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

June, 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.

†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

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 Iskhakov, Rust and Schjerning (2016) 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 Iskhakov, Rust and Schjerning (2016) 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, Aguirregabiria and Mira (2007) 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 Aguirregabiria and Mira (2007) 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 Iskhakov, Rust and Schjerning (2018). 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 Computing all equilibria in directional dynamic games

In this section we set up a classic dynamic discrete game of incomplete information relying on Aguirregabiria and Mira (2007), and introduce additional assumptions which characterize directional dynamic games as a special case of the class model. We then introduce the dynamic Bertrand pricing game from Iskhakov, Rust and Schjerning (2018) as an example of a directional dynamic game. Using this example, we then summarize recursive lexicographic search (RLS) algorithm to compute all Markov perfect equilibria of directional dynamic games.

2.1 Dynamic games of incomplete information

We follow Aguirregabiria and Mira (2007) to set up a canonical dynamic discrete game of incomplete information as a game of N players in discrete time $t = 1, 2, \ldots$. At each time t, each player i chooses an action a_{it} from a finite set of actions $A = \{1, \ldots, J\}$. In the beginning of each period t each player is characterized by a pair of vectors s_{it} and ε_{it}, where the former is common knowledge among to all players, and ε_{it} is a vector of private information of firm i. The observed state of the game at time t is given by a vector s_t which stacks all the states of all players.

The Markov perfect equilibria (MPE) in the game satisfy

$$\begin{cases} \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta), \\ \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta), \end{cases} \tag{1}$$

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 (12) gives a Markov perfect equilibrium of the game.

2.2 Directional dynamic games

Directional dynamic games were introduced by Iskhakov, Rust and Schjerning (2016) as a subclass of discrete dynamic games. We briefly summarize the theory of directional dynamic games (DDG) in the terms introduced in the previous section. In Aguirregabiria and Mira (2007) notation, the the conditions of the formal definition of DDGs are equivalent to the block-triangularity of the choice-specific state transition probability matrices of the state variables.

Roughly speaking, a game $\mathscr{G}$ is directional if we can single out some dimensions of the state space S which we call "directional" components, such that the transitions between the points in these dimensions can be represented by a *directed acyclic graph* (DAG).

Directionality in the stochastic evolution of the states in a game $\mathscr{G}$ can be captured by defining a *partial order* over the state space S. This partial order of the states will generally be *strategy specific* since the stochastic evolution of the states will generally depend on the strategies σ used by the

players, and we use the symbol $\succ_\sigma$ to indicate this dependence. Most games that we analyze will exhibit directionality only in a subvector of the full vector of state variables. Therefore our definition assumes there is a decomposition of S as a cartesian product of two sets D and X, so a generic element of the state space is written as $s = (d,x)$ where we refer to d as the *directional component* of the state space, and x as the *non-directional* component. The partial order $\succ_\sigma$ is defined over the directional component D of the state space S.

In the definition below, we let $\rho(d'|d,x,\sigma)$ denote the *conditional hitting probability,* i.e. the conditional probability that a state with directional component d' is *eventually* reached given that the process starts in state $s = (d,x)$ and the players use strategy σ. Note that $\rho(d'|d,x,\sigma)$ is different from a single step transition probability. In the terminology of Markov chains, $\rho(d'|d,x,\sigma)$ is the probability that the *hitting time* of the set $(d' \times X) = \{(d',x')|x' \in X\}$ is finite conditional on starting in state (d,x) under strategy σ. The hitting time (or *first passage time*) is the smallest time it takes for the state to travel from state $s = (d,x)$ to some state (d',x') where $x' \in X$.

Definition 1 (Strategy specific partial order over states)**.** *Let σ be a feasible n-tuple of strategies for the players in the dynamic game $\mathscr{G}$. Suppose S is a finite subset of R^k that can be decomposed as a cartesian product $S = D \times X$ where $D \subset R^N$, $X \subset R^{k-N}$ and $N \leq k$. A typical element of S is a point $s = (d,x) \in D \times X$, where we allow for the possibility that D or X is a single point (to capture the cases where S has no directional component and the case where S has no non-directional component, respectively). Then a binary relation $\succ_\sigma$ over the directional components $d \in D$ induced by the strategy profile σ is defined as*

$$d' \succ_\sigma d \;\; \textit{iff} \;\; \exists x \in X \;\; \textit{such that} \;\; \rho(d'|d,x,\sigma) > 0 \;\; \textit{and} \;\; \forall x' \in X \; \rho(d|d',x',\sigma) = 0. \tag{2}$$

It can be shown that the binary relation $\succ_\sigma$ is a partial order of the set D.

The partial order of the states captures the directionality in the game implied by the strategy σ. The statement $d' \succ_\sigma d$ can be interpreted intuitively as saying that the directional component d' comes *after* the directional state d in the sense that there is a positive probability of going from d to d' but zero probability of returning to d from d'. Note that $\succ_\sigma$ will generally not be a *total order* of the directional components D because there may be pairs $(d',d) \in D \times D$ that are *non-comparable* with respect to the partial order $\succ_\sigma$. There are two ways in which a pair of points (d',d) can be non-comparable (a situation that we denote by $d' \not\succ d$): there may be no communication between d and d', i.e. zero probability of hitting state d' from d and vice versa, or there may be a two way transition (a *loop*) connecting d and d', i.e. d' can be reached with positive probability from d and vice versa.

The asymmetry and transitivity conditions guarantee that there cannot be any loops between any of the comparable pairs (d',d) of a strict partial order $\succ_\sigma$. However, loops that may exist between *non-comparable* pairs (d',d) that are not elements of the binary relation $\succ_\sigma$, also need to be ruled out.

Definition 2 (No loop condition)**.** *Let σ be a feasible n-tuple of strategies for the players in the dynamic game $\mathscr{G}$. We say that σ has no loops in the directional component D if the following condition is satisfied for all $d' \neq d \in D$*

$$d' \not\succ_\sigma d \Longrightarrow \forall x \in X \;\; \rho(d'|d,x,\sigma) = 0 \;\; \textit{and} \;\; \forall x' \in X \; \rho(d|d',x',\sigma) = 0. \tag{3}$$

It is not hard to show that when the No Loop Condition is satisfied for a feasible strategy σ, the transitions among the directional components of the state vector d induced by this strategy can be represented by a DAG. Let $D(\mathscr{G},\sigma)$ denote a directed graph with nodes corresponding to elements of D and edges connecting the points d and d' if the hitting probability $\rho(d'|d,x,\sigma)$ is positive. Then if d and d' are comparable with respect to $\succ_\sigma$, there can only be an edge from d to d' or vise versa, and otherwise if d and d' are not comparable there is no edge between them due to no communication by the No Loop Condition. Therefore, the directed graph $D(\mathscr{G},\sigma)$ does not have loops, thus it is a DAG.

Different strategies σ can potentially induce different partial orders of the directional component of the state space D. To be able to construct a common total order for the state recursion algorithm, it is important to ensure that strategy specific partial orders are *consistent* with each other, i.e. that there is no pair of states for which d' follows from state d under strategy σ but d follows from d' under σ'.

Definition 3 (Consistent partial orders)**.** *Let σ and σ' be any two feasible n-tuple of strategies for the players in the dynamic game $\mathscr{G}$ and let $\succ_\sigma$ and $\succ'_\sigma$ be the two corresponding induced partial orders of the directional component of the state space D. We say that $\succ_\sigma$ and $\succ_{\sigma'}$ are pairwise consistent if and only if for any $d',d \in D$ we have*

$$\begin{aligned} &\text{if } d' \succ_\sigma d \text{ then } d \not\succ_{\sigma'} d', \text{ and} \\ &\text{if } d \succ_\sigma d' \text{ then } d' \not\succ_{\sigma'} d, \end{aligned} \tag{4}$$

in other words if a transition from d to d' is possible under strategy σ, the opposite transition should not be possible under some other strategy σ', and vice versa.

It is worth noting that the definition of consistency is silent about the non-directional component of the state space, allowing for various strategies to induce any transitions between points that only differ in non-directional dimensions. Given the concept of consistent partial orders, we can formally define the concept of a *directional dynamic game* (DDG).

Definition 4 (Directional Dynamic Games)**.** *We say that a dynamic Markovian game $\mathscr{G}$ with finite state space S and set of all feasible strategies $\Sigma(\mathscr{G})$ is a directional dynamic game (DDG) if given the decomposition of the state space into directional and non-directional components $S = D \times X$, the following two conditions hold: (1) every strategy $\sigma \in \Sigma(\mathscr{G})$ has no loops in the directional component D according to Definition 2, and (2) the elements of the set of induced partial orders on D, $\{\succ_\sigma | \sigma \in \Sigma(\mathscr{G})\}$, are pairwise consistent according to Definition 3.*

2.3 Example: leapfrogging game

The Monte Carlo exercises in this paper are based on the dynamic model of Bertrand competition with cost-reducing investments studied in Iskhakov, Rust and Schjerning (2018) (also known as leapfrogging model). 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 (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 consider a stochastic dynamic game where two Bertrand competitors indexed by $i \in \{1,2\}$ set prices of their homogeneous product and decide whether or not to invest in the latest (state of the art) technology that will allow them to produce at marginal cost $c \leq c_i$. Time is discrete and each firm maximizes expected discounted profits over an infinite horizon ($t = 1,2,\ldots,\infty$). Both firms have a common discount factor $\beta \in (0,1)$. The state of the art marginal cost of production start at an initial value c_0 and follows *exogenous* Markov process with transition probability $\pi(c_{t+1}|c_t)$. The state of the art technology can only improve such that $\pi(c_{t+1}|c_t) = 0$ if $c_{t+1} > c_t$

The investment choice is dichotomous. After paying an investment cost of $K(c)$, the investing firm obtains marginal cost c after one period lag. The state of the game is given by the cost structure of the two firms and the state-of-the-art cost, (c_1, c_2, c). Because of the directionality of the state state transitions, the state space is limited to $X = (c_1, c_2, c) \subset R^3$: $c_1 \geq c$, $c_2 \geq c$. These state is common knowledge. All actions are observable, such that the pricing decisions, $p_i \in R^+$ and investment actions $a_i \in \{I, N\}$ are common knowledge. In each period each firm also incurs additive costs (benefits) from not investing and investing, $\eta\varepsilon_{i,I}$ and $\eta\varepsilon_{i,N}$, where $\eta \geq 0$ is a scaling parameter. These idiosyncratic shocks to investment are private information and assumed to have extreme value Type 1 distribution, to be independent across choice, time and firms.

The strategy profile is a pair of Markovian *behavior* strategies $\sigma = (\sigma_1, \sigma_2)$, where

$$\sigma_i = \Big(p_i(c_1,c_2,c), P_i(c_1,c_2,c)\Big) \in \mathbb{R}_+ \times [0,1]. \tag{5}$$

The choice probability $P_i(c_1,c_2,c)$ denotes the probability that firm i invests in state-of-the-art technology, and $p_i(c_1,c_2,c)$ is the Bertrand pricing decision of firm i. Under Bertrand pricing, the equilibrium prices are given by $p_i(c_1,c_2,c) = \max(c_1,c_2)$ and the expected one period profit of firm $i \neq j$ is

$$r^i(c_1,c_2) = \begin{cases} 0, & \text{if } c_i \geq c_j, \\ c_j - c_i, & \text{if } c_i < c_j. \end{cases} \tag{6}$$

The investment function $P_j(c_1,c_2,c)$ must maximize the expected discounted value of firm j's future profit stream taking into account the investment and pricing strategies of its opponent. The value functions for the firm $i = 1$ take the form

$$V_i(c_1,c_2,c) = \max\left\{v_i^I(c_1,c_2,c) + \eta\varepsilon_{i,I};\ v_i^N(c_1,c_2,c) + \eta\varepsilon_{i,N}\right\}, \tag{7}$$

where the deterministic parts of the choice specific value functions are given by

$$v_i^N(c_1,c_2,c) = r^i(c_1,c_2) + \beta EV_i(c_1,c_2,c|N), \tag{8}$$

$$v_i^I(c_1,c_2,c) = r^i(c_1,c_2) - K(c) + \beta EV_i(c_1,c_2,c|I). \tag{9}$$

With extreme value shocks, the investment probability has the well known logit form

$$P_i^I(c_1,c_2,c) = \frac{\exp\{v_I^I(c_1,c_2,c)/\eta\}}{\exp\{v_i^I(c_1,c_2,c)/\eta\} + \exp\{v_i^N(c_1,c_2,c)/\eta\}}. \tag{10}$$

The expected value functions are given by the following expressions, which perform integration over future state-of-the-art cost c, investment choices of the firm $j = 2$, and idiosyncratic components $\varepsilon_{i,I}$ and $\varepsilon_{i,N}$ in the following period:

$$\begin{aligned} EV_i(c_1,c_2,c|N) &= \int_0^c \left[P_j^I(c_1,c_2,c)H_i(c_1,c,c') + P_j^N(c_1,c_2,c)H_i(c_1,c_2,c')\right]\pi(dc'|c), \\ EV_i(c_1,c_2,c|I) &= \int_0^c \left[P_j^I(c_1,c_2,c)H_i(c,c,c') + P_j^N(c_1,c_2,c)H_i(c,c_2,c')\right]\pi(dc'|c), \end{aligned}$$

where the logsum function is denoted

$$H_i(c_1,c_2,c) = \eta \log\left[\exp\left(v_i^N(c_1,c_2,c)/\eta\right) + \exp\left(v_i^I(c_1,c_2,c)/\eta\right)\right]. \tag{11}$$

The Markov perfect equilibrium of Bertrand investment stochastic game is a pair of strategy profiles $\sigma^* = (\sigma_1^*, \sigma_2^*)$ and a pair of *value functions* $V(s) = \Big(V_1(s), V_2(s)\Big)$, $V_i : S \to R$ such that the Bellman equations are satisfied for each firm, and strategies σ_1^* and σ_2^* constitute mutual best responses, in particular assign positive probabilities only to the actions in the set of maximizers of the Bellman equations. Stacking equations (8), (9) and (10) over all possible states (c_1, c_2, c) leads to the shorthand notation (12) for the leapfrogging game.

2.4 Full solution of directional dynamic games

The recursive lexicographical search (RLS) algorithm of Iskhakov, Rust and Schjerning (2016) 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 Iskhakov, Rust and Schjerning (2016) 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 (topmost) level. Every path from trunk to any leaf goes through every level, and each level corresponds to a point in the state space.

3 Estimating games with multiple equilibria

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

Using notation of Aguirregabiria and Mira (2007), the Markov perfect equilibria (MPE) in a dynamic stochastic game can be written as

$$\begin{cases} \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta), \\ \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta), \end{cases} \tag{12}$$

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 (12) gives a Markov perfect equilibrium of the game.

Denote the set of all solution to the system (12) and thus all MPE by

$$SOL(\Psi, \theta) = \left\{ (\mathbf{P}, \mathbf{V}) \;\middle|\; \begin{array}{l} \mathbf{V} = \Psi^{\mathbf{V}}(\mathbf{V}, \mathbf{P}, \theta) \\ \mathbf{P} = \Psi^{\mathbf{P}}(\mathbf{V}, \mathbf{P}, \theta) \end{array} \right\}. \tag{13}$$

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

$$\mathscr{L}(\mathbf{Z}, \theta) = \max_{\mathbf{P} \in SOL(\Psi, \theta)} \sum_{m=1}^{M} \sum_{t=1}^{T} \sum_{i=1}^{N} \left\{ \log \mathbf{P}_i(a_i^{mt} | \mathbf{x}^{mt} m, \theta) + \log f(\mathbf{x}^{mt} | \mathbf{x}^{mt-1}, a_i^{mt-1}, \theta) \right\} \tag{14}$$

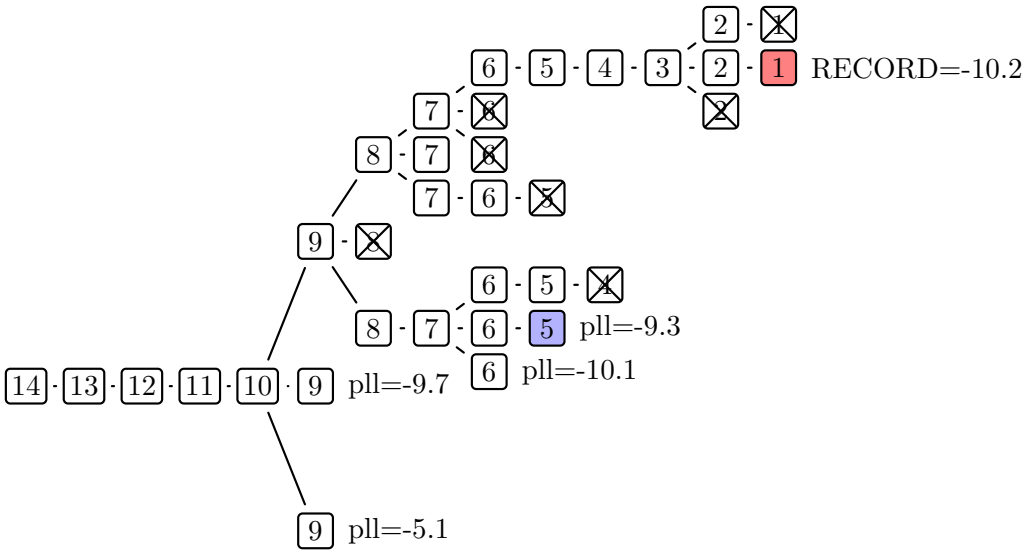
$$= \max_{\mathbf{P} \in SOL(\Psi, \theta)} \mathscr{L}^{(1)}(\mathbf{Z}, \theta, \mathbf{P}) + \mathscr{L}^{(2)}(\mathbf{Z}, \theta), \tag{15}$$

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

$$\hat{\theta}^{ML} = \arg\max_{\theta \in \Theta} \mathscr{L}(\mathbf{Z}, \theta). \tag{16}$$

The proposed ML estimator (16) 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 1: 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 (16). 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 (14)-(15). The need to compute all solutions to (12) 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 (15) 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 (14) 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 (15). 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

$$\mathscr{L}^{(1)}(\mathbf{Z}_k, \theta, \mathbf{P}) \geq \mathscr{L}^{(1)}(\mathbf{Z}_{k'}, \theta, \mathbf{P}) \text{ for } \forall k < k' \in \{1, \ldots, K\}. \tag{17}$$

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 2 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.

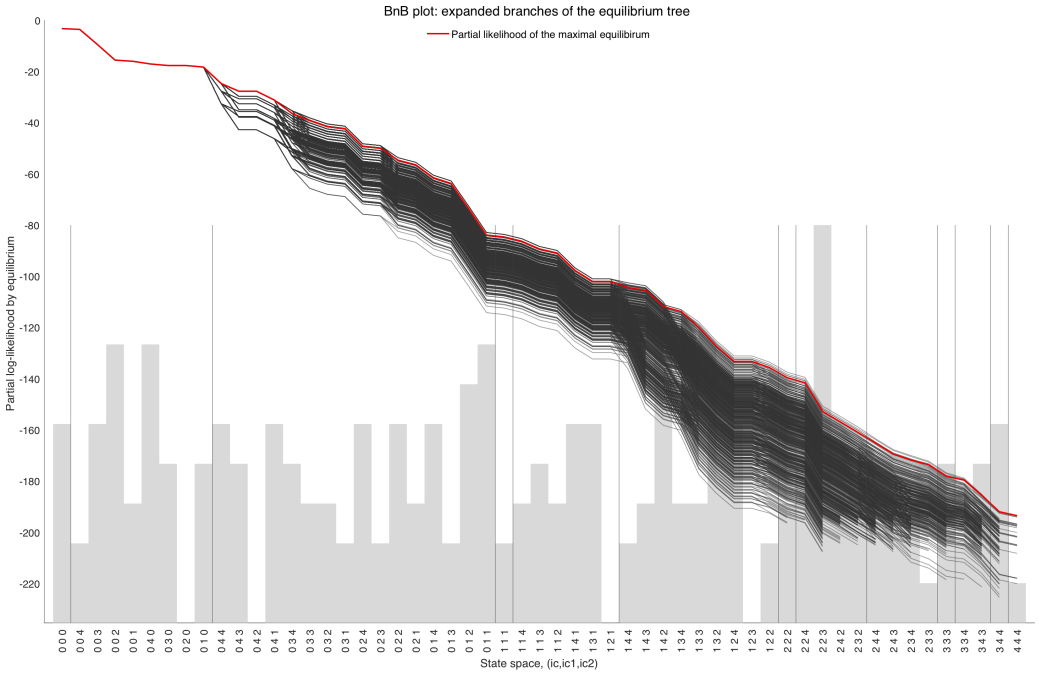


Figure 2: Partial log-likelihood for different MPE

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 2 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.

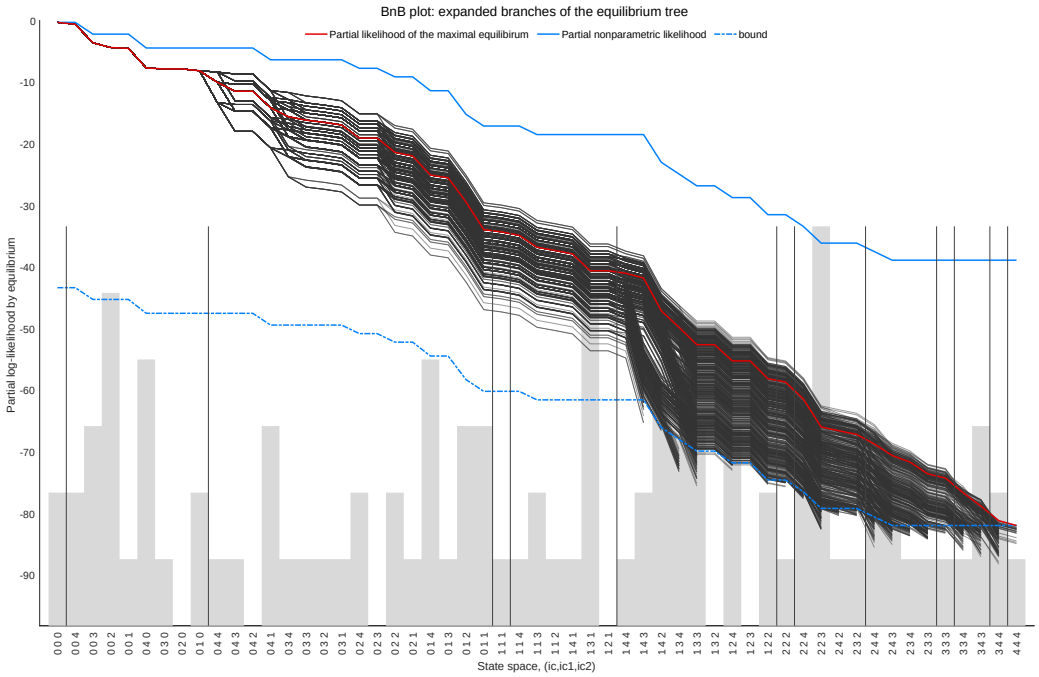


Figure 3: Parial nonparametric log-likelihood (solid blue line) and optimistic computation boundary (dashed blue line)

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 3 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 3 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.

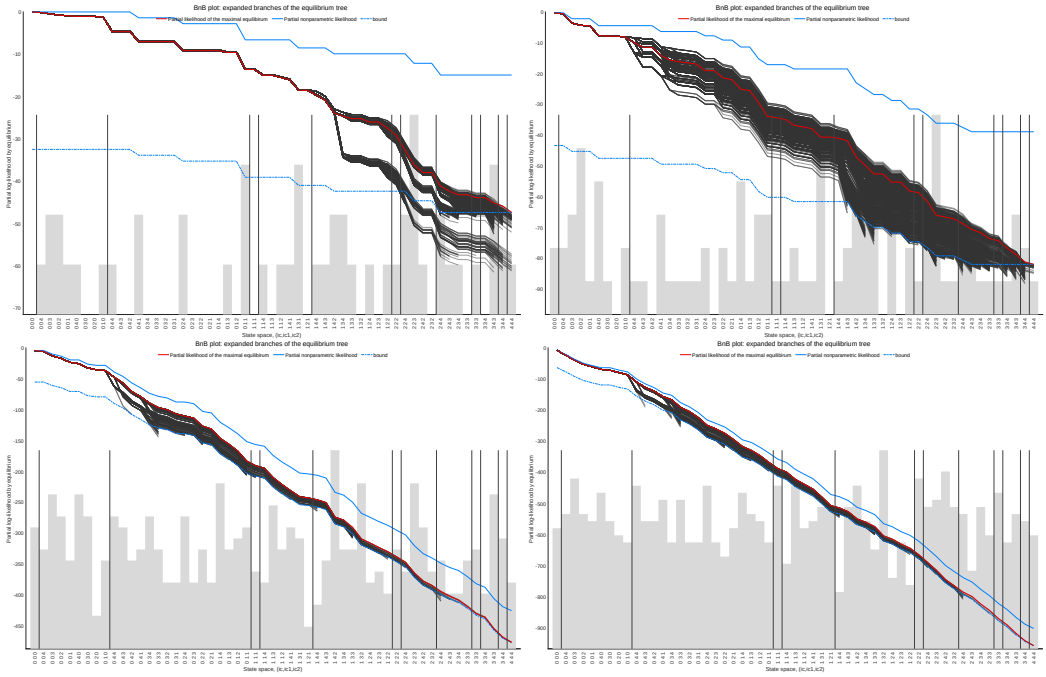


Figure 4: Effect of the sample size on the amount of computations

As sample size increases, the nonparametric loglikelihood line and the pseudo-likelihood line converge to the likelihood line, see Figure 4.

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 Statistical properties of the NRLS estimator

3.2.1 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 $a^{(1)}_{m,t} \in \{0,1\}$ and $a^{(2)}_{m,t} \in \{0,1\}$ 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 $a^{(1)}_{m,t}$ and $a^{(2)}_{m,t}$ 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

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

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 $\mathscr{L}_{M,T}^{(x)}(\theta_x)$, thereby effectively treating the state dynamics as known to us.

3.2.2 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

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

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 equilibrium 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 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

$$\sup_{e\in\mathscr{E}} \sup_{\theta\in\Theta(e)} |\mathscr{L}_{M,T}(\theta, e) - \mathscr{L}_T(\theta, e)| \to^P 0 \tag{18}$$

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 (18), $\mathscr{L}_T(\theta_0,e_0) - \hat{\mathscr{L}}_T(\theta_0,e_0) = o_P(1)$. 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 (18),

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

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

$$s(z_{m,t};\theta,e) = \frac{\partial \ell(z_{m,t};\theta,e)}{\partial \theta}, \quad h(z_{m,t};\theta,e) = \frac{\partial^2 \ell(z_{m,t};\theta,e)}{\partial \theta \partial \theta'}, \tag{19}$$

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 $\sum_{t=1}^{T} E\left[\left\|s(z_{m,t};\theta_0,e_0)\right\|^2\right] < \infty$; $\left\|h(z_{m,t};\theta,e_0)\right\| \leq b(z_{m,t})$ 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 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 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

$$S_{M,T}(\theta,e)=\frac{1}{MT}\sum_{m=1}^{M}\sum_{t=1}^{T}s(z_{m,t_1};\theta,e),\quad H_{M,T}(\theta,e)=\frac{1}{MT}\sum_{m=1}^{M}\sum_{t=1}^{T}h(z_{m,t_1};\theta,e). \tag{20}$$

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 $E[s(z_{m,t};\theta_0,e)|z_{m,t-1},z_{m,t-2},...]=0$ 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:

$$\begin{aligned}M\times\mathscr{LR}_{M,T} &= M\left\{\mathscr{L}_{M,T}\left(\hat{\theta},\hat{e}\right)-\mathscr{L}_{M,T}(\theta_0,\hat{e})\right\} \\ &\simeq MS_{M,T}\left(\hat{\theta},\hat{e}\right)(\hat{\theta}-\theta_0)+\frac{M}{2}(\hat{\theta}-\theta_0)'H_{M,T}\left(\bar{\theta},\hat{e}\right)(\hat{\theta}-\theta_0) \\ &= \frac{M}{2}(\hat{\theta}-\theta_0)'H_{M,T}\left(\bar{\theta},\hat{e}\right)(\hat{\theta}-\theta_0)\to^d\chi^2_{d_\theta},\end{aligned} \tag{21}$$

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

3.2.3 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

$$\hat{e}=\arg\max_{e\in\mathscr{E}}\mathscr{L}_{M,T}\left(\hat{\theta}(e),e\right), \tag{22}$$

where

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

Observe that (22) 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:

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

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

$$\begin{aligned} & M\left\{\mathscr{L}_{M,T}\left(\hat{\theta}(e),e\right)-\mathscr{L}_{M,T}(\theta_0(e),e)\right\} \qquad (23)\\ \simeq\ & MS_{M,T}\left(\hat{\theta}(e),e\right)(\hat{\theta}(e)-\theta_0(e))+\frac{M}{2}(\hat{\theta}(e)-\theta_0(e))'H_{M,T}\left(\bar{\theta}(e),e\right)(\hat{\theta}(e)-\theta_0(e))\\ =\ & \frac{M}{2}(\hat{\theta}(e)-\theta_0(e))'H_{M,T}\left(\bar{\theta}(e),e\right)(\hat{\theta}(e)-\theta_0(e)) \to^d \chi^2_{d_\theta}, \end{aligned}$$

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

$$\sqrt{M}(\mathscr{L}_{M,T}(\theta_0(e),\cdot)-\mathscr{L}_T(\theta_0(e),\cdot)) \to^d Z(\cdot)\in R^{|\mathscr{E}|}\sim N\left(0,\mathrm{Var}\left(\frac{1}{T}\sum_{t=1}^{T}\ell(z_{m,t};\theta,e)\right)\right), \quad (24)$$

where

$$Z(\cdot)\sim N\left(0,\mathrm{Var}\left(\frac{1}{T}\sum_{t=1}^{T}\ell(z_{m,t};\theta,\cdot)\right)\right). \quad (25)$$

Combining (23)–(24),

$$Z_M(\cdot):=\sqrt{M}\left\{\mathscr{L}_{M,T}\left(\hat{\theta}(\cdot),\cdot\right)-\mathscr{L}_T(\theta_0(\cdot),\cdot)\right\} \to^d Z(\cdot), \quad (26)$$

Thus,

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

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 (26), we have

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

where

$$\hat{e}_\infty=\arg\max_{e\in\mathscr{E}}\left\{\sqrt{M}\mathscr{L}_T(\theta_0(e),e)+Z(e)\right\}. \quad (28)$$

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 $\mathrm{Var}\left(\frac{1}{T}\sum_{t=1}^{T}\ell(z_{m,t};\theta,\cdot)\right)$ in (25) by its sample variance. Similarly, the deterministic component $\mathscr{L}_T(\theta_0(e),e)$ is consistently estimated

by $\mathscr{L}_{M,T}\left(\hat{\theta}\left(e\right),e\right)$. Given these estimates, the conditional choice probabilities (CCP's) $\Pr\left(\hat{e}^{*}=e\right)$ 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, $\mathscr{E}_0=\left\{e\in\mathscr{E}:\mathscr{L}_{M,T}\left(\hat{\theta}\left(\hat{e}\right),\hat{e}\right)-\mathscr{L}_{M,T}\left(\hat{\theta}\left(e\right),e\right)\leq K\right\}$ for some cut-off point $K<\infty$.

3.2.4 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:

3.2.5 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

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

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 $\Pr\left(\varepsilon_{t+1}=1\right)=1-\Pr\left(\varepsilon_{t+1}=0\right)>0$. This can be rewritten as

$$x_{t+1}=\mu+x_t+\tilde{\varepsilon}_{t+1}, \tag{29}$$

where $\mu=E\left[\varepsilon_t\right]=\Pr\left(\varepsilon_{t+1}=1\right)$ 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

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

and

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

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 (29), 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.

3.2.6 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 $\mathbf{e}_{0,M}$.

3.2.7 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 $\Pr\left(e_{0,m} = e | x_{m,0} = x\right) = p\left(e|x\right)$, $e \in \mathscr{E}$. That is, we treat $e_{0,m}$ as a latent random variable and we then wish to identify the probabilities $p\left(e_{0,m} = e | x_{m,0}\right)$, $e \in \mathscr{E}$. In this case, the likelihood for market m takes the form

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

where we wish to estimate θ and $\{p\left(e|x\right) : e \in \mathscr{E}, x \in X\}$. 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 ($p\left(e|x_{m,0}\right) > 0$ for $e \in \mathscr{E}_0 \subset \mathscr{E}$ and zero otherwise) to resolve this issue.

3.3 Existing estimators

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 Su and Judd (2012); lin Su (2013); Egesdal, Lai and Su (2015), 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 Aguirregabiria and Mira (2007). In this section we briefly summarize these estimators.

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, Aguirregabiria and Mira (2007) propose an NPL estimator. A NPL fixed point $(\tilde{\theta}, \tilde{\mathbf{P}})$ satisfies the conditions

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

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 Monte Carlo experiments

In this section we present some *very preliminary* Monte Carlo evidence to illustrate the properties of NRLS relative to existing estimation methods. We first consider the case where there is multiple equilibria in theoretical model at the true parameters, but only one of these equilibria is played in the data. In this case we make a head to head comparison between mathematical programming with equilibrium constraints (MPEC) of Su and Judd (2012), a two step pseudo maximum likelihood estimator (PML2step) and the sequential Nested Pseudo Likelihood estimator (NPL) implemented using the NPL algorithm suggested by Aguirregabiria and Mira (2007). We then consider the performance of NRLS when allowing for different equilibria to be played at different markets. For the case where multiple equilibria are played in the data we only investigate the properties for NRLS, since it is less straight forward to extend MPEC and PML2step to this case.

4.1 Data generating process

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 Iskhakov, Rust and Schjerning (2018). The technological progress is indexed by

Table 1: Monte Carlo 1: moderate multiplicity

	MPEC	PML2step	NPL	NRLS
Parameter: k1	1.1090	0.9985	1.0009	0.9815
Bias	0.1090	-0.0015	0.0009	-0.0185
Log-likelihood	-11,102.9	-11,102.9	-11,101.2	-11,092.0
$\|\|\Psi^{\mathbf{P}}(P)-P\|\|$	0.00	0.03	0.01	0.00
$\|\|\Psi^{\mathbf{V}}(v)-v\|\|$	0.00	0.54	0.13	0.00
Runs converged	100.00	100.00	100.00	100.00
CPU time, sec	2.35	0.04	0.23	11.13
K-L divergence	0.01	0.01	0.01	0.00

$\sigma = 0.5$. Number of equilibria in the model: 5. Number of equilibria in the data: 1. True value of parameter k1: 1.0

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.[2] 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 $\beta = exp(-0.05) = 0.9512$.

4.2 One equilibrium played in the data

We conduct three experiments using a sampling scheme where only one equilibrium is played in the data. The results are presented in Table 1, 2 and . For each experiment, we constructed $S = 100$ data sets with 10,000 independent of observations all drawn from the same equilibrium.

Table 1 presents the results of an experiment where there is only 5 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

[2]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 $c_{t+1} > c_t$.

Table 2: Monte Carlo B, run 2: larger multiplicity

	MPEC	PML2step	NPL	NRLS
Parameter: k1	1.0910	0.9948	1.0045	0.9970
MCSD	0.3202	0.0113	0.0094	0.0065
Bias	0.0910	-0.0052	0.0045	-0.0030
Log-likelihood	-6,714.9	-6,714.3	-6,722.5	-6,695.7
$\|\|\Psi^{\mathbf{P}}(P)-P\|\|$	0.00	0.10	0.12	0.00
$\|\|\Psi^{\mathbf{V}}(v)-v\|\|$	0.00	1.22	0.94	0.00
Runs converged	100.00	100.00	5.00	100.00
CPU time, sec	7.72	0.04	0.32	13.79
K-L divergence	0.27	0.01	0.02	0.00

$\sigma = 0.5$. Number of equilibria in the model: 95. Number of equilibria in the data: 1. True value of parameter k1: 1.0

Table 3: Monte Carlo B, run 3: moderate multiplicity, far starts

	MPEC	PML2step	NPL	NRLS
Parameter: k1	6.7685	0.9948	1.0045	0.9970
MCSD	11.3401	0.0113	0.0094	0.0065
Bias	5.7685	-0.0052	0.0045	-0.0030
Log-likelihood	-6,709.4	-6,714.3	-6,722.5	-6,695.7
$\|\|\Psi^{\mathbf{P}}(P)-P\|\|$	0.00	0.10	0.12	0.00
$\|\|\Psi^{\mathbf{V}}(v)-v\|\|$	0.00	1.22	0.94	0.00
Runs converged	30.00	100.00	5.00	100.00
CPU time, sec	10.43	0.06	0.37	18.47
K-L divergence	8.63	0.01	0.02	0.00

$\sigma = 1$. Number of equilibria in the model: 5. Number of equilibria in the data: 1. True value of parameter k_1: 1.0

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.

In Table 2 we present Monte Carlo results where we have reduced the scale of the idiosyncratic investment shocks to $\sigma = 0.5$. This results in more multiplicity (95 equilibria instead of 5), since the taste shocks no longer dampen the strategic incentives to avoid duplicative investments in this anti-coordination game. 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.

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

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 1 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 3. 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.

The current (preliminary) implementation of MPEC is admittedly not ideal since we for example do not implement analytical derivatives of for the involved optimization problems. While this this will affect computational speed for MPEC, analytical derivatives not solve the problem with convergence to local maxima with satisfied constraints. 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.

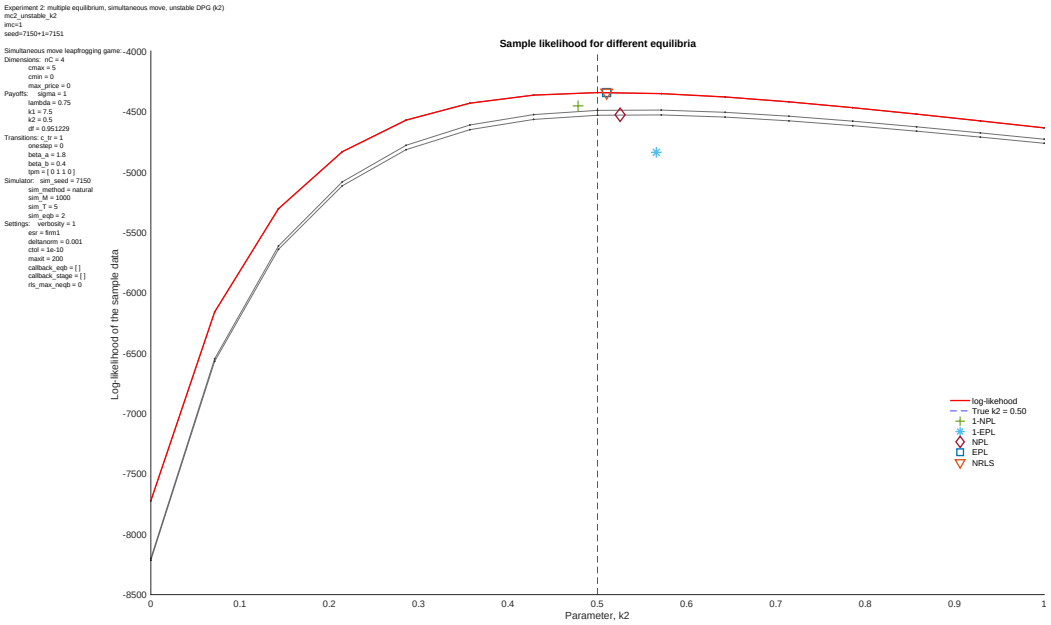
4.3 Graphical evidence

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.

4.4 Multiple equilibria in the data

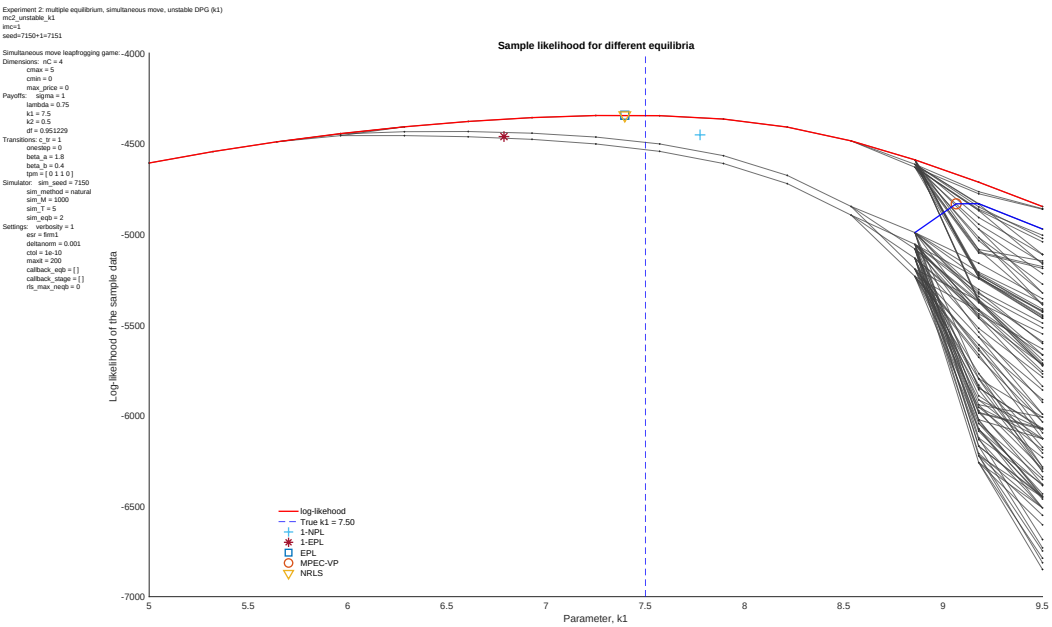
For the case where multiple equilibria are played in the data we only investigate the properties for NRLS, since it is less straight forward to extend MPEC and PML2step to this case. This would require market specific estimates of conditional choice probabilities for all players to initialize each of these algorithms. Allowing for different equilibria to be played at each market, would leading to significant small sample bias for two step estimators and even less ideal starting values for PML2step

Figure 5: Profile likelihood and estimator convergence I



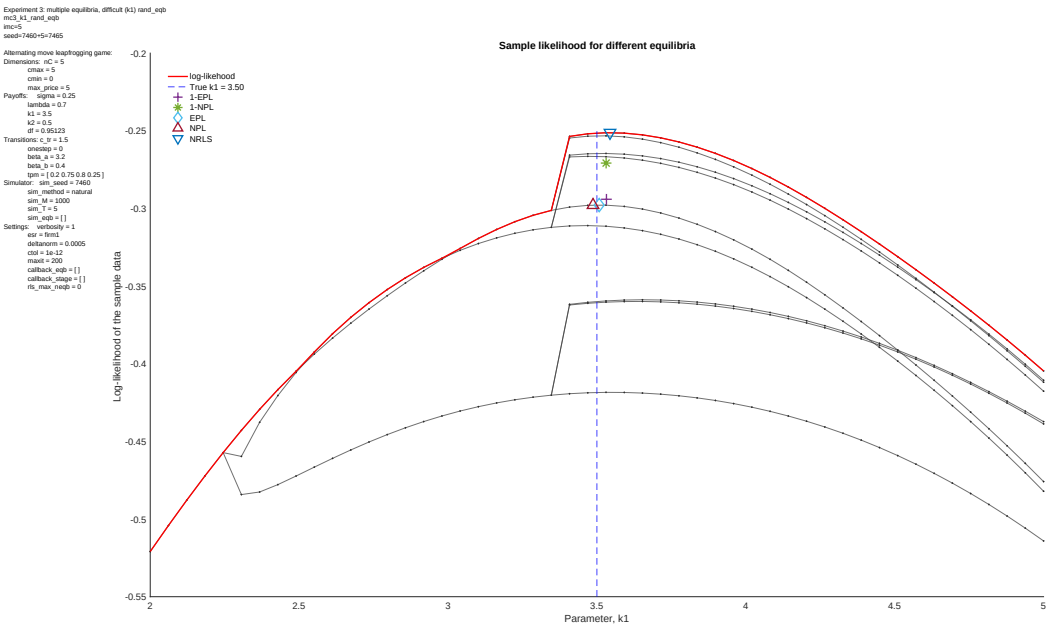
The data is generated from the unstable equilibrium in the model with three equilibria at the true parameter. NPL converged to one of the stable equilibria. Two-step estimators do not fall to any likelihood profile (because they do not impose equilibrium constraints). NRLS and EPL show similar performance.

Figure 6: Profile likelihood and estimator convergence II



The model exhibits a large number of equilibria not far from the true parameter. MPEC converges to a local maximum given by a wrong quilibrium in the parameter interval with high multiplicity. Two-step estimators do not fall to any likelihood profile (because they do not impose equilibrium constraints). NRLS and EPL show similar performance.

Figure 7: Profile likelihood and estimator convergence III



The likelihood function is discontinuous in the neighborhood of the true parameter due to disappearance of several equilibria at a certain parameter value. All estimators experience difficulties: MPEC fails to converge, NPL and EPL converge to an equilibrium which exists for all parameter values, but is not a data generating equilibrium. Two-step estimators do not fall to any likelihood profile (because they do not impose equilibrium constraints). NRLS is the only estimator that robustly converges to the equilibrium played in the data.

and MPEC. For MPEC, the computational complexity of the constrained optimization problem is also significantly amplified since the number of variables and constraints increase by a factor M, since now all choice probabilities and value functions are market specific.

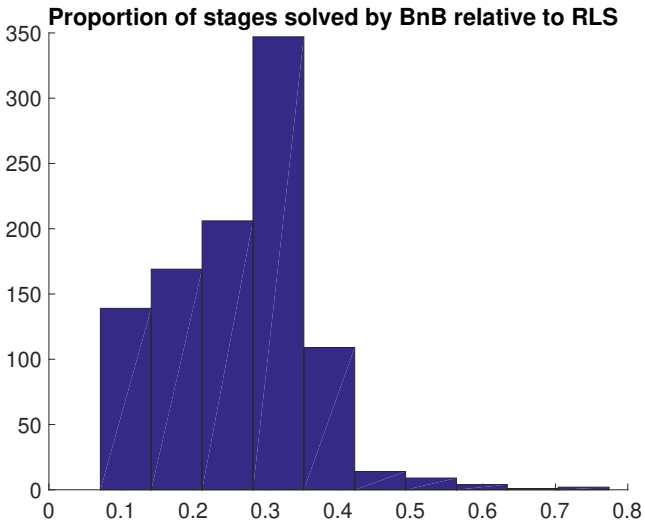
NRLS allows us to relax assumptions for the equilibrium selection rules to allow for different equilibria to be played at different markets without much additional computational cost. To illustrate the performance of NRLS, we have generated $S = 1000$ random samples generated from 3 different equilibria played at $M = 3$ different markets with 100 observations at from each market. In the experiment we increased the investment cost to $K = k_1 = 10$ resulting in 109 MPE in total (at the true parameters).[3]

Figure 8 shows the distribution of computational reduction factor for most basic version of the Branch and Bound algorithm. The average fraction of stages solved by BnB relative to RLS is 0.263 (std=0.09725) and only a tiny fraction of the Monte Carlo runs requires BnB to examining more than half of the stages of the game. The average fraction of MPE computed by BnB relative to full RLS enumeration is only 0.321 (std=0.11635).

Our implementation of NRLS in this experiment is preliminary. It is based on the most basic version of NRLS that ignores that the branch and bound algorithm is nested in the loop that searches over θ. As we mentioned above, this 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. We expect these numbers to be significantly improved

[3]Otherwise the experimental design is similar to the benchmark case studied in Table 1.

Figure 8: Distribution of computational reduction factor



when NRLS is implemented with a recursively updated the bounding function. As the search over the parameter space progress the computational burden of computing the likelihood function will decrease and only require examining the few branches of the solution tree that will potentially result in an increase of the likelihood function.

5 Discussion

6 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 Iskhakov, Rust and Schjerning (2016) 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 (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75, No. 1, 1–53.

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*.

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.

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.

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.

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.

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