**.** *The conditional transition matrix $F(a)$ is a $|X| \times |X|$ matrix with a typical element $f(x'|a,x)$ equal to the transition probability from common knowledge state x to state x' given the action profile a. The unconditional transition matrix F is a $|X| \times |X|$ matrix with a typical element $F_{x,x'} = \sum_{a \in A^N} \left(\prod_{i=1}^N P_i^\sigma\left(a_i|x\right)\right) f(x'|a,x)$ equal to the transition probability from common knowledge state x to state x' averaged over all possible action profiles a with the corresponding joint choice probabilities.*[3]

The right hand side of equation (7) defines a *policy valuation* operator $\Gamma\colon P \mapsto V^\sigma$, which maps the space of choice probabilities into the space of value functions implementing the Hotz-Miller inversion. Define the $\Psi\colon P \mapsto P$ operator by the composition $P \mapsto V^\sigma \mapsto v^\sigma \mapsto P$ which maps choice probabilities for all (a,x) for all players $i \in \{1,\ldots,N\}$ to the corresponding value functions using Γ operator, then choice specific value functions using equation (3), and finally back to the choice probabilities using equation (5). (AM2007, Representation Lemma) establishes that the MPE equilibria of the game can be equivalently characterized as fixed points of both Λ and Ψ operator, i.e. $P^\star = \Psi(P^\star) = \Lambda(P^\star)$, with computational implementation of the Ψ operator being less costly.

2.2 Directionality in dynamic games

IRS2016 introduced directional dynamic games as a subclass of discrete stochastic games that possess an additional structure in the state space. Using notation of the previous section, they consider the case when the state space X can be decomposed as $X = D \times X'$, and used a directed graph $\mathscr{G}(\sigma)$ to capture the structure of positive transition probabilities between the points of D conditional on the strategy profile σ. By definition, for a game to be *directional*, it has to satisfy two conditions: (1)

[3]The unconditional transition matrix F can also be computed as a weighted sum of the conditional transition matrices $F(a)$ with the similar weights, see the proof of Corollary 2.2.

apart from self-loops, each graph $\mathscr{G}(\sigma)$ has to be a acycling, and (2) no two graphs $\mathscr{G}(\sigma)$ and $\mathscr{G}(\sigma')$ contain edges in opposite directions between any two vertices for all feasible strategies σ and σ'.

In this section we present an alternative definition for directional dynamic games that is based on the properties of the conditional and unconditional transition probabilities of the common knowledge state X as defined in the previous section. We use the representation of the graph $\mathscr{G}(\sigma)$ by an adjacency matrix, and draw on the direct one-to-one correspondence between these two objects to express the definitions and propositions from (IRS2016) in terms of the transition probability matrices introduced in the previous section for discrete dynamic games of incomplete information.

Proposition 2.1 (Directional dynamic games of incomplete information)**.** *A discrete dynamic game of incomplete information given by the primitives* $\left(N, A, X, f(x'|a,x), \eta, \beta, \Pi_i(a,x), g_i(\varepsilon_i), i \in \{1,\ldots,N\}\right)$ *is a* directional dynamic game *(IRS2016) if and only if the unconditional transition probability matrix F given by Definition 2 is upper block-triangular, or can be converted to an upper block-triangular form by a permutation applied to its rows and columns, and all the blocks are of the same size.*

Proof. The proof follows from the fact that the upper block-triangularity of the unconditional transition probability is equivalent to the non-cyclicality of the directed graph induced by the common coarse refinement of the strategy-specific partial orders on the directional component of the state space, as defined in (IRS2016).

□

Corollary 2.2. *For a directional dynamic game of incomplete information to be a DDG it is sufficient for all conditional transition probability matrices $F(a)$ to be upper block-triangular, or be convertible to an upper block-triangular form by a common permutation applied to their rows and columns, and all the blocks to be of the same common size.*

Proof. Following the Definition 2, the unconditional transition probability matrix F is a weighted sum of the conditional transition probability matrices $F(a)$, where the weights are given by a $|X| \times |X|$ matrices with identical columns $\prod_{i=1}^{N} P_i^{\sigma}(a_i)$, and the product is element-wise. The proof follows from the fact that the any element that is zero in $F(a)$ for all a is also a zero in F. □

2.3 Example: the leapfrogging game

The empirical literature on dynamic games has multiple examples of DDGs, including (Dube, Hitsch and Chintagunta, 2010; Judd, Schmedders and Yeltekin, 2012; Anderson, Rosen, Rust and Wong, 2025). Here we lay out the model of Bertrand competition with cost-reducing investments, which is a dynamic directional game where a typical equilibrium involves the two asymmetric duopolists to leapfrog each other in terms of the marginal cost of production, hence the model's shorthand name. Depending on the specification and parameter values, this relatively simple directional dynamic game may have very different number of equilibria, ranging from a single equilibrium (IRS2018, Theorem 4, pp. 21) up to hundreds of millions of equilibria in some computed examples. With such flexibility, the leapfrogging model present a very useful testbed for running experiments aimed at exploring the effects of multiplicity of equilibria on the performance of a battery of standard estimators applicable for finite state stochastic games. We use this model as a convenient data generating process in the Monte Carlo experiments of Section 4.

In the context of the generic discrete dynamic games of incomplete information, consider the game between $N = 2$ Bertrand competitors who interact over time by *simultaneously* deciding whether or not to invest in the latest (state of the art) production technology c at a cost $K(c)$, that is $A = \{I, N\}$ corresponding to the two options. If either of the firms invests, their current technology characterized by the marginal cost of production c_i is replaced by the state-of-the-art technology c after one period lag. The common knowledge state of the game is given by $x = (c_1, c_2, c)$ where the state-of-the-art cost c evolves exogenously according to a Markov transition process $\pi(c'|c)$ where $\pi(c'|c) = 0$ for all $c' > c$ reflecting the non-reversibility of technological progress. Assuming myopic demand and normalizing the market size to one, the instantaneous payoffs follow the outcome of the static Bertrand competition with the indirect profit function given by

$$\Pi_i(a,x) = (c_j - c_i)\mathbb{1}\{c_i < c_j\} - K(c)\mathbb{1}\{a_i = I\} = \pi_i(a_i,x),\ i,j \in \{1,2\}, i \neq j. \tag{8}$$

In words, the cost leader receives a profit equal to the difference in the duopolists' marginal costs, the cost follower receives zero profit, and the tie-breaking rule for the case $c_1 = c_2$ does not effect the payoffs. The last equality is due the time-to-build assumption ensuring that current actions are not affecting the current payoffs, see equation (1).

For generality we also consider the case of differentiated products where consumers choose the lower price subject to an idiosyncratic EV1 shock with scale λ, such that the probability of choosing product i is given by the logit formula $q_i(p_1^\lambda, p_2^\lambda, \lambda) = \exp(-p_i^\lambda/\lambda)/\big(\exp(-p_1^\lambda/\lambda) + \exp(-p_2^\lambda/\lambda)\big)$. In this case, the prices $(p_1^\lambda, p_2^\lambda)$ charged by the two firms constitute Bayesian-Nash equilibrium in a static pricing game, and are given as a solution to the following system of equations

$$\begin{array}{rcl} p_1^\lambda & = & \arg\max_p q_1(p, p_2^\lambda, \lambda)(p - c_1), \\ p_2^\lambda & = & \arg\max_p q_2(p_1^\lambda, p, \lambda)(p - c_2), \end{array} \tag{9}$$

in which case the indirect profit functions are given by $\Pi_i^\lambda(a,x) = q_i(p_1^\lambda, p_2^\lambda, \lambda)(p_i^\lambda - c_i) - K(c)\mathbb{1}\{a_i = I\}$. Even though technically the strategies in this case also include pricing decisions, we treat the profit functions implied by the static Bayesian-Nash equilibrium as indirect payoffs and continue focusing on the investment decisions $a_i \in A = \{I, N\}$ as the only strategic actions. It is straightforward to show that as $\lambda \to 0$, the differentiated product case converges to pure Bertrand competition with $p_1^\lambda \to \max(c_1, c_2)$, $p_2^\lambda \to \max(c_1, c_2)$, and $\Pi_i^\lambda(a,x) \to \Pi_i(a,x)$ given in equation (8).

The state space X of the leapfogging game is discretized with K points in each of three dimensions, and is given by $X = \{(c_1, c_2, c) \in \mathbb{R}^3 : c_1 \geqslant c, c_2 \geqslant c\}$, incorporating the natural technological constraints. The cardinality of the state space is consequently equal $|X| = K(K+1)(2K+1)/6$. Whereas c evolves as an exogenous Markov process, transitions of (c_1, c_2) are governed directly by the investment decisions of the two firms: for both i the decision $a_i = N$ causes $c_i' = c_i$, and the decision $a_i = I$ causes $c_i' = c \leqslant c_i$. It is already obvious that none of the described transitions can lead to any of the state variables to increase, hence the directionality of the state transisions. To avoid tedious specification of $f(x'|a,x)$, we present the choice specific value functions v_i^σ in equation (3) that embody all

the necessary information on state transitions:

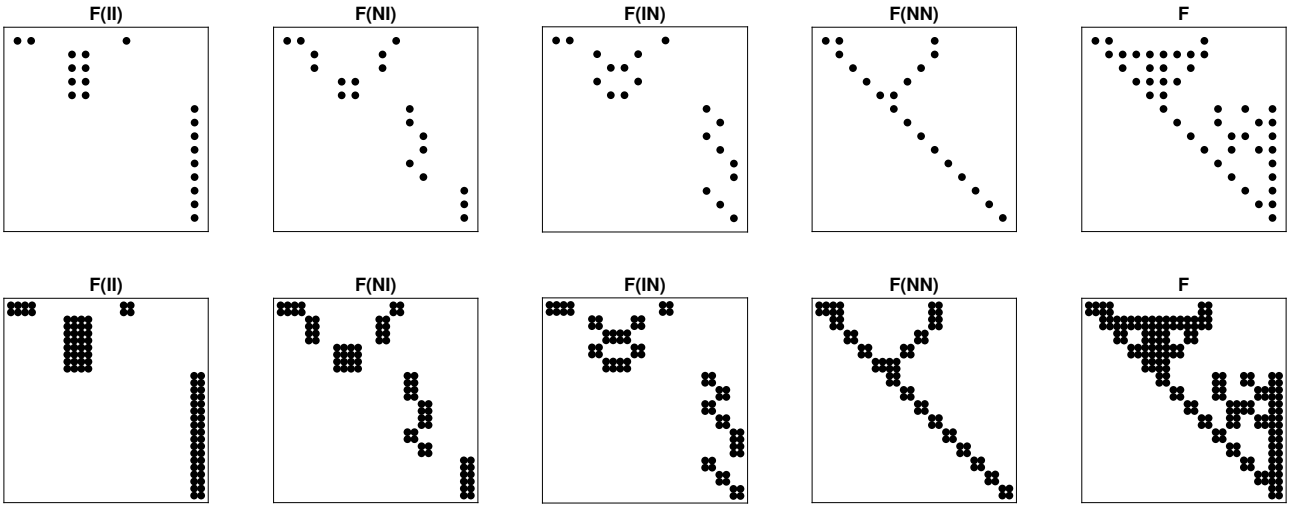
$$\begin{aligned} v_i^\sigma(N,x) &= \pi_i(N,x) + \beta \sum_{c'} \left[P_j(I,x) V_i^\sigma(c_1,c,c') + P_j(N,x) V_i^\sigma(c_1,c_2,c') \right] \pi(c'|c), \\ v_i^\sigma(I,x) &= \pi_i(I,x) + \beta \sum_{c'} \left[P_j(I,x) V_i^\sigma(c,c,c') + P_j(N,x) V_i^\sigma(c,c_2,c') \right] \pi(c'|c). \end{aligned} \tag{10}$$

The top row of Figure 1 shows the sparsity structure of the conditional and unconditional transitions probabilities $F(a)$ and F per Definition 2 for the leapfrogging game with simultaneous moves, $K = 3$. When stacking transition probabilities over the points in the state space X, we used the following ordering rules: (c_1, c_2, c) precedes (c_1', c_2', c') if (a) $c > c'$; or (b) $c = c'$, $c_1 > c$ and $c_2 > c$ while $c_1' = c'$ or $c_2' = c'$; and using arbitrary ordering otherwise. The resulting four conditional transition probability matrices $F(a)$ for $a \in \{(I,I),(N,I),(I,N),(N,N)\}$ and the unconditional F clearly have upper triangular form (block-triangular with 1×1 blocks).

In addition, we consider the *alternating move* version of the leapfrogging game where at each time period only one of the two players can make the investment decision. It requires an additional state variable $m \in \{1,2\}$ that indicates which player is making the decision at the current time period. We assume that m is exogenously determined by the Markov process with transition probabilities $f(m'|m)$ which includes deterministic alternation of moves but may also be stochastic. In the terms of IRS2016, the directional component of the state is (c_1, c_2, c) idetical to the simultaneous move game, and together with the non-directional component of the state space $m \in \{1,2\}$, the common knowledge state space is given by a cartesian product $X \times \{1,2\}$. If the state points are indexed such that the m-dimension is the last, the transition probability matrices have the 2×2 blocks corresponding to each element in the simultaneous move game, as shown in the bottom row of Figure 1. Yet, the alternating move leapfrogging game is still directional by Proposition 2.1.

More details on the setup of the leapfrogging game can be found in IRS2016, with the theoretical properties of the model investigated in IRS2018. Online Appendix NNN contains the full specification of both simultaneous and alternating move versions of the leapfrogging game, in the terms of Section 2.1.

Figure 1: Conditional and unconditional transition probability matrices in the leapfrogging game.



Notes: The table presents the non-zero elements of the conditional and unconditional transition probability matrices $F(a)$ and F in the simultaneous move version of the leapfrogging game, and the bottom row presents the same matrices in the alternating move version of the leapfrogging game. The action space is given by $\{I,N\}$ for each of the two players, resulting in four possible combinations of actions. Number of points in the state space is 14 and 28 for the two versions of the game, respectively, where in the alternating move game the additional state variable $m \in \{1,2\}$ is not directional, resulting in 2 blocks in the transition probability matrices.

3 Full solution maximum likelihood estimator

In this section we define the full solution maximum likelihood estimator for the direcitonal dynamic games and establish its statistical properties. The estimator is based on repeated solutions of the dynamic game using the RLS algorithm, therefore in line with nested fixed point estimator of Rust (1987) we term it *nested RLS* estimator (NRLS).

3.1 Nested RLS estimator (NRLS)

Using notation of AM2007, the Markov perfect equilibria (MPE) in a dynamic stochastic game can be written as

$$\begin{cases} \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta), \\ \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta), \end{cases} \tag{11}$$

where the first line represents the set of Bellman equations that characterize the optimal intertemporal choice by all the players, and the second line represents the set of equations describing the players beliefs about the actions of other players. More specifically, operator $\Psi^{\mathbf{V}}$ indexed by the structural parameters θ maps the space of value functions V and choice probabilities P of the players into the space of value functions. Operator $\Psi^{\mathbf{P}}$ also indexed by θ maps the same space into the space of choice probabilities of the players. Each solution of the system (11) gives a Markov perfect equilibrium of the game.

Denote the set of all solution to the system (11) and thus all MPE by

$$SOL(\Psi, \theta) = \left\{ (\mathbf{P}, \mathbf{V}) \;\middle|\; \begin{array}{l} \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta) \\ \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta) \end{array} \right\}. \tag{12}$$

Given the data $\mathbf{Z} = \{\mathbf{a}^{mt}, \mathbf{x}^{mt}\}_{m \in \mathscr{M}, t \in \mathscr{T}}$, where $\mathscr{M} = \{1, ..., M\}$ and $\mathscr{T} = \{1, ..., T\}$ is the set observed of markets and time periods, the log-likelihood function is given by

$$\mathscr{L}(\mathbf{Z}, \theta) = \max_{\mathbf{P} \in SOL(\Psi, \theta)} \sum_{m=1}^{M} \sum_{t=1}^{T} \sum_{i=1}^{N} \left\{ \log \mathbf{P}_i(a_i^{mt} | \mathbf{x}^{mt} m, \theta) + \log f(\mathbf{x}^{mt} | \mathbf{x}^{mt-1}, a_i^{mt-1}, \theta) \right\} \tag{13}$$

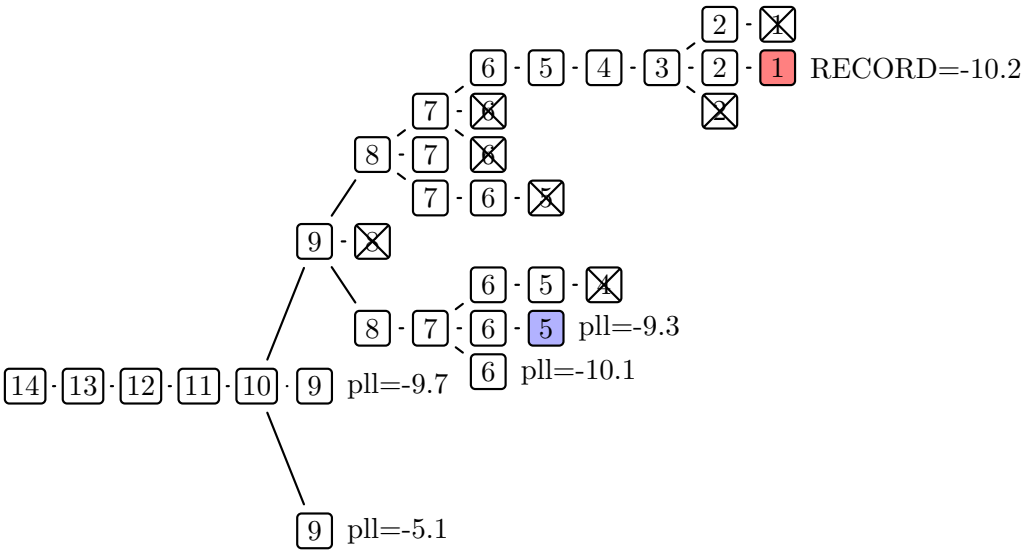
$$= \max_{\mathbf{P} \in SOL(\Psi, \theta)} \mathscr{L}^{(1)}(\mathbf{Z}, \theta, \mathbf{P}) + \mathscr{L}^{(2)}(\mathbf{Z}, \theta), \tag{14}$$

where $\mathbf{P} = (\mathbf{P}_1, ..., \mathbf{P}_N)$ and $\mathbf{a}^{mt} = \left(a_1^{mt}, ..., a_N^{mt}\right)$ contain the players' choice probabilities and actions, respectively, and $f(\cdot|\cdot)$ denotes the transition density of the controlled state process. The maximum likelihood estimator of structural parameters θ is given by

$$\hat{\theta}^{ML} = \arg\max_{\theta \in \Theta} \mathscr{L}(\mathbf{Z}, \theta). \tag{15}$$

The proposed ML estimator (15) has the nested structure identical to the NFXP estimator of Rust (1987): on the outer loop the likelihood function is maximized with respect to the structural parameters, while each evaluation of the likelihood requires the solution of the dynamic problem, forming the inner loop. The major difference is that unlike in the single agent problems where the Bellman

Figure 2: Example of partially traversed RLS tree using branch and bound algorithm.



Notes: The figure shows an intermediate stage of the branch and bound algorithm with a partially traversed RLS tree where two branches were fully extended, and the current best likelihood value of -10.2 is recorded in the second one. Hypothetical values of partial log-likelihood (pll) are shown next to the branches which are yet to be traversed, including the current location of the traversal algorithm in the tree marked with blue shading.

operator is typically a contraction mapping and therefore admits single solution, the inner loop of the NRLS estimator typically has multiple solutions collected in the set $SOL(\Psi,\theta)$ in (15). It is therefore that the estimator has *an additional maximum* over the set of model equilibria that correspond to the particular value of parameter θ, as seen in (13)-(14). The need to compute all solutions to (11) and find the most likely equilibrium conditional on the data, has hindered the development of full solution nested estimators in dynamic games until recently when RLS algorithm solved this problem in the class of directional dynamic games.

The computational burden of the proposed estimator is large and hard to predict upfront, as the number of equilibria may differ for different values of θ. In order to make NRLS estimator computationally feasible, the computation of all MPE for the given value of θ can be combined with the maximization in the first component of (14) by the means of the branch and bound method (Land and Doig, 1960).

Let $\mathbf{Z}_k = \left\{\mathbf{a}_k^{mt}, \mathbf{x}_k^{mt}\right\}_{m\in\mathscr{M},t\in\mathscr{T}}$ denote the subset of data observed at the points of the state space *up to and including the point indexed* $k = \{1,\dots,K\}$.We assume that the indexing of the state points follows the directionality of the game, and thus the order of the nodes of the RLS tree. The point $k = 1$ is the origin of the tree (base of the root), and point $k = K$ corresponds to the level of the leafs of the tree. Obviously, the full dataset corresponds to $k = K$, i.e. $\mathbf{Z}_K = \mathbf{Z}$.

The likelihood of the subset of the data is given by the likelihood function (13) computed on $\mathbf{Z}_k$. Because $\mathscr{L}^{(2)}(\mathbf{Z},\theta)$ is invariant to the particular equilibrium being played, we focus on the first component of (14). The key property that grants the use of branch and bound algorithm for solving the discrete maximization problem $\max_{\mathbf{P}\in SOL(\Psi,\theta)} \mathscr{L}^{(1)}(\mathbf{Z},\theta,\mathbf{P})$, is the monotonicity of the *partial*

likelihood function $\mathscr{L}^{(1)}(\mathbf{Z}_k, \theta, \mathbf{P})$, that is

$$\mathscr{L}^{(1)}(\mathbf{Z}_k, \theta, \mathbf{P}) \geq \mathscr{L}^{(1)}(\mathbf{Z}_{k'}, \theta, \mathbf{P}) \text{ for } \forall k < k' \in \{1, \ldots, K\}. \tag{16}$$

Indeed, as more data is added, partial likelihood can only decrease. This implies that if for some k and particular equilibrium $\mathbf{P}$ computed in the points of the state space up to k (according to the RLS algorithm), the partial likelihood falls below the full likelihood of some other equilibrium $\mathbf{P}'$ given by $\mathscr{L}^{(1)}(\mathbf{Z}, \theta, \mathbf{P}')$, it is not worth spending computing $\mathbf{P}$ any further. In terms of traversal of RLS tree, if the partial likelihood $\mathscr{L}^{(1)}(\mathbf{Z}_k, \theta, \mathbf{P})$, of some branch $\mathbf{P}$ traversed up to level k falls below the same measure of a fully traversed branch $\mathbf{P}'$, it is not worth traversing it any further.

Overall, applying the branch and bound algorithm to the RLS tree, and using the partial likelihood function as the bounding criterion, it is possible to find the most likely equilibrium in the set $SOL(\Psi, \theta)$ without exhausting all solutions with RLS method.

However, the power of BnB algorithm can be significantly improved by the use of a more informative bounding rule we describe next.

By definition, the partial likelihood collects the likelihood of the data over an already expanded part of any branch of the RLS tree.

Let θ denote structural parameters, and ω index the equilibria of the game, and $P_j^{\times}(k, \theta, \omega)$ stand for the parametric choice probabilities to invest, $P_j^I(k, \theta, \omega)$, and not to invest, $P_j^N(k, \theta, \omega)$, for firm j in the state point k.

$$L(\theta, \omega) = \sum_{i=0}^{K} \sum_{j=1}^{2} L_{ij}(\theta, \omega) = \sum_{i=0}^{K} \sum_{j=1}^{2} \left[n_j^I(i) \log P_j^I(i, \theta, \omega) + n_j^N(i) \log P_j^I(i, \theta, \omega) \right],$$

Consider any node in the RLS tree, let the corresponding point in the state space have index k. Given an ordering of state points required by the RLS algorithm, let $k = 0$ denote the initial point (one of the points in the very first stage, the apex in the leapfrogging game). Let K be the total number of points in the state space, and $k = K$ correspond to the terminal point (one of the points in the absorbing stage). The root of the RLS tree is in terminal state point $k = K$, while every leaf of the RLS tree corresponds to the initial state point $k = 0$.

Partial likelihood $L(k, \theta, \omega)$ at node k equals the likelihood of all data in points $i \geq k$.

$$L(k, \theta, \omega) = \sum_{i=k}^{K} \sum_{j=1}^{2} L_{ij}(\theta, \omega).$$

If there are several branches originating at node k, i.e. point k admits multiple stage equilibria, we associate multiple values of partial likelihood to this point. Partial likelihood computed at any node k is shared by all MPE equilibria stemming from this node, which thus share the path from the root of the RLS tree up to node $k-1$.

Partial likelihood at the leaves of the RLS tree equals the likelihood of the equilibrium corresponding to that leaf (and given by the path in the RLS tree from the root to the leaf).

In Figure 3 partial likelihood is plotted over the state space on the horizontal axes, where the points are ordered by their index k in reverse order from terminal point $k = K$ to the initial point $k = 0$. Partial likelihood is shown on the vertical axes. Each black line corresponds to a single MPE in the game.

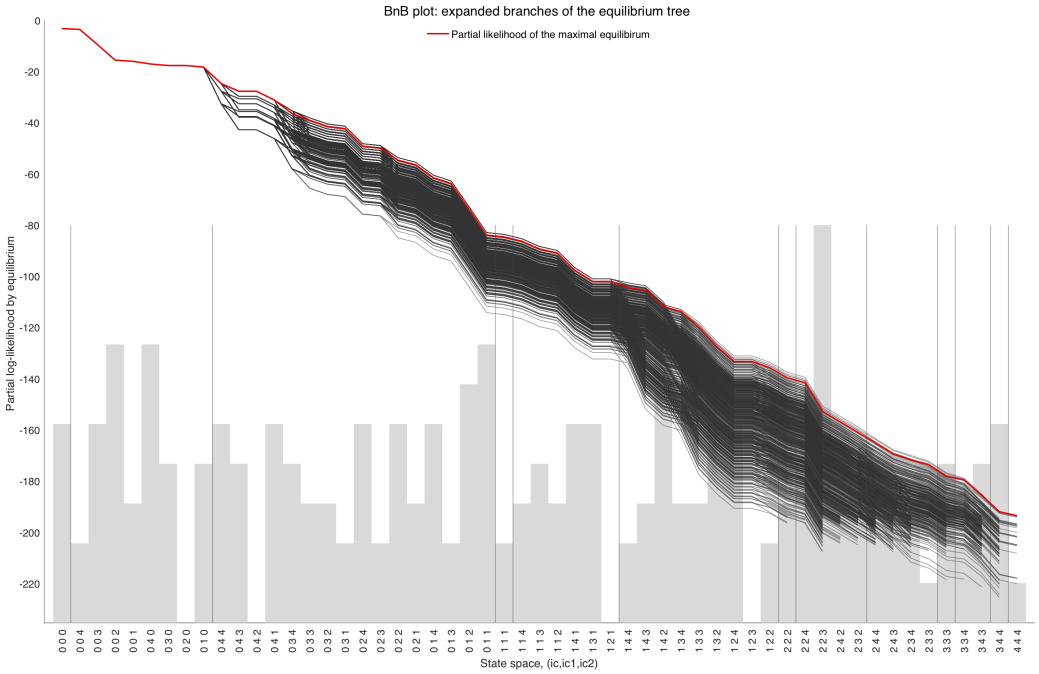


Figure 3: Partial log-likelihood for different MPE

For each line in the partial likelihood plot, the right-most value is the likelihood of the MPE represented by that line. Highlighted in red is the maximal likelihood equilibrium line, which corresponds to the maximum value of the loglikelihood for the whole sample (highest value at the leaf of the RLS tree on the right side of the plot). Some lines do not reach the last $k = 0$ state because they are not expanded by the branch-and-bound (BnB) algorithm after crossing the current at that moment likelihood record.

The histogram in the background of Figure 3 shows the distribution of the (synthetic) observed data points across the state space.

Before solving the model for any equilibria, we can compute the *partial nonparametric log-likelihood* in the same way that partial likelihood is computed over subsets of data (solid blue line in the plot).

We use the following formula below for computing nonparametric loglikelihood L^e. Denote $n_j^I(k)$ as the number of observed investment choices and $n_j^N(k)$ as the number of observed non-investment choices by firm j in the point of the space k.

$$L^e = \sum_{i=0}^{K} \sum_{j=1}^{2} L_{ij}^e = \sum_{i=0}^{K} \sum_{j=1}^{2} \left[n_j^I(i) \log \frac{n_j^I(i)}{n_j^I(i) + n_j^N(i)} + n_j^N(i) \log \frac{n_j^N(i)}{n_j^I(i) + n_j^N(i)} \right],$$

where we set $L_{ij}^e = 0$ whenever $n_j^I(i) = 0$ or $n_j^N(i) = 0$ or $n_j^I(i) + n_j^N(i) = 0$. In other words, we set the nonparametric likelihood to zero for the points of the state space which do not have data or

choice variation. Note that nonparametric likelihood is independent of both the structural parameters and equilibrium selection.

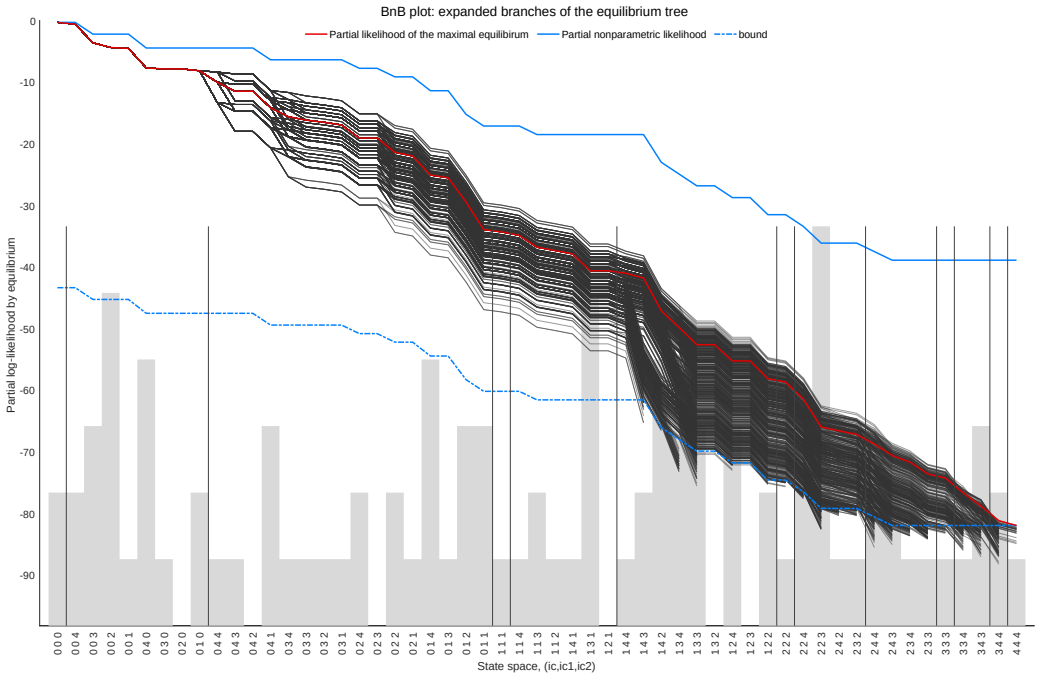


Figure 4: Parial nonparametric log-likelihood (solid blue line) and optimistic computation boundary (dashed blue line)

We refer to the nonparametric log-likelihood $L^e(k)$ computed for the data not yet accounted in the partial likelihood at any node k as *remaining nonparametric likelihood*.

$$RL^e(k) = \sum_{i=0}^{k-1} \sum_{j=1}^{2} L^e_{ij}$$

It is straightforward to show that the nonparametric likelihood is an upper boundary for the likelihood of any parametric model, and thus the blue line in Figure 4 is never declining steeper than the any of the black lines.

Therefore, remaining nonparametric likelihood can be used to forecast the remaining likelihood contribution of each equilibrium, and thus compare the likelihood of any given partially expanded branch and the corresponding MPE as the computation progresses. This is effectively a refinement of the bounding rule in the BnB algorithm.

Let the nonparametric log-likelihood be zero for the points of the state space that do not have data or do not have variation in the observed choices. Because the nonparametric likelihood is an upper bound on the likelihood of the parameteric model, there is no loss of generality of using it as the bounding rule in BnB. That is, to add the remaining nonparametric likelihood at each node of the RLS tree to forecast the value of complete likelihood for that equilibrium and compare it to the current at that moment record.

Equivalently, each current record can be "back-casted" to form the bounding rule for each node of the tree. This is shown in Figure 4 by the dashed blue line in the above plot. Drawn from the actual maximized value of the likelihood, it displays the best theoretical bounding rule of the BnB

solver. In reality, the bound is updated together with the best achieved likelihood. Greedy algorithm of choosing the branch with maximum partial likelihood at each branching node is the way to approach the theoretically best bounding rule.

The standard BnB rule requires expanding each branch until the partial likelihood is strictly below the current record. This corresponds to a completely flat nonparametric likelihood as if no data is available, or data shows no variation in choices.

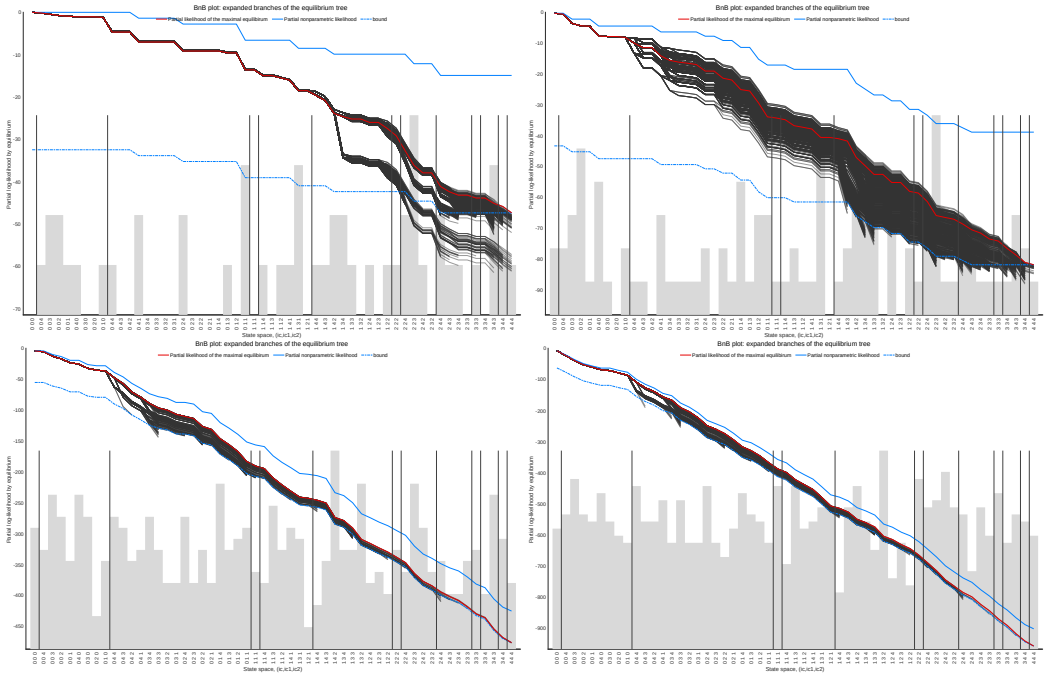


Figure 5: Effect of the sample size on the amount of computations

As sample size increases, the nonparametric loglikelihood line and the pseudo-likelihood line converge to the likelihood line, see Figure 5.

This implies *much sharper* BnB bounding rule, resulting in much fewer branches to be expanded! This is indicated by the width of the band between the blue lines in the plots below.

In some sense, as more data allows for a more precise two-step estimator, for NRLS more data implies fewer computations. Large sample size opens the possibility to estimate much larger models.

With a larger sample size, most of the branches do not have to be expanded at all!

Thus, the branch and bound method with the bounding rule based on the nonparametric NRLS

3.2 Recursive lexicographic search algorithm (RLS)

The recursive lexicographical search (RLS) algorithm of IRS2016 is designed to find all Markov perfect equilibria of a directional dynamic game by decomposing the problem of finding all equilibria into a series of subproblems, each of which corresponds to a stage of the original dynamic game. Using a specially designed *successor function* the decomposition algorithm traverses the lexicographically ordered list of equilibrium selection rules (ESR) to search for the next MPE of the game.

In this section we show that the RLS algoritm of IRS2016 is nothing more than a depth-first traversal of a specific connected an acyclic graph i.e. an RLS tree.

The RLS tree nodes correspond to the points in the (finite) state space of the game. The root of the tree is one of the points in the absorbing final stage of the game (within each stage the states/nodes can have arbitrary order, as per RLS algorithm). Each branching of the tree happens in the states which admit multiple stage equilibria. Maximum number of branches in any node equals the maximum number of stage equilibria in the game.

A single MPE is a complete path in the RLS tree connecting the root with any leaf. Any leaf of the tree corresponds to a point in the very first stage of the game (apex of the state pyramid in the leapfrogging game). The RLS tree is balanced in the sense that all leafs are placed in the final (top-most) level. Every path from trunk to any leaf goes through every level, and each level corresponds to a point in the state space.

3.3 Inner loop of the NRLS estimator: branch and bound approach

3.4 Statistical properties of the NRLS estimator

General framework

We here consider a more general class of games – not necessarily directional ones – since the inference tools developed below apply more generally.

Consider M games (markets), each with two players indexed by $i = 1,2$, observed over T time periods. In the mth game at time $t \geq 1$, let $x_{m,t}$ be the set of state variables and $a_{m,t}^{(1)} \in \{0,1\}$ and $a_{m,t}^{(2)} \in \{0,1\}$ be the actions of the two players. We treat the dynamics of $x_{m,t}$ as known (pre–estimated).

We take as given a parametric model for the game of interest which produces conditional choice probabilies (CCP's) for $a_{m,t}^{(1)}$ and $a_{m,t}^{(2)}$ at a given equilbrium of the model $e(\theta)$ which are known up to some parameter $\theta \in \Theta$. Let $P^{(i)}(x_{m,t}; e(\theta), \theta)$, $\theta_1 \in \Theta_1(e)$ denote the CCP of player i choosing $a_{m,t}^{(i)} = 1$ given the market is in equilibrium $e(\theta)$, $i = 1,2$. Importantly, the set of possible equilibria are parameter–dependent. In order to have a full characterization and unique labelling of these equilbria across different parameter values, let $\mathscr{E} = \{e_1, ..., e_J\}$ be the set of all possible J equilibria across all values of $\theta \in \Theta$. Furthermore, let $\mathscr{E}(\theta) \subseteq \mathscr{E}$ denote the subset that exists for a particular value of θ. In particular, $\mathscr{E} = \bigcup_{\theta \in \Theta} \mathscr{E}(\theta)$. Finally, let $\Theta(e) \subseteq \Theta$ be the parameter subset for which equilbrium $e \in \mathscr{E}$ exists.

Given a random sample of markets, each of which is observed over T time periods, $a_{m,t}, x_{m,t}$, $1 \leq m \leq M$ and $1 \leq t \leq T$, we can then express the log-likelihood conditional on θ being the data-generating structural parameter value and $\mathbf{e}_M = (e_1, ..., e_M) \in \mathscr{E}^M$ being the set of equilibria being played across the M markets as

$$\mathscr{L}_{M,T}(\theta, \mathbf{e}_M) = \frac{1}{MT} \sum_{m=1}^{M} \sum_{t=1}^{T} \ell(z_{m,t}; \theta, e_m),$$

with $z_{m,t} = (a_{m,t}, x_{m,t})$, and

$$\ell(z_{m,t}; \theta, e) = \sum_{i=1}^{2} a_{m,t}^{(i)} \log P^{(i)}(x_{m,t}, e; \theta) + \left(1 - a_{m,t}^{(i)}\right) \log\left(1 - P^{(i)}(x_{m,t}, e; \theta)\right).$$

We can here think of the particular equilibria played across the different markets, $\mathbf{e}_M$, as unknown market specific fixed effects. Note that $\mathscr{L}_{M,T}(\theta,\mathbf{e}_M)$ is not well-defined for $\theta \notin \bigcap_{m=1}^{M}\Theta(e_m)$. We therefore define the parameter set as

$$\mathscr{B}_M = \left\{ (\theta,\mathbf{e}_M) \in \Theta \times \mathscr{E}^M : \theta \in \bigcap_{m=1}^{M} \Theta(e_m) \right\}$$

and the corresponding maximum–likelihood estimator as

$$(\hat{\theta},\mathbf{e}_M) = \arg \max_{(\theta,\mathbf{e}_M)\in\mathscr{B}_M} \mathscr{L}_{M,T}(\theta,\mathbf{e}_M).$$

Finally, note that if the state dynamics contain unknown parameters, say, θ_x, to be estimated then

$$\mathscr{L}^{(x)}_{M,T}(\theta_x) = \frac{1}{MT}\sum_{m=1}^{M}\sum_{t=1}^{T} \log f(x_{m,t}|x_{m,t-1},a_{m,t-1};\theta_x)$$

needs to be added to $\mathscr{L}_{M,T}(\theta,\mathbf{e}_M)$. We will in the following focus on estimation of θ and so ignore $\mathscr{L}^{(x)}_{M,T}(\theta_x)$, thereby effectively treating the state dynamics as known to us.

Same equilibrium across all markets

Consider first the scenario where the same equilibirium is played in all markets. Under this scenario the data requirements are quite weak. In particular, we can let T stay fixed and only let $M \to \infty$. As a consequence, we do not need to make any assumptions about the time series properties of $z_{m,t}$; in particular, it is allowed to be non–stationary and so the theory covers directional games.

With a single equilibrium played across all markets, we have

$$e_{0,m} = e_0 \in \mathscr{E}, m = 1,...,M.$$

Suppose furthermore that the econometrician knows this and so computes

$$(\hat{\theta},\mathbf{e}_M) = \arg \max_{(\theta,e)\in\mathscr{B}} \mathscr{L}_{M,T}(\theta,e),$$

where

$$\mathscr{L}_{M,T}(\theta,e) = \mathscr{L}_{M,T}(\theta,e,....,e), \quad \mathscr{B} = \{(\theta,e) \in \Theta \times \mathscr{E} : \theta \in \Theta(e)\}.$$

With

$$\mathscr{L}_T(\theta,e) = E\left[\mathscr{L}_{M,T}(\theta,e)\right] = \frac{1}{T}\sum_{t=1}^{T} E\left[\ell(z_{m,t};\theta,e)\right]$$

denoting the asymptotic limit of $\mathscr{L}_{M,T}(\theta,e)$ as $M \to \infty$, we impose the following assumption on the model:

Assumption 1. (i) Θ is compact and $\mathscr{E}$ is finite with $\theta_0 \in \Theta(e_0)$ and $e_0 \in \mathscr{E}$; (ii) $\mathscr{L}_T(\theta,e) < \mathscr{L}_T(\theta_0,e_0)$ for any $\theta \in \Theta(e)$ and $e \in \mathscr{E}$ with $(\theta,e) \neq (\theta_0,e_0)$; (iii) $\theta \mapsto \ell(z_{m,t};\theta,e)$ is almost surely continuous on $\Theta(e)$ for all $e \in \mathscr{E}$; (iv) $\|\ell(z_{m,t};\theta,e)\| \leq b(z_{m,t})$ for all $\theta \in \Theta(e)$ and $e \in \mathscr{E}$, where $\sum_{t=1}^{T} E[b(z_{m,t})] < \infty$.

Assumption 1 is fairly standard for showing consistency of M–estimators, c.f. Section 2 of Newey and McFadden (1993): Parts (i)-(ii) ensures that the data–generating parameter value θ_0 and equilbrium e_0 are well-separated from the other potential ones of the model and so are identified. Parts (i) and (iii)-(iv) imply that $\mathscr{L}_{M,T}(\theta,e)$ converges uniformly in (θ,e) towards $\mathscr{L}_T(\theta,e)$ in probability. Under Assumption 1, consistency holds:

Theorem 3.1. *Under Assumption 1,* $\hat{\theta} \to^P \theta_0$ *and* $\Pr(\hat{e} = e_0) \to 1$*, as* $M \to \infty$.

Proof. First note that Assumption 1(i), (iii) and (iv) imply $\sup_{\theta\in\Theta(e)} |\mathscr{L}_{M,T}(\theta,e) - \mathscr{L}_T(\theta,e)| \to^P 0$ for all $e \in \mathscr{E}$, c.f. Lemma 2.1 in Newey and McFadden (1993). This in turn implies that

$$\sup_{e\in\mathscr{E}} \sup_{\theta\in\Theta(e)} |\mathscr{L}_{M,T}(\theta,e) - \mathscr{L}_T(\theta,e)| \to^P 0 \tag{17}$$

since $\mathscr{E}$ is a finite set. Second, Assumption 1(i)-(iii) imply that, for any given $\varepsilon > 0$, there exists a $\delta > 0$ such that $\|\theta - \theta_0\| > \varepsilon$ and $e \neq e_0$ implies $\mathscr{L}_T(\theta,e) \leq \mathscr{L}_T(\theta_0,e_0) + \delta$. Note this in particular holds true if $\theta \notin \Theta(e)$ since in this case $\mathscr{L}_T(\theta,e) = -\infty$.

We wish to show that for any $\varepsilon > 0$, $P\left(\{||\hat{\theta} - \theta_0|| > \varepsilon\} \bigcup \{\hat{e} \neq e_0\}\right) \to 0$ as $M \to \infty$. As shown above, for a given $\varepsilon > 0$, there exists a $\delta > 0$ such that $\|\theta - \theta_0\| > \varepsilon$ or $e \neq e_0$ implies $\mathscr{L}_T(\theta,e) \leq \mathscr{L}_T(\theta_0,e_0) + \delta$. This in turn implies $|\mathscr{L}_T(\theta,e) - \mathscr{L}_T(\theta_0,e_0)| \geq \delta$. Thus,

$$\Pr(||\hat{\theta} - \theta_0|| > \varepsilon, \hat{e} \neq e_0) \leq \Pr(\mathscr{L}_T(\hat{\theta},\hat{e}) \leq \mathscr{L}_T(\theta_0,e_0) + \delta) \leq \Pr(|\mathscr{L}_T(\hat{\theta},\hat{e}) - \mathscr{L}_T(\theta_0,e_0)| \geq \delta).$$

We then wish to show that the final probability in above display converges to zero which is equivalent to $\mathscr{L}_T(\hat{\theta},\hat{e}) \to^P \mathscr{L}_T(\theta_0,e_0)$. Since (θ_0,e_0) is the unique maximiser of $\mathscr{L}_T(\theta,e)$, we know that $\mathscr{L}_T(\theta_0,e_0) \geq \mathscr{L}_T(\hat{\theta},\hat{e})$. Thus,

$$|\mathscr{L}_T(\hat{\theta},\hat{e}) - \mathscr{L}_T(\theta_0,e_0)| = \mathscr{L}_T(\theta_0,e_0) - \mathscr{L}_T(\hat{\theta},\hat{e}) = \left\{\mathscr{L}_T(\theta_0,e_0) - \hat{\mathscr{L}}_T(\theta_0,e_0)\right\} + \left\{\hat{\mathscr{L}}_T(\theta_0,e_0) - \mathscr{L}_T(\hat{\theta},\hat{e})\right\}$$

where, by (17), $\mathscr{L}_T(\theta_0,e_0) - \hat{\mathscr{L}}_T(\theta_0,e_0) = o_P(1)$. For the second righ–hand side term of above display, first observe that, by the definition of $(\hat{\theta},\hat{e})$, $\hat{\theta} \in \Theta(\hat{e})$ for some $\hat{e} \in \mathscr{E}$. Thus, again using (17),

$$\hat{\mathscr{L}}_T(\theta_0,e_0) - \mathscr{L}_T(\hat{\theta},\hat{e}) \leq \hat{\mathscr{L}}_T(\hat{\theta},\hat{e}) - \mathscr{L}_T(\hat{\theta},\hat{e}) \leq \max_{e\in\mathscr{E}} \sup_{\theta\in\Theta(e)} \left|\hat{\mathscr{L}}_T(\theta,e) - \mathscr{L}_T(\theta,e)\right| \to^P 0.$$

In conclusion, $\mathscr{L}_T(\hat{\theta},\hat{e}) \to^P \mathscr{L}_T(\theta_0,e_0)$ as desired. □

Note here that $\hat{e} = e_0$ w.p.a.1. Thus, we can treat e_0 as known when showing asymptotic normality of $\hat{\theta}$. That is, $\hat{\theta}$ is first–order equivalent to the oracle estimator where the true equilibrium is known to us.

To state the formal result, we first introduce

$$s(z_{m,t};\theta,e)=\frac{\partial \ell(z_{m,t};\theta,e)}{\partial \theta},\quad h(z_{m,t};\theta,e)=\frac{\partial^2 \ell(z_{m,t};\theta,e)}{\partial \theta \partial \theta'}, \tag{18}$$

and

$$H_T(\theta,e)=\frac{1}{T}\sum_{t=1}^{T} E\left[h(z_{m,t};\theta,e)\right].$$

Assumption 2. (i) $\theta_0 \in \text{int}\Theta(e_0)$; (ii) $\theta \mapsto \ell(z_{m,t};\theta,e_0)$ is almost surely twice continuously differentiable in a neighbourhood of θ_0; (iii) $\sum_{t=1}^{T} E\left[s(z_{m,t};\theta_0,e)\right]=0$ and $\sum_{t=1}^{T} E\left[\left\|s(z_{m,t};\theta_0,e_0)\right\|^2\right]<\infty$; $\left\|h(z_{m,t};\theta,e_0)\right\| \le b(z_{m,t})$ where $\sum_{t=1}^{T} E\left[b(z_{m,t})\right]<\infty$ for all $\theta \in \Theta_1(e_0)$ and $H_T(\theta_0,e_0)$ has full rank.

Theorem 3.2. *Under Assumption 1,* $\Pr\left(\hat{\theta}=\hat{\theta}_0\right)\to 1$, *where* $\hat{\theta}_0=\arg\max_{\theta\in\Theta(e_0)} \mathscr{L}_{M,T}(\theta,e_0)$ *is the oracle estimator. Thus, if in addition Assumption 2 also holds,*

$$\sqrt{M}(\hat{\theta}-\theta_0)=\sqrt{M}(\hat{\theta}_0-\theta_0)+o_P(1)\to^d N\left(0,H_{T,0}^{-1}\Omega_{T,0}H_{T,0}^{-1}\right),$$

where $H_{T,0}=H_T(\theta_0,e_0)$ *and* $\Omega_{T,0}=\sum_{t_1,t_2=1}^{T} E\left[s(z_{m,t_1};\theta_0,e_0)\,s(z_{m,t_2};\theta_0,e_0)'\right]/T^2$. *If the model is correctly specified then* $\Omega_{T,0}=H_{T,0}$.

Proof. As part of the proof of Theorem 3.1, it was shown that $P(||\hat{\theta}-\theta_0||<\varepsilon,\hat{e}=e_0)\to 1$ as $M\to\infty$. In particular, $\hat{\theta}=\hat{\theta}_0$ with probability appoaching one (w.p.a.1). Now, using Assumption 2(i)-(ii) and the mean–value theorem,

$$0=S_{M,T}\left(\hat{\theta},\hat{e}\right)=S_{M,T}\left(\hat{\theta},e_0\right)=S_{M,T}(\theta_0,e_0)+H_{M,T}\left(\bar{\theta},e_0\right)(\hat{\theta}-\theta_0),$$

w.p.a.1, for some $\bar{\theta}$ lying on the line segment conneting $\hat{\theta}$ and θ_0, where

$$S_{M,T}(\theta,e)=\frac{1}{MT}\sum_{m=1}^{M}\sum_{t=1}^{T} s(z_{m,t_1};\theta,e),\quad H_{M,T}(\theta,e)=\frac{1}{MT}\sum_{m=1}^{M}\sum_{t=1}^{T} h(z_{m,t_1};\theta,e). \tag{19}$$

By Assumption 2 (iii)-(iv) together with the CLT and the ULLN, $\sqrt{M}S_{M,T}(\theta_0,e_0)\to^d N(0,\Omega_{T,0})$ and $H_{M,T}\left(\bar{\theta},e_0\right)\to^p H_{T,0}$. The result now follows by Slutsky's Theorem. □

Remark. Above theorem assumes that $\theta_0 \in \text{int}\Theta(e_0)$, c.f. Assumption 2(i). If instead θ_0 is situated on the boundary of $\Theta(e_0)$ then the asymptotic distribution of the MLE becomes non–standard; see, e.g., Andrews (1999). The distribution depends on the curvature of the boundary of and so in order to develop such a theory, additional assumptions about the form of $\Theta(e_0)$ have to be imposed. If, for example, θ is a scalar and $\Theta(e_0)=[\theta_u,\theta_l]$ with $\theta_0=\theta_l$ situated on the boundary then $\sqrt{M}(\hat{\theta}-\theta_0)\to^d \min\{0,Z\}$, where $Z\sim N\left(0,H_{T,0}^{-1}\Omega_{T,0}H_{T,0}^{-1}\right)$.

Remark. If we strengthen Assumption 2(iii) to $E\left[s(z_{m,t};\theta_0,e)\,|z_{m,t-1},z_{m,t-2},...\right]=0$ then $\Omega_{T,0}$ simplifies to $\Omega_{T,0}=\sum_{t=1}^{T} E\left[s(z_{m,t};\theta_0,e_0)\,s(z_{m,t};\theta_0,e_0)'\right]/T^2$..

Assuming the model is correctly specified, above result allows us to use standard inference tools for θ_0. We can, for example, use the log-likelihood ratio statistic to test relevant hypotheses:

$$\begin{aligned} M \times \mathscr{LR}_{M,T} &= M\left\{\mathscr{L}_{M,T}\left(\hat{\theta},\hat{e}\right)-\mathscr{L}_{M,T}\left(\theta_0,\hat{e}\right)\right\} \\ &\simeq MS_{M,T}\left(\hat{\theta},\hat{e}\right)\left(\hat{\theta}-\theta_0\right)+\frac{M}{2}\left(\hat{\theta}-\theta_0\right)'H_{M,T}\left(\bar{\theta},\hat{e}\right)\left(\hat{\theta}-\theta_0\right) \\ &= \frac{M}{2}\left(\hat{\theta}-\theta_0\right)'H_{M,T}\left(\bar{\theta},\hat{e}\right)\left(\hat{\theta}-\theta_0\right)\rightarrow^{d}\chi^2_{d_\theta}, \end{aligned} \tag{20}$$

where $d_\theta = \dim(\theta)$.

Inference on equilbrium

The above analysis assumes that e_0 is well–separated from the other equilibria of the model in the population. This in turn implies that in large samples, we can treat e_0 as known and do not have to take into account any uncertainty about the true equilbrium. In small samples, , however, the log-likelihood evaluated at two different equilibria may be close to indistinguishable. We therefore here develop inference tools for the equilibrium.

First, write the estimated equilibrium as

$$\hat{e} = \arg\max_{e\in\mathscr{E}} \mathscr{L}_{M,T}\left(\hat{\theta}(e),e\right), \tag{21}$$

where

$$\hat{\theta}(e) = \arg\max_{\theta\in\Theta(e)} \mathscr{L}_{M,T}(\theta,e).$$

Observe that (21) is a random discrete choice problem. We will now show that in large samples $\hat{e}$ can be represented as the choice in a multinomial choice model. As a consequence, the distribution of $\hat{e}$ can be computed using tools from this literature. In order to achieve this, we need the following additional assumption:

$$\sqrt{M}(\hat{\theta}(e)-\theta_0(e))\rightarrow^{d} N(0,V_T(e)), \quad V_T(e) = H^{-1}_{T,0}(e)\,\Omega_{T,0}(e)\,H^{-1}_{T,0}(e),$$

where $\theta_0(e) = \arg\max_{\theta\in\Theta(e)}\mathscr{L}_T(\theta,e)$, $H_{T,0}(e) = H_T(\theta_0(e),e)$ and $\Omega_{T,0}(e) = \sum_{t_1,t_2=1}^{T} E\left[s(z_{m,t_1};\theta_0(e),e)\,s(z_{m,t_2}\right.$
That is, in any given equilibrium $e\in\mathscr{E}$, not just in the data–generating one, the log–likelihood $\mathscr{L}_{M,T}(\theta,e)$ is regular enough so that it has a unique maximum, $\theta_0(e)$, in the limit and its smooth around this maximum. Under this assumption, we have

$$\begin{aligned} & M\left\{\mathscr{L}_{M,T}\left(\hat{\theta}(e),e\right)-\mathscr{L}_{M,T}\left(\theta_0(e),e\right)\right\} \\ \simeq\ & MS_{M,T}\left(\hat{\theta}(e),e\right)\left(\hat{\theta}(e)-\theta_0(e)\right)+\frac{M}{2}\left(\hat{\theta}(e)-\theta_0(e)\right)'H_{M,T}\left(\bar{\theta}(e),e\right)\left(\hat{\theta}(e)-\theta_0(e)\right) \\ =\ & \frac{M}{2}\left(\hat{\theta}(e)-\theta_0(e)\right)'H_{M,T}\left(\bar{\theta}(e),e\right)\left(\hat{\theta}(e)-\theta_0(e)\right)\rightarrow^{d}\chi^2_{d_\theta}, \end{aligned} \tag{22}$$

for any $e \in \mathscr{E}$ while, by the CLT,

$$\sqrt{M}(\mathscr{L}_{M,T}(\theta_0(e),\cdot) - \mathscr{L}_T(\theta_0(e),\cdot)) \to^d Z(\cdot) \in R^{|\mathscr{E}|} \sim N\left(0, \text{Var}\left(\frac{1}{T}\sum_{t=1}^{T}\ell(z_{m,t};\theta,e)\right)\right), \quad (23)$$

where

$$Z(\cdot) \sim N\left(0, \text{Var}\left(\frac{1}{T}\sum_{t=1}^{T}\ell(z_{m,t};\theta,\cdot)\right)\right). \quad (24)$$

Combining (22)–(23),

$$Z_M(\cdot) := \sqrt{M}\left\{\mathscr{L}_{M,T}\left(\hat{\theta}(\cdot),\cdot\right) - \mathscr{L}_T(\theta_0(\cdot),\cdot)\right\} \to^d Z(\cdot), \quad (25)$$

Thus,

$$\hat{e} = \arg\max_{e\in\mathscr{E}}\{M\mathscr{L}_T(\theta_0(e),e) + Z_M(e)\}, \quad (26)$$

where $M\mathscr{L}_T(\theta_0(e),e)$ and $Z_M(e)$ are the deterministic and random component, respectively, of above random discrete choice problem. Due to (25), we have

$$\Pr(\hat{e} = e) = \Pr(\hat{e}_\infty = e) + o_P(1), \quad e \in \mathscr{E},$$

where

$$\hat{e}_\infty = \arg\max_{e\in\mathscr{E}}\left\{\sqrt{M}\mathscr{L}_T(\theta_0(e),e) + Z(e)\right\}. \quad (27)$$

This last problem corresponds to the multinomial Probit model, where $M\mathscr{L}_T(\theta_0(e),e)$ is the deterministic component of the "utility" of equilibrium e and $Z(e)$ is the normally distributed "taste shock". The distribution of $Z(\cdot)$ can be estimated by replacing $\text{Var}\left(\frac{1}{T}\sum_{t=1}^{T}\ell(z_{m,t};\theta,\cdot)\right)$ in (24) by its sample variance. Similarly, the deterministic component $\mathscr{L}_T(\theta_0(e),e)$ is consistently estimated by $\mathscr{L}_{M,T}\left(\hat{\theta}(e),e\right)$. Given these estimates, the conditional choice probabilities (CCP's) $\Pr(\hat{e}^* = e)$ can be computed using existing simulation–based methods for computation of CCP's of the multinomial Probit; see, e.g. Börsch-Supan and Hajivassiliou (1993). These methods are computationally very demanding though when $|\mathscr{E}|$ is large. One way of reducing this computational burden is to replace $\mathscr{E}$ by a subset $\mathscr{E}_0$ that only includes the most "likely" equilibria; for example, $\mathscr{E}_0 = \left\{e \in \mathscr{E} : \mathscr{L}_{M,T}\left(\hat{\theta}(\hat{e}),\hat{e}\right) - \mathscr{L}_{M,T}\left(\hat{\theta}(e),e\right) \le K\right\}$ for some cut-off point $K < \infty$.

Multiple equilibria played

Consider now the case where multiple equlibria are played across the different markets. We here discuss different estimators that can be used in this setting. As before θ_0 denote the true value of the structural parameters while $\mathbf{e}_{0,M} = (e_{0,1},....,e_{0,M})$ denotes the equilbria played across the M markets.

The most robust estimation method is to treat the equilibria played in the different markets as fixed effects leading to the following estimator,

$$(\hat{\theta}, \hat{\mathbf{e}}_M) = \arg \max_{\theta \in \Theta, \mathbf{e}_M \in \mathscr{E}^M} \mathscr{L}_{M,T}(\theta, \mathbf{e}_M).$$

This approach is pursued by, for example, Bajari, Benkard and Levin (2007). Hahn and Moon (2010) provide an asymptotic theory for this estimator in the case where $z_{m,t}$ is stationary and mixing. This property is essential for their analysis since it implies that the estimated equilibrium in a given market is consistent as $T \to \infty$ under regularity conditions similar to the ones found in Assumption 1 of the previous section (together with stationarity). However, their stationarity assumption is violated in our directional games since $x_{m,t}$ has an absorbing state which it reaches in finite time with probability 1. This in turn implies that their fixed effects estimator of the equilbrium played in market m cannot be consistent in our setting.

Below, we discuss potential alternatives:

Treating model with absorbing state as an approximation

Suppose that the state variable of the actual model of interest, say, $x_{m,t}$, is an explosive process without an absorbing state so that

$$\pi(x_{t+1} = x_t + 1 | x_t) = 1 - \pi(x_{t+1} = x_t | x_t) > 0.$$

This corresponds to a random walk with drift,

$$x_{t+1} = x_t + \varepsilon_{t+1},$$

where ε_{t+1} are i.i.d. with $\Pr(\varepsilon_{t+1} = 1) = 1 - \Pr(\varepsilon_{t+1} = 0) > 0$. This can be rewritten as

$$x_{t+1} = \mu + x_t + \tilde{\varepsilon}_{t+1}, \tag{28}$$

where $\mu = E[\varepsilon_t] = \Pr(\varepsilon_{t+1} = 1)$ and $\tilde{\varepsilon}_{t+1} = \varepsilon_{t+1} - \mu$.

The current leapfrogging model can be thought of as an "approximation" to above model where x_t is replaced by $\hat{x}_{K,t}$ with dynamics

$$\pi(\hat{x}_{K,t+1} = \hat{x}_{K,t} + 1 | \hat{x}_{K,t}) = 1 - \pi(\hat{x}_{K,t+1} = \hat{x}_{K,t} | \hat{x}_{K,t}) > 0 \text{ for } \hat{x}_{K,t} \leq K,$$

and

$$\pi(\hat{x}_{K,t+1} = K | \hat{x}_{K,t} = K) = 1.$$

As $K \to \infty$, $\hat{x}_{K,t} \to x_t$.

We could now proceed to do inference in the "exact model" where x_t evolves according to (28), and then treat the numerical implementation/estimation, where x_t is replaced by $\hat{x}_{K,t}$ as a numerical

approximation. My conjecture is that the results of Hahn and Moon (2010) can be extended to handle random walk type behaviour.

Partial identification inference

One could instead treat the problem as a partially identified one and analyze the identified set for θ_0 for a given fixed T; this could be done along the lines of Otsu, Pesendorfer, Takahashi and Sasaki (2022); see also Chesher, Rosen and Zhang (2024). However, their analysis does not allow for inference regarding $\mathbf{e}_{0,M}$.

Treat equilibrium selection as mixture model

Instead of being completely agnostic about the equilibrium selection and so treating the equilbria played in different markets as fixed effects, a perhaps more attractive option is to impose some restrictions on the equilibrium selection. One possible approach would be to assume that $e_{0,m}$ is drawn from some probability distribution $\Pr(e_{0,m} = e|x_{m,0} = x) = p(e|x)$, $e \in \mathscr{E}$. That is, we treat $e_{0,m}$ as a latent random variable and we then wish to identify the probabilities $p(e_{0,m} = e|x_{m,0})$, $e \in \mathscr{E}$. In this case, the likelihood for market m takes the form

$$\sum_{e \in \mathscr{E}} \left[\prod_{t=0}^{T} P^{(i)}(x_{m,t}, e; \theta)^{a^{(i)}_{m,t}} \left(1 - P^{(i)}(x_{m,t}, e; \theta)\right)^{1-a^{(i)}_{m,t}} \right] p(e|x_{m,0}),$$

where we wish to estimate θ and $\{p(e|x) : e \in \mathscr{E}, x \in X\}$. This corresponds to a mixture model. Identification results have been developed in Luo, Xiao and Xiao (2022) for finite T; see also Higgins and Jochmans (2023). However, above is a high–dimensional estimation problem when $|\mathscr{E}|$ is large. One could consider imposing some type of sparsity ($p(e|x_{m,0}) > 0$ for $e \in \mathscr{E}_0 \subset \mathscr{E}$ and zero otherwise) to resolve this issue.

4 Monte Carlo experiments

This section compares the performance of the NRLS estimator with a battery of exisiting estimators for dynamic games, using Monte Carlo experiments. We run several experiments of increasing complexity, starting from a game with unique equilibrium, several examples of games with multiple equilibria at the true parameter values, where data is generated from a single equilibrium, and finally presenting the case of multiple equilibria played in different markets. Before presenting the results we briefly summarize the existing estimators using the general dynamic game framework in section 2.1. Our results demonstrate that unlike the existing estimators NRLS is fully robust to the number of equilibria and reliably delivers maximum likelihood estimates of the structural parameters while identifying the equilibria played in the data.

4.1 Existing estimators

2-step PML and NPL estimator

Two-step pseudo maximum likelihood (PML) and nested pseudo likelihood (NPL) estimators are the most widely used estimators for dynamic games (**?**Pakes, Ostrovsky and Berry, 2007; Sanches, Silva and Srisuma, 2016).

EPL estimator

The efficient pseudo likelihood (EPL) estimator is a recent development in the literature on estimation of dynamic games Dearing and Blevins (2025).

MPEC estimators

Mathematical programming with equilibrium constraints (MPEC) is a general framework for structural estimation, which has been used in the literature to estimate various static and dynamic models including dynamic games Su and Judd (2012); lin Su (2013); Egesdal, Lai and Su (2015)

From the slides:

In the Monte Carlo exercises we compare the performance of the NRLS algorithm with a battery of standard estimation methods, namely mathematical programming with equilibrium constraints (MPEC) by

two step pseudo maximum likelihood estimator (PML2step) by Pakes, Ostrovsky and Berry (2007), and the sequential Nested Pseudo Likelihood estimator (NPL) implemented using the NPL algorithm suggested by AM2007. In this section we briefly summarize these estimators.

Implementation details from the slides:

- *Two-step estimator, NPL and EPL*
  - Matlab unconstrained optimizer (with numerical derivatives)
  - CCPs from frequency estimators
  - Max 120 iterations (for NPL and EPL)
- *MPEC*

  - Matlab constraint optimizer (interior-point) with analytic derivatives
  - MPEC-VP: Constraints on both values and choice probabilities (as in Egesdal, Lai and Su, 2015)
  - MPEC-P: Constraints in terms of choice probabilities + Hotz-Miller inversion (twice less variables)
  - Starting values from two-step estimator
- Estimated parameter k_1
- Sample size: *1000 markets in 5 time periods*
- Parameters are chosen to ensure good coverage of the state space and non-degenerate CCPs in all states

MPEC

It has been stated in the literature that constrained optimization approach does not require to solve for all the equilibria at each guess of structural parameter vector.

Given data: $\mathbf{Z} = \{\bar{\mathbf{a}}^{mt}, \bar{\mathbf{x}}^{mt}\}_{m\in\mathscr{M},t\in\mathscr{T},}$ the log of the augmented likelihood function is

$$\mathscr{L}(Z,P) = \frac{1}{M}\sum_{i=1}^{N}\sum_{m=1}^{M}\sum_{t=1}^{T}\log P_i(\bar{a}_i^{mt}|\bar{\mathbf{x}}^{mt};\theta)$$

The constrained optimization formulation of the ML estimation problem is

$$\begin{aligned}\max_{(\theta,\mathbf{P},\mathbf{V})} \quad & \mathscr{L}(\mathbf{Z},\mathbf{P}) \\ \text{subject to} \quad & \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V},\mathbf{P},\theta) \\ & \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V},\mathbf{P},\theta)\end{aligned}$$

The constrained optimization approach only needs to find those equilibria together with structural parameters that are local solutions and satisfy the corresponding first-order conditions. Constraints are satisfied (and an equilibrium solved) only at a solution, not at every iteration. These two features eliminate a large set of equilibria together with structural parameters that do not need to be solved.

Two step pseudo maximum likelihood (PML)

Step 1: nonparametrically estimate the conditional choice probabilities, denoted by $\hat{\mathbf{P}}$ directly from the observed data $\mathbf{Z}$

Step 2: solve

$$\begin{aligned}\max_{(\theta,\mathbf{P},\mathbf{V})} \quad & \mathscr{L}(\mathbf{Z},\mathbf{P}) \\ \text{subject to} \quad & \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V},\hat{\mathbf{P}},\theta) \\ & \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V},\hat{\mathbf{P}},\theta)\end{aligned}$$

or equivalently

$$\begin{aligned} \max_{(\theta,\mathbf{V})} \quad & \mathscr{L}(\mathbf{Z}, \Psi^{\mathbf{P}}(\mathbf{V}, \hat{\mathbf{P}}, \theta)) \\ \text{subject to} \quad & \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \hat{\mathbf{P}}, \theta) \end{aligned}$$

The PML2step estimator is defined as

$$\theta^{PML2step} = \arg\max_{\theta} \mathscr{L}(\mathbf{Z}, \Psi^{\mathbf{P}}(\Gamma(\theta, \hat{\mathbf{P}}), \hat{\mathbf{P}}, \theta))$$

Nested pseudo-likelihood (NPL)

The PML estimator can have large biases in finite samples In an effort to reduce the finite-sample biases associated with the PML estimator, AM2007 propose an NPL estimator. A NPL fixed point $(\tilde{\theta}, \tilde{\mathbf{P}})$ satisfies the conditions

$$\begin{aligned} \tilde{\theta} &= \arg\max_{\theta} \mathscr{L}(\mathbf{Z}, \Psi^{\mathbf{P}}(\Gamma(\theta, \tilde{\mathbf{P}}), \tilde{\mathbf{P}}, \theta)) \\ \tilde{\mathbf{P}} &= \Psi^{\mathbf{P}}(\Gamma(\theta, \tilde{\mathbf{P}}), \tilde{\mathbf{P}}, \theta) \end{aligned}$$

The NPL algorithm: For $1 \leq K \leq \bar{K}$, iterate over Steps 1 and 2: Step 1: Given $\tilde{\mathbf{P}}_{K-1}$, solve $\tilde{\theta}_K = \arg\max_{\theta} \mathscr{L}(\mathbf{Z}, \Psi^{\mathbf{P}}(\Gamma(\theta, \tilde{\mathbf{P}}_{K-1}), \tilde{\mathbf{P}}_{K-1}, \theta))$
Step 2: Given $\tilde{\theta}_K$, update $\tilde{\mathbf{P}}_K$ by
$\tilde{\mathbf{P}}_K = \Psi^{\mathbf{P}}(\Gamma(\theta_K, \tilde{\mathbf{P}}_{K-1}), \tilde{\mathbf{P}}_{K-1}, \theta_K)$
increase K by 1.

4.2 Experiment A: Unique equilibrium

To establish the performance baseline we start by briefly presenting the simplest case where there is a unique equilibrium in the model for all values of the parameters, and thus the single equilibrium data.

We use the same starting values for θ for all estimators. These are randomly perturbed true values. For the two-step CCP estimator we estimate the initial CCPs using a simple frequency estimator. For the sequential NPL algorithm, we initialize the CCPs using the same set of frequency estimates. For maximization of the pseudo likelihood we use standard build in unconstrained gradient based optimizer in Matlab (with numerical derivatives). For MPEC, we use the interior-point algorithm in Matlab constrained optimization solver, fmincon. As in Egesdal, Lai and Su (2015), we use both $\mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V},\mathbf{P})$ and $\mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V},\mathbf{P})$ as constraints and solve for the vector of smoothed value functions $\mathbf{V}$ and conditional choice probabilities (CCPs), $\mathbf{P}$, in addition to structural parameters θ. Similar to NPL, we use CCPs obtained from simple frequency estimators to initialize MPEC; in this case as starting values for the variables $\mathbf{P}$ in the constrained optimization problem. As starting values for the smoothed value functions, $\mathbf{V}$, we use the Hotz-Miller inversion to obtain the smoothed value functions implied by the estimated CCPs.

As the data generating process, we use the simultaneous move formulation of the leapfrogging game analyzed in **?**. The technological progress is indexed by state of the art production cost, c_t, that stochastic improves with probability $c_t/(1+c_t) \geq 0$ such that $\pi(c_{t+1}|c_t) > 0$ for $c_{t+1} \leq c_t$. Conditional on technological improvement the new state of the art is drawn from Beta distribution on $[c_t, 0]$ with parameters $a = 1.8$ and $b = 0.4$ allowing for both no, smaller and drastic changes in technology.[4] We discretize c_t in $n = 3$ points on uniform grid resulting in 14 points in state space of the model. The investment costs $K(c) = k_1 = 1$ and the common per period discount factor for the firms' discounted profits is $\beta = exp(-0.05) = 0.9512$.

	1-NPL	NPL	EPL	MPEC-VP	MPEC-P	NRLS
True $k_1 = 3.5$	3.52786	3.49714	3.49488	3.49488	3.49486	3.49488
Bias	0.02786	-0.00286	-0.00512	-0.00512	-0.00514	-0.00512
MCSD	0.10037	0.06522	0.07042	0.07042	0.07078	0.07042
ave log-likelihood	-1.166613	-1.161442	-1.161432	-1.161432	-1.161389	-1.161432
log-likelihood	-5833.066	-5807.208	-5807.159	-5807.159	-5806.945	-5807.159
log-like shortfall	-	-0.050	-0.000	-0.000	-0.000	-0.000
KL divergence	0.03254	0.00021	0.00024	0.00024	0.00024	0.00024
$\|\|P - P_0\|\|$	0.11270	0.00469	0.00495	0.00495	0.00500	0.00495
$\|\|\Psi(P) - P\|\|$	0.161849	0.000000	0.000000	0.000000	0.000000	0.000000
$\|\|\text{Bellman}(v) - v\|\|$	0.870946	0.000000	0.000000	0.000000	0.000000	0.000001
N runs of 100	100	100	100	100	99	100

Table 1: Monte Carlo experiment A: unique equilibrium

	2step	NPL	EPL	MPEC-VP	MPEC-P	NRLS
True k1=7.5	7.55163	7.49844	7.49918	7.65318	7.35124	7.49919
Bias	0.05163	-0.00156	-0.00082	0.15318	-0.14876	-0.00081
MCSD	0.17875	0.06062	0.03413	0.99742	0.47136	0.03413
ave log-likelihood	-0.84779	-0.84425	-0.84421	-0.88682	-0.87541	-0.84421
log-likelihood	-21194.856	-21106.332	-21105.126	-22170.397	-21885.366	-21105.126
log-like shortfall	-	-1.206	-0.000	-1062.740	-776.809	-0.000
KL divergence	0.02557	0.00040	0.00013	0.23536	0.16051	0.00013
$\|\|P-P_0\|\|$	0.11085	0.00490	0.00280	0.17466	0.20957	0.00280
$\|\|\Psi(P)-P\|\|$	0.170940	0.000000	0.000000	0.000000	0.000000	0.000000
$\|\|\text{Bellman}(v)-v\|\|$	1.189853	0.000000	0.000000	0.000000	0.000000	0.000001
N runs of 100	100	100	100	98	97	100

Table 2: Monte Carlo experiment B: small multiplicity with stable equilibrium played in the data.

4.3 Experiment B: Small multiplicity with stable equilibrium in the data

Table 2 presents the results of an experiment where there is only 3 equilibria when the model is solved at true parameters. Here, the scale parameter on the idiosyncratic shocks to investment is relatively high, $\sigma = 1$. All estimators seem to converge for each run; in the sense that tight tolerances are met to ensure that the objective function is locally maximized for all estimators and that equilibrium constraints are satisfied for MPEC.

The two step estimator (PML2step) is known less efficient than MLE and to suffer to from small sample bias, but fast to estimate and globally convergent even in the case with multiple equilibria. Given the large number of observations in this experiments it is not surprising that CCP type estimator produces parameter estimates that are quite close to their true value when CCPs are accurately estimated. While a frequency estimator should consistently estimate the equilibrium CCPs if only a single equilibrium played in the data, the CCPs are still estimated rather that being derived as an equilibrium from the structural model. Therefore equilibrium constraints of the equilibrium will not be satisfied to machine precision in finite samples, which we also confirm in the Monte Carlo experiment where $||\Psi^{\mathbf{P}}(P)-P|| = 0.03$ for PML2step.

The NPL algorithm should improve the efficiency of the two step estimator by recursively updating CCPs and (if it converges) thereby impose constraint $||\Psi^{\mathbf{P}}(P)-P|| = 0$ rather than holding the CCP's fixed at the estimates from the first step. The NPL algorithm does seem to converge, improve on the pseudo likelihood and reduce $||\Psi^{\mathbf{P}}(P)-P||$. Both NRLS and MPEC fully impose the constraints from the model as part of the estimation procedure such that $||\Psi^{\mathbf{P}}(P)-P||$ and $||\Psi^{\mathbf{V}}(V)-V||$ are close to zero (up to the selected convergence criteria). However, while both MPEC (and NPL) seem to converge they converge to local maxima of the likelihood function. Since there are multiple equilibria in the model, it is possible to find a combination of P, V and θ that satisfies the constraints but will not maximize the likelihood. In contrast, NRLS will always select the equilibrium that maximize the likelihood for given value θ.

As mentioned above, we initialize MPEC using the estimated CCPs for $\mathbf{P}$ and the corresponding $\mathbf{V}$ implied by the Hotz-Miller inversion. But we are considering a hypothetical and ideal situation

[4]The process for the state of the art marginal cost of production is directional in the sense that it only improves: making the transition probability matrix for state of the art cost is lower triangular, i.e. $\pi(c_{t+1}|c_t) = 0$ if $c_{t+1} > c_t$.

with a fairly large number of observations uniformly distributed over the entire state space and therefore very accurate estimates of **P**. It is therefore not surprising that MPEC does relatively well. In smaller samples or larger state spaces, estimates of **P** will be more noisy and will therefore provide poorer starting values for MPEC and NPL. To illustrate this problem with lack of convergence and convergence to local minima we redo the experiment from Table **??** with starting values of θ and P that are further away from their true values. We keep the number of observations high to isolate the effect the initialization of the iterative algorithms.

The result of this experiment is presented in Table **??**. Both NPL and MPEC now reveals serious convergence problems. When MPEC does converge to a local maximum with satisfied constraints, it is often at values of the structural parameters that are far from the true ones and at a much lower likelihood values. NRLS on the other hand always converge to a global maximum that (with guarantee) satisfies the constraints, identifies the equilibrium played in the data and estimates the structural parameters without much bias and low variance.

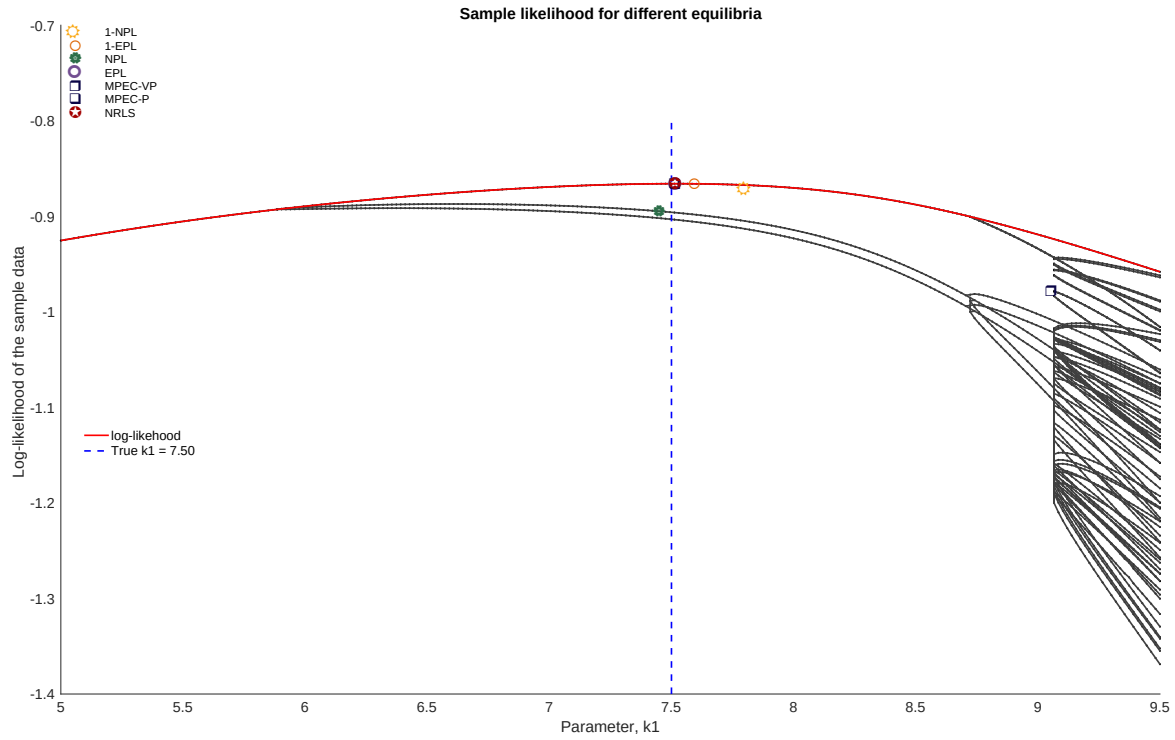
Our implementation of MPEC is admittedly fully optimized but we do implement analytical first derivatives of the objective function. While this this will affect computational speed for MPEC, the convergence properties of the method are not affected. The instability of MPEC is an artifact of using iterative algorithm that needs to search over a large parameter space. A tempting solution is to initialize MPEC with multiple combinations of starting values. However, using multi-start procedures for estimation of models with a large multiplicity of equilibria is by no means as an exhaustive and systematic approach as the one we suggest. Our experimenting with multi starts for MPEC leads to the discouraging result that starting values of MPEC inadvertently imposes artificial equilibrium selection that result in many different local maxima for the same data set often without finding the global maximum.

4.4 Experiment C: Small multiplicity with *unstable* equilibrium in the data

	2step	NPL	EPL	MPEC-VP	MPEC-P	NRLS
True k1=7.5	7.54238	7.39276	7.48044	7.73133	7.63100	7.50176
Bias	0.04238	-0.10724	-0.01956	0.23133	0.13100	0.00176
MCSD	0.17145	0.05608	0.15801	0.72988	0.89874	0.03820
ave log-likelihood	-0.86834	-0.89374	-0.86550	-0.88512	-0.90196	-0.86504
log-likelihood	-21708.592	-22343.584	-21637.535	-22127.906	-22549.056	-21626.121
log-like shortfall	-	-765.242	-11.413	-502.121	-920.643	-0.000
KL divergence	0.02271	0.15996	0.00257	0.11452	0.20182	0.00012
$\|\|P - P_0\|\|$	0.09757	0.20709	0.00619	0.03860	0.02504	0.00307
$\|\|\Psi(P) - P\|\|$	0.160102	0.000000	0.000000	0.000000	0.000000	0.000000
$\|\|\text{Bellman}(v) - v\|\|$	1.126738	0.000000	0.000000	0.000000	0.000000	0.000001
N runs of 100	100	18	100	99	98	100

Table 3: Monte Carlo experiment C: small multiplicity with unstable equilibrium played in the data.

Figure 6: Profile likelihood and estimator convergence I



Note: The data is generated from the unstable equilibrium in the model with three equilibria at the true parameter. NPL converged to one of the stable equilibria. Two-step estimators do not fall to any likelihood profile (because they do not impose equilibrium constraints). NRLS and EPL show similar performance. MPEC converged to a local maximum among many equilibria on the right.

4.5 Experiment D: Discontinuous likelihood function

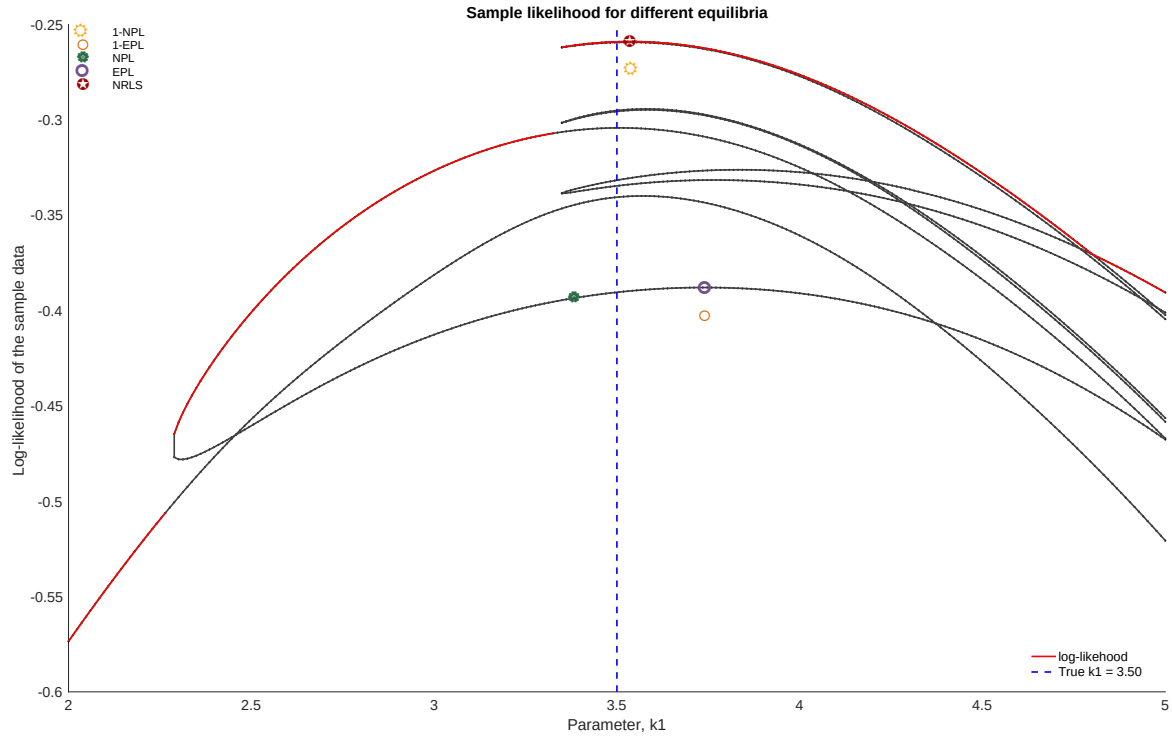
	2step	NPL	EPL	MPEC-VP	MPEC-P	NRLS
True k1=3.5	3.49739	3.55144	3.64772	3.65943	3.67027	3.50212
Bias	-0.00261	0.05144	0.14772	0.15943	0.17027	0.00212
MCSD	0.13999	0.07133	0.12900	0.12693	0.11583	0.03255
ave log-likelihood	-0.27494	-0.29474	-0.29528	-0.30330	-0.30257	-0.25086
log-likelihood	-1374.721	-1473.695	-1476.425	-1516.503	-1512.847	-1254.320
log-like shortfall	-	-219.375	-222.104	-270.999	-267.523	-0.000
KL divergence	0.01512	0.04889	0.04495	0.04102	0.04078	0.00016
$\|\|P - P_0\|\|$	0.62850	0.86124	0.83062	0.66562	0.65879	0.01610
$\|\|\Psi(P) - P\|\|$	0.763764	0.000000	0.000000	0.000000	0.000000	0.000002
$\|\|\text{Bellman}(v) - v\|\|$	0.852850	0.000000	0.000000	0.000000	0.000000	0.000005
N runs of 100	100	100	100	28	27	100

Table 4: Monte Carlo experiment D: Discontinuous likelihood function.

Figure 7

The full solution approach we take allows for investigation the performance of all considered estimators in great detail. In this section we present a series of typical estimation outcomes graphically to illustrate where different estimators fall in the likelihood surface. Of particular interest is the performance of the estimators which impose equilibrium constraints but fail to deliver an unbiased estimate of the structural parameters, as seen in the previous sections. How multiplicity of equilibria in the model impairs the performance of such estimators is seen seen in the following figures which not only plot the true likelihood surface (maximum over parameters and equilibria), but also show the likelihood conditional on equilibrium selection. It then becomes evident that MPEC, NPL and EPS may converge at an equilibrium with lower likelihood than the one that generated the data.

Figure 7: Profile likelihood and estimator convergence I



Note: The data is generated from ...

4.6 Experiment E: Massive multiplicity

	1-NPL	NPL	EPL	NRLS
True k1=3.75	3.70959	3.71272	3.78905	3.74241
Bias	-0.04041	-0.03728	0.03905	-0.00759
MCSD	0.11089	0.06814	0.40716	0.03032
ave log-likelihood	-0.38681557	-0.37348793	-0.45256293	-0.35998461
log-likelihood	-1934.078	-1867.440	-2262.815	-1799.923
log-like shortfall	-	-66.529	-467.607	-0.000
KL divergence	Inf	14.07523	12231.59186	0.32429
$\|\|P - P_0\|\|$	0.82204	0.65580	0.79241	0.07454
$\|\|\Psi(P) - P\|\|$	0.963574	0.000000	0.000000	0.000006
$\|\|\Gamma(v) - v\|\|$	7.020899	0.000000	0.000000	0.000008
N runs of 100	100	18	68	100
CPU time	0.159s	11.262s	4.013s	4.731s

Table 5: Monte Carlo experiment E: Massive multiplicity

In Table 5

Evidently, the problem with local optima for MPEC persists and we begin to see serious convergence problems for the NPL algorithm, although it delivers reasonable precise estimates when it does converge. NRLS always converges and is able to deliver precise estimates of the parameters.

4.7 Experiment F: Convergence properties of NRLS

4.8 Experiment G: Multiple equilibria in the data

5 Discussion and conclusions

This draft contains preliminary results on comparing the new full solution maximum likelihood estimator for directional dynamic games, with a battery of existing estimators for the models of strategic interaction. Our estimator is based on the recursive lexicographic search (RLS) algorithm developed in IRS2016 which is capable of computing all Markov perfect equilibria in the class of directional dynamic games. With all equilibria numerically computed by this method for each value of parameters, the likelihood can be maximized both with respect to the usual vector of structural parameters, and the discrete equilibrium selection rule that can be thought of as an additional nuisance parameter.

We document limited performance by all standard estimation methods which worsens as the number of equilibria increases in the theoretical model at true value of parameters. The proposed nested RLS estimator (NRLS) is robust to the number of equilibria in the theoretical model. We also show that it can identify the exact data generating equilibria in the setup where data contains several equilibria played at different types of markets.

References

AGUIRREGABIRIA, V. AND P. MIRA (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70(4), 1519–1543.

——— (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75, No. 1, 1–53.

ANDERSON, A., J. ROSEN, J. RUST AND K.-P. WONG (2025): "Disequilibrium Play in Tennis," *Journal of Political Economy*, 133(1), 190–251.

ANDREWS, D. (1999): "Estimation when the Parameter is on the Boundary," *Econometrica*, 67, 1341–1383.

BAJARI, P., C. L. BENKARD AND J. LEVIN (2007): "Estimating dynamic models of imperfect competition," *Econometrica*, 75(5), 1331–1370.

BESANKO, D., U. DORASZELSKI, Y. KRYUKOV AND M. SATTERTHWAITE (2010): "Learning-by-doing, organizational forgetting, and industry dynamics," *Econometrica*, 78(2), 453–508.

BORKOVSKY, R. N., U. DORASZELSKI AND Y. KRYUKOV (2010): "A user's guide to solving dynamic stochastic games using the homotopy method," *Operations Research*, 58(4-part-2), 1116–1132.

BÖRSCH-SUPAN, A. AND V. A. HAJIVASSILIOU (1993): "Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models," *Journal of Econometrics*, 58(3), 347–368.

CHESHER, A., A. ROSEN AND Y. ZHANG (2024): "Robust Analysis of Short Panels," *arXiv:2401.06611*.

DEARING, A. AND J. R. BLEVINS (2025): "Efficient and Convergent Sequential Pseudo-Likelihood Estimation of Dynamic Discrete Games," *The Review of Economic Studies*, 92(2), 981–1021.

DUBE, J.-P. H., G. J. HITSCH AND P. K. CHINTAGUNTA (2010): "Tipping and concentration in markets with indirect network effects," *Marketing Science*, 29(2), 216–249.

EGESDAL, M., Z. LAI AND C.-L. SU (2015): "Estimating dynamic discrete-choice games of incomplete information," *Quantitative Economics*, 6(3), 567–597.

HAHN, J. AND H. R. MOON (2010): "Panel Data Models with Finite Number of Multiple Equilibria," *Econometric Theory*, 26(3), 863–881.

HIGGINS, A. AND K. JOCHMANS (2023): "Identification of mixtures of dynamic discrete choices," *Journal of Econometrics*, 237, 105462.

HOTZ, V. J. AND R. A. MILLER (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60(3), 497.

ISKHAKOV, F., J. RUST AND B. SCHJERNING (2016): "Recursive Lexicographical Search: Finding all Markov Perfect Equilibria of Finite State Directional Dynamic Games," *The Review of Economic Studies*, 83(2), 658–703.

——— (2018): "The Dynamics of Bertrand Price Competition with Cost-reducing Investments," *International Economic Review*, 59(4), 1681–1731.

JUDD, K. L., K. SCHMEDDERS AND E. YELTEKIN (2012): "Optimal Rules for Patent Races," *International Economic Review*, 53(1), 23–52.

KASAHARA, H. AND K. SHIMOTSU (2012): "Sequential Estimation of Structural Models with a Fixed Point Constraint," *Econometrica*, 80(5), 2303–2319.

LAND, A. H. AND A. G. DOIG (1960): "An Automatic Method of Solving Discrete Programming Problems," *Econometrica*, 28(3), 497–520.

LIN SU, C. (2013): "Estimating Discrete-Choice Games of Incomplete Information: Simple Static Examples," Discussion paper, SSRN.

LUO, Y., P. XIAO AND R. XIAO (2022): "Identification of dynamic games with unobserved heterogeneity and multiple equilibria," *Journal of Econometrics*, 226, 343–367.

MA, Q. AND J. STACHURSKI (2021): "Dynamic Programming Deconstructed: Transformations of the Bellman Equation and Computational Efficiency," *Operations Research*, 69(5), 1591–1607.

NEWEY, W. AND D. MCFADDEN (1993): "Large sample estimation and hypothesis testing," in *Handbook of Econometrics, Vol. 4*, ed. by R. Engle and D. McFadden, chap. 36. Elsevier Science B.V.

OTSU, T., M. PESENDORFER, Y. TAKAHASHI AND Y. SASAKI (2022): "Estimation of (static or dynamic) games under equilibrium multiplicity," *International Economic Review*, 63, 1165–1188.

PAKES, A. AND P. MCGUIRE (1994): "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," *The Rand Journal of Economics*, pp. 555–589.

PAKES, A., M. OSTROVSKY AND S. BERRY (2007): "Simple estimators for the parameters of discrete dynamic games (with entry/exit examples)," *the RAND Journal of Economics*, 38(2), 373–399.

PESENDORFER, M. AND P. SCHMIDT-DENGLER (2010): "Sequential Estimation of Dynamic Discrete Games: A Comment," *Econometrica*, 78(2), 833–842.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

SANCHES, F. A. M., D. J. SILVA AND S. SRISUMA (2016): "Ordinary Least Squares Estimation of a Dynamic Game Model," *International Economic Review*, 57(2), 623–633.

SHAPLEY, L. S. (1953): "Stochastic Games," *Proceedings of the National Academy of Sciences of the United States of America*, 39(10), 1095–1100.

SU, C.-L. AND K. L. JUDD (2012): "Constrainted Optimization Approaches to Estimation of Structural Models," *Econometrica*, 80(5), 2213–2230.