

SEMI-PARAMETRIC LIKELIHOOD INFERENCE FOR WORST CASE DYNAMIC DISCRETE GAMES

ZHENGYUAN GAO ¹

AND KEES JAN VAN GARDEREN

This paper considers a dynamic discrete game model with bounded rational agents. Since the appearance of bounded rationality does not lead to a consistent two-step parametric estimator, we propose an alternative two-step inference method. The first step is to control the complexity of the nonparametric approximations and estimate value and policy functions in the worst case fixed point. In the second step, given the fitted value function in the worst case, pseudo true structural parameters are consistently estimated via local Empirical Likelihood with non-differentiable moment constraints. Within the local neighborhood of pseudo true parameters, the estimator is asymptotic normal and achieves the lower bound for the maximum risk. Localization renders the computational algorithm feasible and improves stability and speed.

KEYWORDS: Dynamic Programming, Empirical Likelihood, Local Asymptotic Quadratic family, Linear-quadratic Optimization.

1. INTRODUCTION

1.1. *Relevant Literature*

Dynamic discrete choice models build upon the agents' rational behavior, namely their future expectation and utility maximization. The models exploit the intrinsic evolution structure and capture the endogenous effect of agents' actions. Besides self-expectation, agent decision may depend on the expectation of his opponents' behavior. Concerns about strategic interactions along with the evolution process have motivated a new line of research in dynamic discrete games. Many techniques and estimation methods are introduced to handle various issues in dynamic models, e.g. multiple equilibria, the curse of dimensionality,

Tinbergen Institute and Quantitative Economics Department, University of Amsterdam, z.gao@uva.nl
Tinbergen Institute and Quantitative Economics Department, University of Amsterdam

¹Corresponding author. The paper is based on two chapters of first author's Ph.D thesis. First author would like to thank Peter Boswijk, Wouter den Haan, Kenneth Judd, Paulo Parente, Che-Lin Su, and participants from seminars in Tinbergen Institute, University of Amsterdam for useful comments. First author like to thank Nan Yang for helpful discussions.

and heterogeneity. This paper identifies a robustness issue in dynamic discrete choices and games with boundedly rational agents and then develops a tractable inference approach under a more flexible assumption on consumers' optimal behavior and a more flexible model specification.

Estimation methods for single agent dynamic discrete choice models were initiated by [Wolpin \(1984\)](#), [Pakes \(1986\)](#) and [Rust \(1987\)](#). Estimating the parameters of these structural models requires solving an optimization problem with a nested dynamic programming constraint with an inner and outer loop. The computation is non-trivial for both the likelihood function evaluation in the outer-loop and the fixed point iteration in the inner-loop. The computational complexity grows exponentially when the model allows for strategic interaction. This complexity mainly comes from two sources, the exponential growth of the state space and the existence of multiple equilibria ([Aguirregabiria and Mira, 2009](#)). Although facing these computational difficulties in large scale problems, [Ericson and Pakes \(1995\)](#) set up an oligopolistic competition model under a framework of dynamic discrete games with completely informed participants. Recent work on dynamic discrete games relaxes the complete information assumption and suggests using two step methods to estimate incomplete information models. Incomplete information models are able to explain the heterogeneity in players' actions by introducing random disturbances. These random disturbances, like private information state variables, are assumed to be independently distributed across players such that integration of these random disturbances does not require numerical evaluations over an increasing state space or complete information case. In this way, randomization can break down the curse of dimensionality caused by the multivariate integration in the dynamic programming. [Keane and Wolpin \(1994\)](#) suggest solving the dynamic programming problem by Monte Carlo integration and interpolation. Later, [Rust \(1997\)](#) proves that a certain random Bellman operator with particular structure can break the curse of dimensionality.

In dynamic discrete games, the incomplete belief setting plays important roles in reducing the computational burden and in capturing the heterogenous behavior across agents. The prevailing approaches for solving the incomplete models are slightly different from those in deterministic models. Estimations in complete information models require computing of the full set of equilibria for the candidate parameter values and then select the optimal value. In incomplete information models, the estimation procedure often includes two steps. In first

step, people try to establish objective functions with estimated or approximated components. The second step is to solve either constrained or unconstrained optimization problem with respect to the structural parameters. The objective function for the optimization problems come from the first step. [Pesendorfer and Schmidt-Dengler \(2003\)](#) and [Aguirregabiria and Mira \(2007\)](#) exploit the mapping between the conditional choice probabilities and the choice specific value functions and use nonparametrically estimated probability functions to recover the value function for a specific agent in their first step. [Bajari, Benkard, and Levin \(2007\)](#), [Pakes, Ostrovsky, and Berry \(2008\)](#) and [Bajari, Chernozhukov, Hong, and Nekipelov \(2009\)](#) nonparametrically obtain consistently policy functions in the first step. Then the objective function in the second step can be set up as a pseudo likelihood or a minimum distance function with or without constraints from equilibrium conditions. All these two step approaches concern the optimal value of the objective function in the second step from a set of solutions satisfying the equilibrium constraints.

Apart from applying first step nonparametric estimation for policy functions, nonparametric techniques are used to identify the structural parameters ([Magnac and Thesmar, 2002](#)) and to obtain an approximate solution for the value function ([Rust, Traub, and Wozniakowski, 2002](#)). The infinite dimensional parameter space implicit in the nonparametric setting, induces an infinite space of candidate models. A suitable model for nonparametric inference should balance the bias and the variance of estimators and approximating functions. Therefore, a key issue in nonparametric statistics is how to make an optimal choice in order to minimize e.g. the mean square error (MSE). However, in dynamic discrete games estimation, it seems that it is not advisable to select candidate functions based on the MSE criterion. There are two main reasons. The first reason is that the mis-specification in the second step objective function may be caused by the first step bias. In dynamic discrete games, the second step objective functions are either pseudo likelihood or approximate moment constraints. They are not exact parametric functions or constraints. A researcher conducting inference based on structural parameter estimators in the second step should be concerned about possible bias from the first step estimators. [Fernandez-Villaverde, Rubio-Ramrez, and S. Santos \(2006\)](#) show that the higher order bias from first step dynamic programming solution gives a first order effect on the likelihood function in the second step. The sequence of approximating likelihoods did not converge to the exact likelihood so that the sequence of

approximating policy functions does not converge to the exact one either. Therefore, achieving an exact fit in the first step has the advantage of mitigating the bias in the second step, but at the cost since the exact fit makes the nonparametric model less flexible. The second reason is that prediction and testing can be misleading due to the small sample size. In micro-econometrics, data resources, especially industrial data, are often very limited and it may be unwise to solely rely on pure statistical inferential decisions to judge the correctness of the models. “The statistical mentality that ‘all structural models are rejected, therefore none of them are any good’ does a disservice and contributes to a radicalization of some members of the profession”-Rust (2008). Given the complicated nesting dynamic programming structure, people often prefer to “fit” a given model nonparametrically rather than to obtain a nonparametric function with inferential ability. With an exactly fitting first step approximation it will be easier to obtain a satisfactory result. Due to the importance of the consistency in the first step nonparametric approximation or estimation, there seems little benefit in focussing on the flip side which is to consider the variance of estimated nonparametric functions.

However, a more flexible approach in the first step will allow for more robust inference, also in the second stage. We consider the bounded rationality which is known as *the worst case* of Bellman’s optimality principle in dynamic models, where the numerically approximated solution is a distance ϵ away from the fixed point value function. In Figure 1 from Arifovic (1994), an example based on experimental data (Wellford, 1989) shows that under a complete information design, after a number of periods, a fixed point for profits is approximately achieved. However, under an incomplete information design, the fixed point does not seem to converge. On the other hand, it does not seem to diverge either, but the profit remains within a band¹. Except for a recent theoretical study by Rust, Traub, and Wozniakowski (2002) there is relatively little literature analyzing the dynamic economic model in the worst case.

¹Wellford (1989) shows that in experiment (b) price fluctuates within the region defined by the competitive price and Cournot price. In some periods the market reaches the rational expectations equilibrium level, but does not remain there for the whole duration.

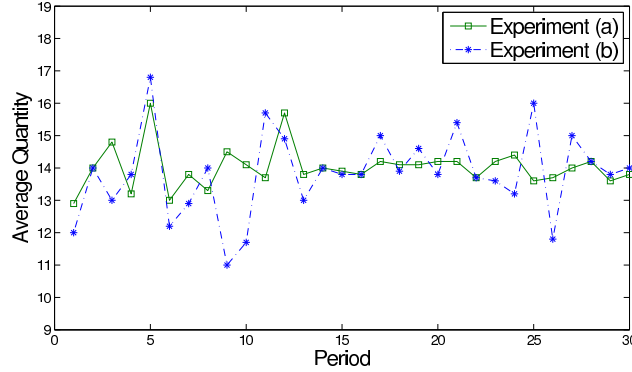


FIGURE 1.— Wellford’s Experiment 2 and 3

The design of Wellford’s experiment is that participants (sellers) make supply decisions under the rational expectation of market clearing price. The price is computed using exogenously given demand schedule and total market supply. Experiment (a) is a complete information model such that sellers realize the exogenously demand schedule. Experiment (b) is an incomplete information model such that sellers receive information only about last-period price and their own profits but no information about total quantity.

1.2. Contributions

The first contribution of this paper is to propose an alternative approach to estimate a dynamic discrete choice model with boundedly rational agents. In games, departures from rationality induce so called ϵ -equilibrium or *robust equilibrium* (Radner, 1981, 1986), which means that each player is satisfied to get within a small distance of his best response function. In other words when no player can gain more than ϵ in his expected payoff by altering his strategy, the strategy profile is said to be an ϵ -equilibrium. In dynamic games, the ϵ -equilibrium is converted to an ϵ -variation in the fixed point condition. Suppose one fixed point is equivalent to one equilibrium model, then the set of models induced by the ϵ -variation complicates the dynamic programming problem. This paper introduces a method to select a single representative model. The selected model represents all those models that satisfy the ϵ -fixed point condition.

The second contribution is to exploit the asymptotic properties of the solution of the dynamic programming problem with boundedly rational agents. The asymptotic behavior is not standard since the solution is inconsistent within a small “ball” but the solution will not get out of this “ball” in probability. This so called “non-trivial consistency” results from the ϵ -fixed point condition.

The third contribution is to develop a localization method as an alternative method for handling the moment constraints with non-smooth population functions, since simulated methods² and other sample smoothing methods³ require differentiable population moment function. This localization method transfers an infeasible problem of structural parameters into a feasible problem of local parameters. Provided the original estimator is consistent and efficient, we prove that the transformation preserves consistency and efficiency. In this sense there is no information loss from the transformation and it is intuitive that this will hold approximately for non-trivial consistency.

The constrained likelihood approach for discrete choice models or dynamic games has also been considered in the Nested Fixed Points (NFXP) approach by [Rust \(1987\)](#), and the Mathematical Programming with Equilibrium Constraints (MPEC) approach by [Su and Judd \(2008\)](#), and the Nested Pseudo Maximum Likelihood (NPML) approach by [Aguirregabiria and Mira \(2007\)](#). None of these approaches, however, consider the worst case fixed point condition with nonparametric probability functions of heterogeneous beliefs.

Essentially, our approach avoids computing the conditional choice probability which causes a huge computational burden⁴. Conditional choice probabilities are used to set up transition densities for the likelihood function⁵ or to derive a closed form policy function⁶ or both. In the

²e.g. [Pakes and Pollard \(1989\)](#).

³e.g. [Kitamura, Tripathi, and Ahn \(e.g. 2004\)](#); [Bajari, Chernozhukov, Hong, and Nekipelov \(e.g. 2009\)](#).

⁴The MPEC proposed by [Su and Judd \(2008\)](#) shares a similar motivation. MPEC can find the optimal solution without a redundant specification of the relation between σ and structural parameters. [Su and Judd \(2008\)](#) show that a constrained optimization problem:

$$\begin{aligned} \max_{(\theta, \sigma)} \mathcal{L}(\theta, \sigma; \mathbf{s}) \\ \text{s.t. } T(\theta, \sigma) = 0 \end{aligned}$$

can be solved by good optimizers without additional efforts of specifying an algorithm for computing $\{\Sigma: \sigma = \Sigma(\theta)\}$. A good solver in an optimization program will implicitly define Σ via $T(\theta, \Sigma(\theta)) = 0$ and implement the augmented likelihood through $\mathcal{L}(\theta, \Sigma(\theta); \mathbf{s})$. When F has unknown functional form, the constraint plays a role as a regularized operator for the inverse problem $G = A^{-1}F$. In addition, [Su and Judd \(2008\)](#) claim that the algorithm in the constrained optimization case will have quadratic convergent rate which is faster than the linear convergent rate of common iterative algorithms. Hence, the constraints in the second step estimation can improve the algorithm and regularize the solution set.

⁵see e.g. [Rust \(1987\)](#); [Aguirregabiria and Mira \(2007\)](#).

⁶see e.g. [Hotz and Miller \(1993\)](#); [Bajari, Benkard, and Levin \(2007\)](#)

first step of our approach, we use the ϵ -approximation to nonparametrically solve the dynamic programming problem. The nonparametric function abandons the necessity of a closed form policy function. Moreover, the bias from the ϵ -approximation will be accumulated in the second step estimation if the second step is based on a parametric likelihood. Therefore, neither the conditional choice probability is necessary, nor the parametric assumption on the private shocks.

We propose a two-step approach to do inference in structural dynamic models with a discrete state space. The approach does not rely on strong parametric assumptions on the specification of the incomplete information and provides flexibility to the potential issue of model mis-specification. The approach also achieves robustness in the presence of ϵ deviation in Bellman's optimality principle. Note that ϵ here is a *deterministic* value determined by prior deliberation. In principle, one could check or optimize ϵ , we do not pursue this in the paper. We also provide a solution to the feasibility issue in the model with non-differentiable moment constraints and solve it using the localization technique developed in this paper.

2. STOCHASTIC DYNAMIC GAMES

We introduce the model considered here with an example of dynamic competition among oligopolistic competitors. The framework with oligopolistic competition dates back to [Ericson and Pakes \(1995\)](#) and has been improved and generalized in the literature. [Doraszelski and Pakes \(2007\)](#) provide a fruitful review. The major feature of this framework is that actions taken in a given period may affect both current profits and the future strategic interactions. The evolution of an industry with heterogeneous firms is modeled via the dynamic programming problem with discrete time and infinite time horizon $t = 1, 2, \dots, \infty$. Dynamic competitions represent in terms of entry, exit, and investment decisions in each period. Unlike the reduced form model, this dynamic structural model is able to exploit the intrinsic evolution process and construct a controllable scheme.

The model includes N firms, denoted $i = 1, \dots, N$. In period t the state variable of firm i is denoted by s_{it} . The state vector among all firms is commonly observed as $\mathbf{s}_t \in \mathcal{S}$, where $\mathcal{S} \in \mathbb{R}^N$ is the entire state space. Depending on the specific application, relevant state variables might include the firms' capacity, market share or investments. Given the current common information \mathbf{s}_t , firm i will choose its strategy $a_{it} \in \mathcal{A}_i$. We assume the firms

choose actions $\mathbf{a}_t = (a_{1t}, \dots, a_{Nt})$ simultaneously in each period. The actions can be a binary choice, e.g. entry and exit decision or a continuous choice, e.g. investment quantities, product prices, etc. Firms face private shocks or opportunities in practice. To econometricians, these shocks $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})$ are unobservable information, therefore ε_{it} is treated as a *random variable*. The distribution associated with ε_{it} is $G_i(\cdot)$ on \mathbb{R} . The different G_i across agents is also known as heterogenous belief in rational expectation models.

Given an action a_i , the state s_i will transfer to the next state s'_i with a certain transition probability $p_i(s'|s; a)$. An action trajectory (a_{i1}, \dots, a_{iT}) will reduce the evolution process from state s_{i1} to s_{iT} to a Markov chain $\mathbf{p}_i = (p_i(s_2|s_1; a_{i1}), \dots, p_i(s_{T+1}|s_T; a_{iT}))$. In the paper, the Markov transition probability is assumed to be homogenous, namely given action a and states s and s' , the transition probability $p_i(s'|s; a)$ is homogenous in time. With this Markovian structure, the firms play stationary Markov strategies in this game. For simplicity, we denote a, s as the current action and state and a', s' as the next period variables. The dependence of $\mathbf{P}(\cdot|\mathbf{s}; \mathbf{a})$ on actions \mathbf{a} are not always necessary. It is obvious that entry/exit decision will affect the firm's next status but it may not so obvious that a short term adjustment of capacity will affect a long term strategic target.

Let $\pi_i(a_i, a_{-i}, \mathbf{s}, \varepsilon_i)$ be current observable profit of firm i with opponents' action a_{-i} . A firm makes its decision to maximize the expected future profit,

$$(2.1) \quad \mathbb{E} \left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} \pi_i(\mathbf{a}_\tau, \mathbf{s}_\tau, \boldsymbol{\varepsilon}_{i\tau}) | \mathbf{s}_t \right],$$

where $\beta \in (0, 1)$ is the discount factor. The primitives of the model include the discount factor β , the transition probability $p(\cdot|\cdot, \mathbf{a})$ ⁷ and the parametric profit functions $\{\pi_i(\cdot, \theta)\}_N$

⁷To assume public informed transition probability seems to be very strange at first glance, but this assumption is just a compromise of relaxing specification of measure $dG(\cdot)$. In boundedly rational case, G is not consistently estimated. A misspecified G does not lead to convincing $p(\cdot|\cdot, \mathbf{a})$ since $p(\cdot|\cdot, \mathbf{a})$ is usually obtained by conditional choice probabilities $F(a|\mathbf{s})$ as [Hotz and Miller \(1993, Lemma 3.1\)](#) where $F(a|\mathbf{s})$ is very sensitive to specification of G . Note that $F(a|\mathbf{s})$ comes from

$$F(a|\mathbf{s}) = \int \mathbb{I}(\sigma = a, \mathbf{s}, \varepsilon) dG(\varepsilon),$$

where the expression of $F(a|\mathbf{s})$ depends on G .

which we introduce in the following.. The parametric expected function is:

$$(2.2) \quad \mathbb{E} \left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} \pi_i(\mathbf{a}_{\tau}, \mathbf{s}_{\tau}, \theta) | \mathbf{s}_t \right],$$

where θ are the structural parameters. The difference are that heterogeneity is not taken into account here and profit function is parameterized as structural parameter θ . The expectation is only taken on transition states. To construct moment constraints of θ in second step, we will use $\pi_i(\mathbf{a}_{\tau}, \mathbf{s}_{\tau}, \theta)$ in (2.2).

In this model, we consider Markov Perfect Equilibria (MPE). MPE guarantee the optimal actions taken by the agents based on the state of the system at that time only. By MPE condition, each firm's action depends on the current state and its current private shock. Let $\boldsymbol{\sigma}(\mathbf{s}, \boldsymbol{\varepsilon}) = (\sigma_1(\mathbf{s}, \varepsilon_1), \dots, \sigma_N(\mathbf{s}, \varepsilon_N))$ be a profile of Markov strategy function or decision rules for N firms such that $\sigma_i(\cdot, \cdot) : \mathcal{S} \times \mathbb{R} \rightarrow \mathcal{A}_i$. The Bellman's equation for the dynamic model (2.1) is

$$(2.3) \quad V_i(\mathbf{s}, \varepsilon_i) = \max_{a_i \in \mathcal{A}_i} \left\{ \pi_i(a_i, a_{-i}, \mathbf{s}, \varepsilon_i) + \beta \int \int V_i(\mathbf{s}', \varepsilon'_i) dP(\mathbf{s}' | \mathbf{s}; a_i, a_{-i}) dG(\varepsilon_i) \right\},$$

where $V_i(\mathbf{s}, \varepsilon_i)$ is the value function. However, the continuously distributed unobservable ε raises serious dimensionality problems since the numerical integral is difficult to approximate the underlying distribution in the high dimension case. By Bellman's principle of optimality, Rust (1987) suggest substitute the decision profile into (2.3) and integrate out the private shocks, then we have:

$$(2.4) \quad V_i(\mathbf{s}; \boldsymbol{\sigma}(\mathbf{s}, \bar{\boldsymbol{\varepsilon}})) = \mathbb{E}_{\varepsilon} \left[\pi_i(\boldsymbol{\sigma}(\mathbf{s}, \boldsymbol{\varepsilon}), \mathbf{s}, \varepsilon_i) + \beta \int V_i(\mathbf{s}'; \boldsymbol{\sigma}(\mathbf{s}', \bar{\boldsymbol{\varepsilon}})) dP(\mathbf{s}' | \mathbf{s}; \boldsymbol{\sigma}(\mathbf{s}, \boldsymbol{\varepsilon})) \middle| \mathbf{s} \right].$$

This is the integrated Bellman's equation. Let $\boldsymbol{\sigma}$ denote $\boldsymbol{\sigma}(\mathbf{s}, \bar{\boldsymbol{\varepsilon}})$. Note that $\bar{\boldsymbol{\varepsilon}}$ is a vector of deterministic value. $V_i(\mathbf{s}; \boldsymbol{\sigma})$ is called the ex ant value function which illustrates that expected profits at the beginning of a period satisfy Bellman optimal principle. $V_i(\mathbf{s}; \boldsymbol{\sigma})$ is not a function of ε , so that solving the Bellman's equation (2.4) does not require integration over the unknown function $V_i(\mathbf{s}, \varepsilon_i)$. Since the policy function σ does not depend on current private information ε , estimation of σ is of little practical interest⁸. Furthermore, we assume following standard conditions:

⁸An alternative way of solving (2.4) is suggested by Bajari, Benkard, and Levin (2007). They estimated

ASSUMPTION 1 *Conditional Independence (CI):* Conditional on \mathcal{S} , the distribution G_t is independently distributed across agents. G_{it} for agent i is independently and identically distributed on time t .

ASSUMPTION 2 *Additive Separability (AS):* Private information appears additively in the profit function. $\pi_{it}(\theta, \mathbf{a}_t, \mathbf{s}, \varepsilon_{ti}) = \pi_{it}(\theta, \mathbf{a}_t, \mathbf{s}) + \varepsilon_{it}(a_{it})$.

In this paper, we use an ϵ -approximating value function \hat{V} to solve the dynamic programming problem and estimate the structural parameter without the step of estimating parametric distributions of private information ε . In section 3, we compare this nonparametric approximation method with the policy function iteration and show that they are asymptotic equivalent when the worst deviation ϵ goes to zero.

3. KERNEL-BASED VALUE FUNCTIONS AND POLICY FUNCTIONS

For the 1st-step estimation, in subsection 3.1 we apply a kernel-based approximation to the value function of Bellman's equation while in subsection 3.2 we show that, in the worst case fixed point condition, the approximation problem can be dually represented by a constrained optimization problem of minimizing approximating function's complexity s.t. an ϵ -deviation fixed point condition. The first method is of theoretical interest, since the method controls the growth of problem's complexity and breaks the curse of dimensionality. The second method achieves our main interest by solving the ϵ -fixed point problem in the Bellman's equation. We prove the solution of this constrained optimization problem is uniformly *nontrivial* consistent to that of the kernel smoothing method in 3.1. Furthermore the exponential bounds of both methods are equivalent, thus the worst case solution breaks the curse of dimensionality. At the end of this section, we derive a policy function iteration algorithm for the worst case Bellman's equation.

the best response function $\sigma(\mathbf{s}, \varepsilon)$ nonparametrically and then use $\hat{\sigma}$ and the integrated Bellman's equation to solve the dynamic programming problem. In this case, the estimated $\hat{\sigma}$ is a function depending on \mathbf{s} and ε . Recovering $\sigma(\mathbf{s}, \varepsilon)$ has practical meanings.

3.1. The Kernel-based Approximation

The estimator derived from the ϵ -deviation fixed point condition does not have a standard asymptotic property, thus it is difficult to assess its convergent result. The purpose of this subsection is to show another nonparametric technique solving dynamic programming problems, the kernel approximation. The asymptotic result of this kernel method is used to compare with that of the worst case estimator in subsection 3.2.

The kernel-based approximation for the dynamic programming is introduced by [Ormoneit and Sen \(2002\)](#) in statistical learning. Their algorithm assigns value function estimates to the states in a sample trajectory $\mathcal{S}_i = \{\mathbf{s}_1, \dots, \mathbf{s}_T\}$ for agent i and updates these estimates by kernel-based averaging. In economics, [Rust \(1997\)](#) gives a local averaging approach to approximate the true Bellman operator and shows that it circumvents the curse of dimensionality in dynamic programming. The track is to impose a normalized structure for the random Bellman operator and then average the value function over the random sample points. Kernel-based averaging is a kind of local averaging with certain functional structures on the dot product space $k : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$, [Ormoneit and Sen \(2002\)](#) Lemma 1 prove that given fixed bandwidth, the kernel-based approximation also has a polynomial growth function. Therefore, the estimator of kernel-based approximation method is an appropriate benchmark for solving dynamic programming in the nonparametric setting.

The recursive relation in (2.4) can be expressed briefly as an Bellman operator Γ mapping from V' to V . In the infinite horizon problem, there is no terminal period from which the backward induction starts to carry out the dynamic programming algorithm in (2.4). Bellman equation can be compactly written as a fixed point condition $V = \Gamma V$ or $V = \Gamma V'$. In practice, a long finite horizon dynamic programming problem can approximately solve the infinite horizon problem. The idea of kernel-based approximation is to use a random operator $\hat{\Gamma}$ to obtain a fixed point condition $V = \hat{\Gamma} V'$ so that given the actions or policy functions, $\hat{\Gamma}$ converges to Γ .

Suppose \mathcal{S}^a is a collection of m_a historical state transitions from s to s' given action a such that $\mathcal{S}_i^a = \{(s_{ij}, s'_{ij}) | j = 1, \dots, m_a\}$. The kernel function $k_{\mathcal{S}^a, b}(\mathbf{s}_i, \mathbf{s}')$ is centered at \mathbf{s}_i ,

$$(3.1) \quad k_{\mathcal{S}^a, b}(\mathbf{s}_i, \mathbf{s}') := \phi\left(\frac{\|\mathbf{s}_i - \mathbf{s}'\|}{b}\right) / \sum_{(\mathbf{s}_u, \mathbf{s}'_u) \in \mathcal{S}^a} \phi\left(\frac{\|\mathbf{s}_u - \mathbf{s}'\|}{b}\right)$$

where ϕ is a “mother kernel” function and b is the bandwidth parameter. ϕ is in the radial basis function (RBF) class including the normalized class used in Rust (1997). The kernel approximation function is:

$$(3.2) \quad \hat{\Gamma}_a(V_i)(\mathbf{s}', \theta_0) = \pi_i(\mathbf{a}, \mathbf{s}, \theta_0) + \sum_{(\mathbf{s}_u, \mathbf{s}'_u) \in \mathcal{S}^a} k_{\mathcal{S}^a, b}(\mathbf{s}, \mathbf{s}') [\beta V_i(\mathbf{s}', \theta_0)].$$

The operator $\hat{\Gamma}_a : \mathcal{B}(\mathcal{S}) \rightarrow \mathcal{B}(\mathcal{S})$ is a random operator based on historic realizations of outcome given action a . Note that the transition probability is not used in (3.2), because we want a single kernel to smooth the product measure involving the transition measure and the private information measure. For simplicity, the profit function from (2.2) is parameterized by true θ_0 so that the model can be calibrated. What we concern of this method is the asymptotic behavior of the approximated operator not its realistic implementation thus we do not bother to consider the unknown θ case.

Given the optimal policy rule σ , the kernel-based function approximates the ex ant value function in (2.4) and $\|\hat{\Gamma}_\sigma - \Gamma\| \rightarrow 0$ in L^∞ . The RBF averages both the random sample points and private shocks in the integrated Bellman equation. The smoothness of the approximation is controlled by the bandwidth choice. Ormoneit and Sen (2002) show that the optimal bandwidth has a shrinkage rate of $O(m_a^{-2/N})$.

THEOREM 3.1 *Given assumption A.1 and A.2, the sequence $\hat{\Gamma}_a V - \Gamma_a V$ uniform converges to zero.*

Equation (3.2) provides a model-free approximation, while the structural parameters are assumed to be known. Since the kernel-based approximation satisfies the fixed point condition asymptotically, one can control the policy function to evaluate the degree of fitting,

$$(3.3) \quad V_i(\mathbf{s}, \theta_0) - [\pi_i(\sigma, \mathbf{s}, \theta_0) + \hat{\Gamma} V_i(\sigma, \mathbf{s}'; \theta_0)].$$

Equation (3.3) is fitting error of kernel-based approximation. The optimal policy function will minimize the fitting error. To obtain the optimal policy function, one need to apply the value iteration update algorithm for a system of linear regression:

$$(3.4) \quad V_i = \pi_i(\mathbf{a}, \mathbf{s}; \theta_0) + \mathcal{T} [K_{\mathcal{S}^a, b}(\beta V_i(\mathbf{s}'; \theta_0))].$$

If $|\mathcal{A}_i| = M$. $K_{S^a,b}$ is a $m_a \times m_a \times M$ tensor with entry $k_{S^a,b}(\mathbf{s}, \mathbf{s}')$ at location $(\mathbf{s}, \mathbf{s}', a)$. \mathcal{T} is an operator on $m_a \times m_a \times M$ tensor and maximizes over its third dimension. V_a is a $m_a \times M$ matrix. The sparsity and complexity of (3.4) is $O(l \times l \times M)$ if we only consider l -nearest neighborhood. It is very essential in models with high dimensional action states \mathcal{A} to choose a fix neighborhood for local averaging. Therefore, kernel approximation method is a good benchmark to study another nonparametric solution of the Bellman's equation. In the following subsection, we propose a method that shares a similar asymptotic behavior with the kernel approximation approach but significantly improves the computational feasibility and solves dynamic programming problems in the bounded rational case.

3.2. The Kernel-based Constrained Optimization

The idea of constrained optimization procedure is to put the tricky part into the constraints and set up the objective function of controllable parameters. We consider the ϵ -variation fixed condition as a single inequality constraint, since a set of MPE satisfies this constraint. Moreover, instead of using kernel function whose parameter is nonlinear, we use a special local basis function to approximate the unknown ex ant value function.

Given the MPE assumption, we approximate the ex ant value function in (3.6) by a local basis function of state variable \mathbf{s} :

$$(3.5) \quad \hat{V}_i(\mathbf{s}) = \rho^T \Phi_i(\mathbf{s}),$$

where Φ is a basis function whose inner product constructs a kernel function such that $\langle \Phi(\mathbf{s}), \Phi(\mathbf{s}') \rangle = k(\mathbf{s}, \mathbf{s}')$, and ρ is the coefficient for Φ . Since kernel (Gram) matrix $K_{ij} := k(\mathbf{s}_i, \mathbf{s}_j)$ is positive definite, Φ is invertible. Once basis function and local sample points chosen, $\Phi(\mathbf{s})$ is determined. Since V is in Banach space, Φ always exists. The basis function approximation can also be found in [Bajari, Benkard, and Levin \(2007\)](#) where they assume profit function and value function are linear in unknown parameters to simplify computation. However, equation (3.5) has different meanings. The variable ρ is not necessary corresponding to underlying parameters, ρ itself is to describe how vibratile and complex the approximation is. Thus we can adjust the fitting of Φ 's polynomial function by tuning ρ . The norm of ρ , $\|\rho\|$ is a regularizer. We put $\|\rho\|^2/2$ in the objective function in order to penalize too complex approximation.

ASSUMPTION 3 *The set of ϵ -equilibrium includes Markov Perfect Equilibrium.*

Now we convert dynamic programming problem to an expression in ϵ -equilibria. ϵ -equilibria implies that agents are indifferent within a set that covers their best response functions such that any V_i

$$V_i(\mathbf{s}; \boldsymbol{\sigma}) - .5\epsilon \leq V_i \leq V_i(\mathbf{s}; \boldsymbol{\sigma}) + .5\epsilon$$

is feasible for equation (2.4). Therefore, by assumption 3 the profit $\pi_i(\boldsymbol{\sigma}(\mathbf{s}, \boldsymbol{\varepsilon}), \mathbf{s}, \varepsilon_i)$ of boundedly rational agent with private information should be one of the acceptable deviation of integrated Bellman's equation such that:

$$(3.6) \quad \left| V_i(\mathbf{s}; \boldsymbol{\sigma}) - \pi_i(\boldsymbol{\sigma}(\mathbf{s}, \boldsymbol{\varepsilon}), \mathbf{s}, \varepsilon_i) + \beta \int V_i(\mathbf{s}'; \boldsymbol{\sigma}) dP(\mathbf{s}'|\mathbf{s}; \mathbf{a}) \right| \leq \epsilon.$$

If agent follows the rational expectation principle, the observable value should equal to $\mathbb{E}\pi_i(\boldsymbol{\sigma}, \mathbf{s}_j, \varepsilon_i)$ so that the fixed point condition exactly holds, but due to private shocks or heterogenous belief, the gain (loss) differs from the expected value. However, this gain (loss) does not affect agent's policy function. The value of ϵ should be set closely to the absolute value of large variation of ε_i if ε is an additive term. Since ϵ -deviation appearance is caused by $\pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i)$.⁹

Therefore, instead of using the equality function (2.4), we establish an inequality where the solution of integrated Bellman's equation (2.4) is included. Equation (3.6) can be interpreted as a so-called ϵ -insensitive loss function devised by Vapnik (1998) such that:

$$(3.7) \quad |\hat{V} - \Gamma\hat{V}|_\epsilon = \max \{0, |\hat{V} - \Gamma\hat{V}| - \epsilon\}.$$

Note that equation (3.7) implies a set of models \hat{V} . We call equation (3.7) ϵ -fixed point condition. The primal problem of minimizing complexity of the approximating value function

⁹Equation (3.6) and equation (2.4) associated with assumption refassAS imply that $\mu \leq 2\epsilon$ where μ_i is the expectation of ε_i . This inequality implicitly states that ϵ -deviation involves the private information.

s.t. the worst case ϵ -deviation of (3.6) is

$$(3.8) \quad \begin{aligned} \min_{\rho, \xi} \quad & \frac{1}{2} \|\rho\|^2 + \frac{C}{|\mathcal{S}_i|} \sum_{j \in \mathcal{S}_i} (\xi_j + \xi_j^*), \\ \text{s.t.} \quad & \rho^T \Phi_i(\mathbf{s}_j) - \left[\pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i) + \beta \sum (\rho^T \Phi_i(\mathbf{s}')) \mathbf{p}(\mathbf{s}' | \mathbf{s}_j, \mathbf{a}) \right] \leq \epsilon + \xi_j, \\ & -\rho^T \Phi_i(\mathbf{s}_j) + \left[\pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i) + \beta \sum (\rho^T \Phi_i(\mathbf{s}')) \mathbf{p}(\mathbf{s}' | \mathbf{s}_j, \mathbf{a}) \right] \leq \epsilon + \xi_j^*, \\ & \xi_j \geq 0, \xi_j^* \geq 0 \quad \forall j \in \mathcal{S}_i. \end{aligned}$$

ξ_j and ξ_j^* are slack variables. Fixed point condition and ϵ -insensitive loss function set up a 2ϵ width “tube” for fitting curves. ξ_j and ξ_j^* play a role as a soft margin for that tube. Minimization in (3.8) captures the main feature of constructing a robust inference procedure. It states that in order to obtain a small risk, we need to control both empirical risk, via ϵ -insensitive loss function, and model complexity, via penaltizer ρ . The parameter C trades off model complexity and curves fitting.

The key idea to solve the optimization problem in (3.8) is to construct a Lagrangian from the objective function and the corresponding constraints, by introducing a dual set of variables. The Lagrangian function has a saddle point with respect to the primal and dual variables at the solution.

$$(3.9) \quad \begin{aligned} L := & \frac{1}{2} \|\rho\|^2 + \frac{C}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} (\xi_j + \xi_j^*) \\ & - \sum_{j \in \mathcal{S}} \alpha_j \left(\epsilon + \xi_j + \rho^T \Phi_i(\mathbf{s}_j) - \left[\pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i) + \beta \sum (\rho^T \Phi_i(\mathbf{s}')) \mathbf{p}(\mathbf{s}' | \mathbf{s}_j, \mathbf{a}) \right] \right) \\ & + \sum_{j \in \mathcal{S}} \alpha_j^* \left(\epsilon + \xi_j^* - \rho^T \Phi_i(\mathbf{s}_j) + \left[\pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i) + \beta \sum (\rho^T \Phi_i(\mathbf{s}')) \mathbf{p}(\mathbf{s}' | \mathbf{s}_j, \mathbf{a}) \right] \right), \end{aligned}$$

where the dual variables have to satisfy positivity constraints $\alpha^*, \alpha \geq 0$. We can simplify the fitting error as

$$(3.10) \quad \rho^T \Phi_i(\mathbf{s}_j) - \left[\pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i) + \beta \sum (\rho^T \Phi_i(\mathbf{s}')) \mathbf{p}(\mathbf{s}' | \mathbf{s}_j, \mathbf{a}) \right] = \rho^T \Psi(\mathbf{s}_j) - \pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i),$$

where $\Psi(\mathbf{s}_j) := \Phi_i(\mathbf{s}_j) - \beta \sum \Phi_i(\mathbf{s}') \mathbf{p}(\mathbf{s}' | \mathbf{s}_j, \mathbf{a})$. The partial derivatives of L with respect to the primal variables (ρ, ξ_j, ξ_j^*) equal to zero for optimality:

$$(3.11) \quad \begin{aligned} \partial L / \partial \rho &= \rho - \sum_{j \in \mathcal{S}_i} (\alpha_j^* - \alpha_j) \Psi(\mathbf{s}_j) = 0 \\ \partial L / \partial \xi_j &= C / |\mathcal{S}| - \alpha_j = 0 \\ \partial L / \partial \xi_j^* &= C / |\mathcal{S}| - \alpha_j^* = 0 \end{aligned}$$

Substituting (3.11) into (3.9) yields the dual optimization problem,

$$\begin{aligned}
(3.12) \quad & \min_{\alpha, \alpha^*} \quad -\frac{1}{2} \sum_{j, t \in \mathcal{S}_i} (\alpha_k^* - \alpha_k)(\alpha_j^* - \alpha_j) \langle \Psi(\mathbf{s}_t), \Psi(\mathbf{s}_j) \rangle \\
& - \epsilon \sum_{j \in \mathcal{S}_i} (\alpha_j^* + \alpha_j) + \sum_{j \in \mathcal{S}_i} \pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i) (\alpha_j^* - \alpha_j), \\
& s.t. \quad 0 \leq \alpha_j^*, \alpha_j \leq C/|\mathcal{S}_i|.
\end{aligned}$$

The dimensionality of the dual problem has been reduced to only $2|\mathcal{S}_i|$, since only α and α^* are the decision variables. The optimization problem (3.12) is a convex linear quadratic programming. Many available packages can solve this standard linear-quadratic programming problem (3.12). Given the numerical values α and α^* , coefficient ρ and value function can be expressed as:

$$\begin{aligned}
(3.13) \quad & \rho = \sum_{j \in \mathcal{S}} (\alpha_j^* - \alpha_j) \Psi(\mathbf{s}_j), \\
& \hat{V}_i(\mathbf{s}_t) = \rho^T \Phi_i(\mathbf{s}_t) = \sum_{j \in \mathcal{S}_i} (\alpha_j^* - \alpha_j) \left[K_{tj} - \beta \sum \mathbf{p}(\mathbf{s}' | \mathbf{s}_t, \mathbf{a}) K_{tj} \right].
\end{aligned}$$

Note that the solution of V does not compute the $O(m^2 \times M)$ complex inversion. ρ is “parameterized” by dual slack variables α , α^* and basis function Φ . Later on, we informally call ρ “parameter”, although it is not a parameter for the underlying model. ρ , however, captures the substantial influence from bounded rationality via ϵ and preserves the shape of the value function that is an approximated solution of integrated Bellman’s equation. Any adjustment of ϵ , the model will be reflected through the variation of ρ . Therefore, ρ plays a similar role as the parameter in parametric model.

Beside the dual representation, it is necessary to extend the result of uniform convergence in Theorem 3.1 to the constrained case. The constraint determines the possible set of functions so that it can be considered as regularized operator for the available functions. We denote it as Υ such that it maps from a inner product space of $\mathcal{K} := \{k|k : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}\}$ into another inner product space $\Omega(k) := \langle \Upsilon k, \Upsilon k \rangle$ where $\Omega(k)$ is the regularization term. The transformation Υ extracts those parts that should be affected by the regularization. Because those interesting “potential candidate” functions should “almost” satisfy the constraint, they are expected to be stable under Υ transforming.

THEOREM 3.2 *If Assumption A.2 holds and a cost function $c(\cdot, \cdot, \cdot) : \mathcal{A} \times \mathcal{S} \times \mathcal{E} \mapsto \mathbb{R}$ satisfies Assumption A.3, for any agent i and actions \mathbf{a} , we have*

$$\lim_{J \rightarrow \infty} \Pr \left\{ \left| \frac{1}{J} \sum_{j=1}^J \left| c(\mathbf{s}_j, \pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_{ij}), \hat{V}) \right|_{\epsilon} - \frac{1}{J} \sum_{j=1}^J c(\mathbf{s}_j, \pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_{ij}), \hat{\Gamma}V) \right| > 4\epsilon \right\} = 0,$$

where \hat{V} is $\sum_{j \in \mathcal{S}_i} (\alpha_j^* - a_j) K_{t_j}$ in equation (3.13) and $\hat{\Gamma}V$ is in equation (3.2). When the cost function $c(\cdot, \cdot, \cdot)$ equals to the fitting error function (3.10) of the fixed point condition. We say the fixed point condition of \hat{V} is nontrivial consistent to that of $\hat{\Gamma}V$.

The *nontrivial* consistency concept is proposed by Vapnik and Chervonenkis (VC) in e.g. Vapnik (1998). This concept is the consistency in the worst case over all tractable candidate functions. In other words, this is another version of law of large numbers which is uniform over all tractable candidate functions. Classic consistency requires to select the best one from all admissible functions for a given sample. While nontrivial consistency requires that the induction principle be consistent even after the “best” functions have been removed. Because these “best” functions are usually the best in some given samples but not the best in other test samples, moreover, these best functions are often much more complicate than the second best functions. Thus VC propose the consistency concept that rules out a function which uniformly does better than all other functions.

This result is also a very useful condition for capacity control. The capacity term is a property of the function class which can be measured by entropy numbers. In the statistical point of view, the worst case concern depresses the growth rate of entropy numbers¹⁰ so that it avoids using too complex functional forms. Since the worst case estimator make the fixed point condition converge to a small neighborhood of the fixed point condition of the common kernel approximation method, it means that a significant simplification of functional forms does not sacrifice too much consistency. Furthermore, the constrained optimization method is also computational better than the kernel self-approximation. To see this, we note that equation (3.13) can be written as $V = \sum_{i,j} \langle \alpha_i, \alpha_j \rangle \Gamma_a k(s_i, s_j)$, as Vapnik (1998) points out that the linear-quadratic optimization over the coefficient $\langle \alpha_i, \alpha_j \rangle$ is more stable than the inverse operator in (3.4).

¹⁰Uniform convergence often requires a exponential bound which has a factor in terms of entropy number. By the theorem, Υ is a scale operator so that the entropy number changes by scaling.

3.3. Iterative Policy Algorithm

The preceding results show that a kernel-based constrained optimization reduced the computation complexity and almost maintain the uniform consistency for the ex ant value function approximation. The construction bases on a fixed policy rule \mathbf{a} . With explicit kernel-based expressions (3.13), we can set up an iteration algorithm to obtain the estimated policy function $\hat{\sigma}$. Due to the ϵ -loss function, the estimated policy function is unnecessary equivalent to that in recursive forward iteration method. The ϵ -tube avoids “over-fitting” the model thus $\Gamma\hat{V}$ will not fluctuate within the tube by $\pm\epsilon$. Thus the action $a = \arg \max_a [\Gamma\hat{V} \pm \epsilon]$ may be different from $a' = \arg \max_a \hat{\Gamma}V$ with an exact approximation $\hat{\Gamma}$. The algorithm is given below:

1. Set the initial policy \mathbf{a}_0 , select a basis function and its corresponding kernel k .
2. Choose a subset \mathcal{S}_a of states space and ensure the transition probability between any two of them are strictly positive.
3. Given the action a_{it} (a_{i0} for the first evaluation) for agent i , calculate the kernel matrix $K := \{K_{kj} = \langle \Phi(s_k), \Phi(s_j) \rangle \text{ for } i \in I\}$ for any $\mathbf{s}_k, \mathbf{s}_j \in \mathcal{S}_a$.
4. Given the profit function π_i evaluated at policy \mathbf{a}_t and kernel matrix K , solve the optimization problem (3.12) for α^* and α^* .
5. Apply (3.13) to obtain ex ant value function $V_i(\mathbf{s})$, and then calculate the one-step policy improvement: $a_{i,t+1} = \arg \max_a \sum_{j \in \mathcal{S}} (\alpha_j^* - \alpha_j) [K_{ij}(a_{i,t}) - \beta \sum \mathbf{p}(\mathbf{s}' | s_i, a_{i,t}) K_{ij}(a_{i,t})]$. Set next period action to $a_{i,t+1}$, update the policy rule and then go back to Step-3.

The algorithm procedure is similar to the inner iteration of the NFXP (Rust, 1987) except that we implement the kernel-based optimization rather than an approximation. Thus one may expect to compare the results of NFXP with (3.12). However, we have to emphasize that it is unfair to use the policy function $a' = \arg \max_a \hat{\Gamma}V$ to judge the correctness of $a = \arg \max_a [\Gamma\hat{V} \pm \epsilon]$. They focus on different aspects. People who uses the exact approximation has a strong confidence on the fixed point condition and prefers to believe the equilibria happen in the steady condition $V^* = \hat{\Gamma}V^*$. People who uses the robust kernel-based optimization may prefer to be less optimistic and consider the potential actions under imperfect situation.

The results are comparable when ϵ is set to zero in (3.8). Once people is optimistic towards Bellman's principle of optimality, he can set ϵ to zero and assume the fixed point condition is exactly hold. The problem reduces to the kernel-based approximation. We give the following theorem for the equivalence of iterative policy algorithm in parametric and semi-parametric settings.

THEOREM 3.3 *When ϵ goes to zero and the basis function of the kernel matrix equals the value function almost surely, the approximating value function obtained by (3.13) is equivalent to the solution V^* of the following system of linear equations:*

$$V^* = \sum_{a_i \in \mathcal{A}} F_i^*(a_i) [\pi^*(a_i) + e_i^*(a_i)] + \beta \sum_{s'} V^*(\theta, s') \mathbf{P}^*(s' | \mathbf{s}),$$

where e^* is the expectation of ϵ conditional on a_i and $F_i^*(a_i)$ is the conditional choice probability. P^* , π^* , e^* are vectors that stack the corresponding state-specific elements. Star * represents the elements associate with an equilibrium conditional choice probability. In addition, the iterative policy algorithms are also equivalent.

Actually, the iterative policy algorithm does not work as good as value function approximation. A small ϵ gives a very fluctuate policy function and then the wiggling policy function induces many local equilibria. A large ϵ may give a poor approximation for value function. Since the iterative policy algorithm depends on approximating value functions, a large ϵ also leads to a disaster. Thus policy iteration is sensitive to the selection of ϵ . To obtain a stable policy function, either one has a very precise knowledge on ϵ or one need to impose constraints about the shape of policy function. [Cai and Judd \(2009\)](#) show that monotonic and concave constraints on policy functions significantly improve the stability of solutions.

4. THE SECOND STEP ESTIMATION

As in the preceding section, θ is treated as a structural parameter for the integrated Bellman equation and ρ exploits the rest structural specification. There are two obstacles for the estimation of θ based on ρ . Firstly, one has no prior information about the form of the distribution of the heterogeneous random variable, nor of the underlying parameter

spaces. Secondly, the pseudo parameter ρ is irrelevant to θ . In this section, we formulate these obstacles as a constrained optimization problem.

The nonparametric smoothing densities are feasible for the first obstacle. Empirical Likelihood (EL [Owen, 1988, 1990, 2001](#)) generates a so-called implied density function based on model constraints. The usual kernel density assigns weights according to the inner product between two input variables in the feature space, in contrast, EL distributes weights according to the imposed constraints. The pseudo parameters of EL's implied density are the Lagrangian multipliers of the model constraints. The input values satisfying the model constraints will be assigned to higher weights while those violating the model constraints will be assigned to lower weights or be penalized to zero if it is an outlier. But directly using EL or GMM method is not applicable, since the moment constraints in our approach are *non-differentiable*. In this section, we develop a local EL estimator that preserves all major properties of EL but can optimize the likelihood subjected to non-differentiable constraints and improve the computational quality.

We handle the second obstacle by constructing a correlation between nonparametric solutions and parametric functions. Given some assumptions, we can build the link between nonparametric and parametric models via the following equation:

$$\rho^T \Phi_i(\mathbf{s}) = \pi_i(\mathbf{a}, \mathbf{s}, \theta_1) + \mu_{it}(a_{it}) + \beta \int \rho^T \Phi_i(\mathbf{s}') dP(\mathbf{s}'|\mathbf{s}, \mathbf{a}),$$

where ρ is solution of [\(3.8\)](#) and μ is the conditional mean of ε_i . The intention of this equation is to replace observations $\pi_i(\cdot, \varepsilon)$ in [\(2.3\)](#) with a parametric function $\pi_i(\theta)$ in order to specify the correspondence of θ and ρ . We consider such $\hat{\theta}$ as the structural estimator.

4.1. *The Semi-parametric Constraints*

In parametric models, the constraint for the second step estimation in dynamic games usually comes from a fixed point condition of the conditional choice probability ([Pesendorfer and Schmidt-Dengler, 2003; Aguirregabiria and Mira, 2007](#)) or an equilibrium condition of the policy function and structural parameters ([Su and Judd, 2008; Bajari, Benkard, and Levin, 2007; Bajari, Chernozhukov, Hong, and Nekipelov, 2009](#)). A nonparametric model does not specify the functional form of mixing distribution G neither the dependence of θ , thus we have to seek for an alternative constraint.

Assumptions 2 and 1 concern on the primitives structure and are the key assumptions about private shocks. We know that in principle the integrated Bellman's equation should be exactly hold, therefore, replace observed $\pi_i(\mathbf{a}, \mathbf{s}, \varepsilon_i)$ with the parametric $\pi_i(\mathbf{s}_t, \mathbf{a}_t, \theta)$ and ε_{it} .

$$(4.1) \quad \mathbb{E}_\varepsilon \left[\pi_i(\boldsymbol{\sigma}(\mathbf{s}, \boldsymbol{\varepsilon}) = \mathbf{a}, \mathbf{s}, \theta) + \varepsilon_i(a_{it}) + \beta \int \hat{V}_i(\mathbf{s}') dP(\mathbf{s}'|\mathbf{s}, \mathbf{a}) \middle| \mathbf{s} \right]$$

where ρ is the solution of (3.8). When the integrated Bellman's equation exactly holds, we use equality instead of inequality,

$$(4.2) \quad \begin{aligned} \rho^T \Phi(\mathbf{s}_t) &= \mathbb{E}_\varepsilon \left[\pi_i(\mathbf{s}_t, \mathbf{a}_t, \theta_0) + \varepsilon_{it}(a_{it}) + \beta \int \rho^T \Phi(\mathbf{s}_t) dP(\mathbf{s}_{t+1}|\mathbf{s}_t, \mathbf{a}_t) \right], \\ &= \pi_i(\mathbf{s}_t, \mathbf{a}_t, \theta_0^*) + \mu_{0,it}^*(a_{it}) + \beta \int \rho^T \Phi(\mathbf{s}_t) dP(\mathbf{s}_{t+1}|\mathbf{s}_t, \mathbf{a}_t). \end{aligned}$$

Equation (4.2) comes from the assumption 2 and 1 and ε -equilibrium property. In this case, our main interest is to estimate the pseudo true θ_0^* that make (4.2) satisfied. The assumption 3 gives the restriction that the solution of integrated Bellman's equation (2.4) is included in the feasible set of the inequality constraint (3.6). In other words, $\rho^T \Phi(\mathbf{s}_t)$, the solution of inequality (3.6), is *an approximated solution* of (2.4). Therefore the pseudo true value is to minimize the Bellman's error in terms of the approximated value function $\rho^T \Phi(\mathbf{s}_t)$ and the parametric profit function $\pi_i(\mathbf{s}_t, \mathbf{a}_t, \theta)$ and $\mu_{it}(a_{it})$. The assumption 1 states ε is independently distributed across t for action a . Rewrite equation (4.2) as:

$$(4.3) \quad \rho^T \Phi(\mathbf{s}_t) - \left[\pi_i(\mathbf{s}_t, \mathbf{a}_t, \theta) + \beta \int \rho^T \Phi(\mathbf{s}_t) dP(\mathbf{s}_{t+1}|\mathbf{s}_t, \mathbf{a}_t) \right] = \int \varepsilon_i(a) dG(\varepsilon(a))$$

ASSUMPTION 4 *Ergodic $G_i(E)$: Γ is a measure-preserving mapping on the measurable space $(\mathcal{E}_i, \mathcal{G}, G_i)$ for all $i \in I$.*

With ergodic $G(\cdot)$ and Fubini theorem, the sample average of equation (4.3) has such property:

$$(4.4) \quad \begin{aligned} \sum_t \left[\rho^T \Phi(\mathbf{s}_t) - \left(\pi_i(\mathbf{s}_t, \mathbf{a}_t, \theta) + \beta \int \rho^T \Phi(\mathbf{s}_t) dP(\mathbf{s}_{t+1}|\mathbf{s}_t, \mathbf{a}_t) \right) \right] / T \\ = \int \left[\sum_t \varepsilon_t(a) / T \right] dG_i \longrightarrow \int \mu_{it}(a) dG_t = \mu_i(a). \end{aligned}$$

Given fixed ϵ , we denote:

$$m_{it}(\theta) = \left\{ \rho^T \Phi(\mathbf{s}_t) - \left(\pi_i(\mathbf{s}_t, \mathbf{a}_t, \theta) + \mu_i(a) + \beta \int \rho^T \Phi(\mathbf{s}_{t+1}) dP(\mathbf{s}_{t+1} | \mathbf{s}_t, \mathbf{a}_t) \right) \right\}.$$

Note that $m_{it}(\theta)$ is conditional on \mathbf{s}_t . The limit of the sample averaging of $m_{it}(\theta)$ over t is an integration w.r.t. G . The empirical distribution of G can be estimated by EL. Note that $\pi_i(\mathbf{s}_t, \mathbf{a}_t, \theta)$ has different functional form of \mathbf{a} . In this case the derivative of moment constraint $m_{it}(\theta)$ w.r.t. structural parameters θ is not well defined. Therefore, we face a problem of non-smooth moment constraint¹¹. In the following subsection, we propose an approach to handle this non-smooth problem and avoid the curse of dimensionality.

4.2. Empirical Likelihood and Local Empirical Likelihood

The constraint (4.4) is used to identify unknown the distribution function G and speed up the computation convergent rate. We apportion the probabilities $g = (g_1, \dots, g_T)$ for the distribution of G . The sum of log-likelihood function of G is expressed in terms of these weights as $\sum_{t=1}^T \log n g_t$. In addition, g should satisfy the common requirement of probabilities such that $g_t \geq 0$ and $\sum_t g_t = 1$. Given action a for firm i , we obtain the following EL criterion:

$$(4.5) \quad \begin{aligned} & \max_{\theta} \sum_t \log n g_{it}, \\ & \text{s.t. } g_{it} \geq 0, \quad \sum_t g_{it} = 1, \\ & \quad \sum_t g_{it} m_{it}(\theta) = 0, \quad \theta \in \Theta_i. \end{aligned}$$

The constraints of the objective function are made up of a convex hull of a family of multinomial distributions $\{g_i\}_T$ and the nonparametric constraint. We may proceed by the method of Lagrange multipliers:

$$(4.6) \quad L' := \sum_t \log n g_{it} - n\lambda \sum_t m_{it}(\theta) + \gamma(\sum_t g_{it} - 1).$$

By Karush Kuhn Tucker conditions (KKT), an explicit expression can be derived by a Lagrange multiplier argument:

$$(4.7) \quad \tilde{g}_{it}(\theta) = \frac{1}{T} \frac{1}{1 + \lambda m_{it}(\rho, \theta)},$$

¹¹The non-differentiability implies that there is no closed form derivative on \mathbb{R} . It is possible that the functional derivative is feasible and tractable.

where λ is found by numerical search. A feasible λ satisfies:

$$(4.8) \quad \frac{1}{T} \sum_{t=1}^T \frac{m_{it}(\theta)}{1 + \lambda m_{it}(\theta)} = 0.$$

Substitute (4.7) into (4.5), we have a minimax criterion function that is the dual representation of problem (4.5):

$$(4.9) \quad \begin{aligned} & - \min_{\theta} \sum_t \log T(1 + \lambda m_{it}(\theta)), \\ & s.t. \frac{1}{T} \sum_{t=1}^T \frac{m_{it}(\theta)}{1 + \lambda m_{it}(\theta)} = 0. \end{aligned}$$

Problem (4.9) is a standard nonlinear optimization problem. The outer-loop of the optimization is to minimize the empirical log-likelihood with respect to θ and the inner-loop is to obtain numerical value of λ .

If the log-likelihood ratio in (4.5) is replaced by entropy $\sum_t n g_{it} \log n g_{it}$, [Kitamura and Stutzer \(1997\)](#) show that the implied density will become

$$(4.10) \quad \tilde{g}_{it,ET}(\theta) = \frac{\exp \lambda m_{it}(\theta)}{\mathbb{E}_t \exp \lambda m_{it}(\theta)},$$

which is similar to the multinomial choice probability. In the parametric case, people usually assume that the unknown private shocks have multinomial distributions. The implied density is far more flexible to the observations and reassign the weights based on the observations on constraint. EL, like GMM and its relative approaches, constructs a divergence criterion to preserve the identification property and improves the robustness and efficiency of the optimization procedure.

However, all GMM and EL estimations are global approaches and people assume the underlying parameters can be globally found. Global optimization procedure is used for problems with a small number of variables, where computing time is not critical, and the possibility of finding the true global solution is very high. The complexity of global optimization methods grows exponentially with the problem sizes. Since the generated parameter set Θ is not a global character, it is unnecessary to implement a global optimization method.

Moreover, the outer-loop optimization could be discontinuous due to the presence of multiple local optima. In contrary to the optimization problem (3.8), the computational difficulty

of EL is to evaluate the outer-loop. The searching direction of θ in the outer-loop is unstable. Let H and s denote the Hessian and gradient function of $\sum_t \log T(1 + \lambda m_{it}(\theta))$. Newton iteration gives

$$(4.11) \quad \theta^{(k+1)} = \theta^{(k)} - H(\theta^{(k)})^{-1} s(\theta^{(k)}).$$

The evaluation of Hessian matrix $H(\theta^{(k)})$ requires the second derivative of log-likelihood function. Note that there is still no closed form derivative of $m_{it}(\theta)$. The numerical derivative is also difficult to implement and time consuming because of nonlinearity. Because we applied the ϵ -loss function in the first step, the θ may perform quite similarly in a small neighborhood of ϵ . The Hessian matrix in this region is so flat that singularity problems may occur.

In addition, for the validity of Newton method, the Taylor expansion for the log-likelihood function should be guaranteed or essentially that the log-likelihood function is smooth enough. Global “smoothness” generally fails in this incomplete belief dynamic model. Therefore, we suggest a localized EL estimation that inherits the properties of EL but overcomes the non-differentiable issue.

We construct LEL estimator which has invariance limited distribution within the local neighborhood and is asymptotic optimal. It does not require global smooth functions and the optimization only depends on local EL’s values. The estimation concerns on local parameters rather than θ , thus it avoids evaluating non-differentiable or nonlinear function of θ . It is more “computational efficient” to evaluate the local EL than to calculate the second derivative of the objective function or the Lagrangian. The word “local” is to indicate that, given θ , the values θ' is very close to the θ so that it is not possible to separate $\tilde{g}(\theta)$ and $\tilde{g}(\theta')$ easily. We will use $\theta + n^{-1/2}\tau$ instead of θ' , where τ is called a local parameter

Suppose the log-likelihood ratios of implied probabilities can locally approximate to a linear-quadratic formula. The problem reduces to a standard linear-quadratic programming problem:

$$(4.12) \quad \begin{aligned} & - \min_{\tau} \sum_t \left[\tau^T S_t - \frac{1}{2} \tau^T M_t \tau \right], \\ & s.t. \frac{1}{T} \sum_{t=1}^T \frac{m_{it}(\theta) - \mu(a)}{1 + \lambda [m_{it}(\theta) - \mu(a)]} = 0. \end{aligned}$$

S and M are elements calculated by LEL.

4.3. LEL Estimator Construction

Partition the space of θ into several grids and select one value from each grid. Let the selected value θ_{it}^* . Running the following procedure for *every grid*.

1. Find an auxiliary estimate θ_{it}^* with its value in θ .
2. Construct a matrix $M_t = \{M_{t,q,p}\}$, $q, p = 1, 2, \dots, l$, where

$$M_{t,q,p} = - \left\{ \Lambda_t[\theta_t^* + \sqrt{T}(u_q + u_p), \theta_t^*] \right. \\ \left. - \Lambda_t[\theta_t^* + \sqrt{T}u_q, \theta_t^*] - \Lambda_t[\theta_t^* + \sqrt{T}u_p, \theta_t^*] \right\}$$

u_1, \dots, u_l are local unit bases in \mathbb{R}^l . M_t is invertible. $\Lambda_t(\theta_1, \theta_2) = \log(\tilde{g}_{it}(\theta_1)/\tilde{g}_{it}(\theta_2))$.

3. Construct a linear term by the linear-quadratic approximation function:

$$u_q^T S_t = \Lambda_t[\theta_t^* + \sqrt{T}u_q, \theta_t^*] + \frac{1}{2}M_{t,q,q}.$$

Since all the values from RHS are known, S_t can be computed as a statistics.

4. Construct a central estimator:

$$\tilde{\theta}_t = \theta_t^* + \sqrt{T}M_t^{-1}S_t,$$

namely $S_t = M_t(\tilde{\theta}_t - \theta_t^*)/\sqrt{T}$. $\tilde{\theta}_t$ appears as θ_n^* with an added correction.

5. Compute local estimator τ via

$$- \min_{\tau} \sum_t \left[\tau^T S_t(\tilde{\theta}_t - \theta_t^*) - \frac{1}{2}\tau^T M_t \tau \right] \\ \text{s.t. } \frac{1}{T} \sum_{t=1}^T \frac{m_{it}(\theta_t^*)}{1 + \lambda m_{it}(\theta_t^*)} = 0,$$

where θ_t^* is fixed value thus $m_{it}(\theta)$ is a numerical vector. The nonlinear constraint is transferred to a linear constraint of λ . The above problem is a linear-quadratic optimization with linear constraints.

6. Obtain the value of $\sum_t \log n \tilde{g}(\tilde{\theta}_t + n^{-1/2}\tau)$ and compare it with the values from other grids. If $\tilde{\theta}_t + n^{-1/2}\tau$ returns the optimal likelihood value, it is selected.

In section 5, we prove that the constructed Hessian matrix M is invertible. The advantage of LEL is that the gradient vectors and Hessian matrices are available without any differentiation operation. LEL also re-scales the likelihood value and then makes it significant to evaluate. The construction of M bases on the property of logarithm function. Matrix M is

very sensitive to the small change. Let $\tilde{g}(\theta_t^* + \sqrt{T}u)/\tilde{g}(\theta_t^*) = \delta$ and $\delta \in (0, 1)$. The logarithm will magnify the small δ exponentially such that $\log \delta \in (-\infty, 0)$. In contrary, the constructed estimator is robust to the peculiar outliers because large δ values will be mitigated by logarithm operation. To see this, suppose the peculiarities $\{x\}_k$ drive the moment constraints on u_i direction unbounded, there is

$$\begin{aligned} & \left\{ \log \frac{\tilde{g}_{\theta_n^* + \delta_n(u_i + u_j)}}{\tilde{g}_{\theta_n^*}}(\{x\}_k) - \log \frac{\tilde{g}_{\theta_n^* + \delta_n u_i}}{\tilde{g}_{\theta_n^*}}(\{x\}_k) - \log \frac{\tilde{g}_{\theta_n^* + \delta_n u_j}}{\tilde{g}_{\theta_n^*}}(\{x\}_k) \right\} \\ &= \left\{ \log \frac{\tilde{g}_{\theta_n^* + \delta_n u_j}}{\tilde{g}_{\theta_n^*}}(\{x\}_k) - \log \frac{\tilde{g}_{\theta_n^*}}{\tilde{g}_{\theta_n^*}}(\{x\}_k) - \log \frac{\tilde{g}_{\theta_n^* + \delta_n u_j}}{\tilde{g}_{\theta_n^*}}(\{x\}_k) \right\} = 0. \end{aligned}$$

The second step in above equation use the property that outliers are independent of specific values of θ . Therefore, the peculiar set $\{x\}_k$ plays no role in the quadratic construction. We conclude that Local EL is a robust estimation. The calculation of M at every fixed θ^* is independent of the numerical second-order derivatives. The evaluation of log-likelihood on θ is easily to compute, thus LEL has an improvement on computation. In section 5, we prove that the linear-quadratic expansion in (4.12) with M_t and S_t obtained by the LEL algorithm converges to the EL log-likelihood function.

THEOREM 4.1 *If the underlying distribution G have finite mean and variance, $-2 \sum_t \log T \tilde{g}(\tilde{\theta})$ converges in distribution to $\chi_{(1)}^2$ as $T \rightarrow \infty$.*

Theorem 4.1 provides an asymptotic justification for tests that accept the value θ at the α level, when $-2 \sum_t \log T \tilde{g}(\tilde{\theta}) < \chi_{(1)}^{2, 1-\alpha}$. If μ is accepted, the function \tilde{g} at given $\tilde{\theta}$ is the implied distribution for this structural parameter. EL confidence region gives an instruction on how to use the implied distribution. If the confidence region accepts μ , it means that μ is a reliable estimated parameter, and its associated implied distribution is suitable for inferential processes. Otherwise, one should be caution to μ and hesitate to apply the empirical distribution of $\varepsilon(a)$ in (4.4) to inference. EL methods for the mean require some modifications to work for other parameters like the variance. The variance of $[m(\theta) - \mu]^2$. If μ is known or it locates in the confidence region, we can construct EL ratio function for the mean of $[m_t(\rho, \theta) - \mu]^2$. The computation is similar to preceding approach.

The implied probability $\tilde{g}(\tilde{\theta})$ can be considered as “parametric family” with $T - 1$ parameters $(\tilde{g}_1, \dots, 1 - \sum_t^{T-1} \tilde{g}_t)$. Implied probability of EL is very flexible to the increasing data

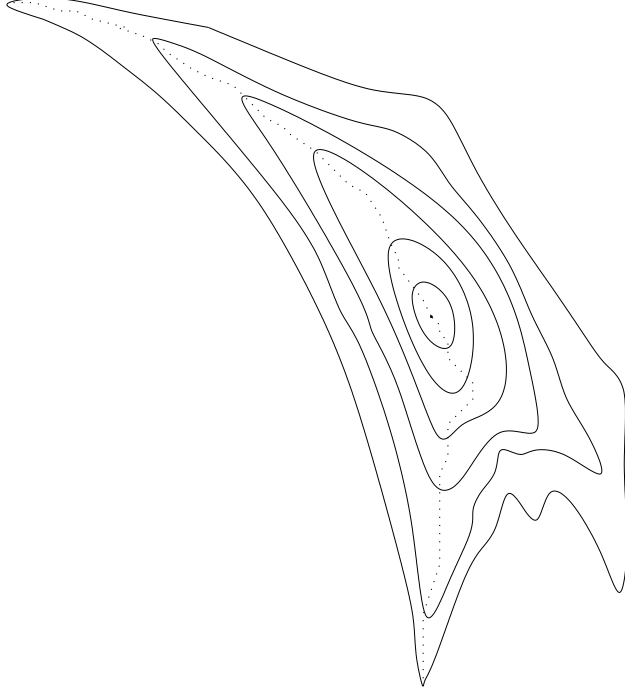


FIGURE 2.— Loglikelihood

The horizontal scale is ten times as large as the vertical scale.

size, since its parameter grows with the sample. Thus for continuous underlying distribution, EL may approach the true parameter values that parametric MLE might not. The growing parameters make EL appear to be very different from parametric likelihood. However, if the underlying distribution is discrete, the EL ratio function is that of a multinomial with only observed values. As t increases, eventually EL reduces to a random data-determined multinomial with an ever-increasing number of parameters.

5. ASYMPTOTIC PROPERTIES

In this section, we only consider parameters θ in order to coincide with other literature on asymptotic theory, moreover ρ is obtained in the first step and plays as fixed values in LEL. Let $M(\theta) = \mathbb{E}m(\theta, x)$, $m_i(\theta) = m(\theta, x_i)$. m_i is $d \times 1$ vector and θ is $k \times 1$ vector with $k \leq d$. $\Lambda_n(\theta_1, \theta_2) = \log(\tilde{g}_{\theta_1}/\tilde{g}_{\theta_2})(x_i)$ and $\Lambda_n(\theta_1) = \log n\tilde{g}_{\theta_1}(x_i)$ where \tilde{g}_θ is the implied probability of $\{m_1(\theta), \dots, m_n(\theta)\}$.

5.1. Consistency without Differentiability Assumption

Huber (1967) first relaxes the differentiable assumption for M -estimator, because some objective functions in this class are not differentiable in \mathbb{R}^N . But differentiability is a common assumption in economics, since most models only depend on smooth expectation of utility functions. Two good benefits for assuming differentiable functions are that people can easily setup the optimization routine of computing Jacobian or Hessian matrices and that the standard *sandwich* estimators' asymptotic results can be applied. Since our moment constraints are not differentiable, the standard procedure of asymptotic proof is not useful here. In this section, we rephrase the consistent proof of EL without assuming any differentiability. In the next section, we derive weakly convergent theorem for LEL under non-differentiable setting.

ASSUMPTION 5 (i) θ_0 is a well-separated point in $M(\theta)$ such that $\sup_{\theta:d(\theta,\theta_0)\geq\epsilon} M(\theta) < M(\theta_0)$. (ii) $m(x, \theta)$ is continuous in θ , $\lim_{\theta' \rightarrow \theta} |m(x, \theta) - m(x, \theta')| = 0$. (iii) $M(\theta)$ exists for all $\theta \in \Theta$ and has a unique zero at $\theta = \theta_0$. $\mathbb{E}m(x, \theta_0)m(x, \theta_0)^T$ is full rank. (iv) There exists a continuous function which is bounded away from zero, such that (1) $\sup_{\theta} |m(x, \theta)|/b(\theta)$ is integrable, (2) $\lim_{\theta \rightarrow \infty} \inf |M(\theta)|/b(\theta) \geq 1$, and (3) $\mathbb{E}[\lim_{\theta \rightarrow \infty} \sup |m(x, \theta) - M(\theta)|/b(\theta)] < 1$.

Assumption 5(i) is a local separable condition. (ii) is used to obtain the continuity of the lagrangian multiplier. (iii) ensures the model is identifiable. (iv) is an envelop assumption; we use it to obtain some dominated convergence results.

THEOREM 5.1 *If assumption 5 holds, then every sequence T_n satisfying*

$$T_n := \arg \sup_{\theta} \sum_{i=1}^n \log n \tilde{g}_i(\theta)$$

will converge to θ_0 almost surely.

The usual proof of EL consistency (Qin and Lawless, 1994) requires twice continuous derivative of $m(x, \theta)$ existed and full ranked. Kitamura and Stutzer (1997); Kitamura, Tripathi, and Ahn (2004) relax the assumptions in Qin and Lawless (1994) and obtain the consistency based on Wald's approach (Wald, 1949). Newey and Smith (2004) assume the differentiability of the Lagrangian rather than $m(x, \theta)$. Schennach (2007) gives another version of consistency proof for the non-differentiable objective function in order to avoid applying

Taylor expansion, however the differentiability of moment constraints is assumed in order to obtain a satisfied approximation for the Lagrangian $\lambda(\theta)$. In this paper, assumption 5 (i)-(iv) are similar with standard L -estimator's conditions in Huber (1981) thus the differentiability assumption is not required.

5.2. Weakly Convergence for LEL Estimator

The classical asymptotic theory in statistics relies heavily on certain linear quadratic approximations to the logarithms of likelihood ratios or the criterion function of M - or Z -estimations. To make the linear quadratic approximations valid, a crucial step is to make sure the function is smooth enough. A global “smoothness” appears to be frail here. Beside smoothness, the navigation of most computation methods need to calculate the Hessian matrix of the linear quadratic approximation. In our problem, the likelihood function is either too flat (ϵ too wide) or too fluctuant (outliers).

The poor-behavior of global method inspires us to think of a local alternative estimation. In this sub-section, we construct a local estimator which has invariance limited distribution within the local neighborhood and is asymptotic optimal. It does not require global smooth functions and the optimization only depends on local EL's values. It is more “computational efficient” to evaluate the local EL than to calculate the second derivative of the objective function or the Lagrangian. Furthermore, it reduces a programming problem with nonlinear constraints to a problem with linear constraints. The transformation significantly mitigates the computational burden and extend the estimation to handle nonstandard constraints.

The word “local” is to indicate that the values θ so close to the point θ_0 that it is not possible to separate $P_{\theta_0,n}$ and $P_{\theta,n}$ easily. The local neighborhoods of a point θ_0 is $\{\theta : |\theta - \theta_0| \leq \delta_n \tau\}$ where τ is called local parameter. The linear-quadratic approximations to the log-likelihood ratios is possibly available in other minimum contrasts estimations, but such construction only leads to asymptotically sufficient estimates when the contrast function can at least locally mimics the log-likelihood function.

If local linear quadratic approximations of log-likelihood ratios existence at θ_0 for every x and moment functions are differentiable, Lemma 1 in Appendix C show that the linear-quadratic expansion will be:

$$(5.1) \quad 2 \sum_{i=1}^n \log \frac{\tilde{g}_{\theta_0 + \delta_n \tau}(x_i)}{\tilde{g}_{\theta_0}} = \tau \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \delta_n \sum_{i=1}^n m_i(\theta_0) \\ - \frac{1}{2} \tau \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \mathbb{E} \frac{\partial m(x, \theta_0)}{\partial \theta} \tau + o_p(1).$$

Asymptotic normality can be deduced from equation (5.1) with additional conditions on the continuity or the boundness of second derivative of the moment restriction functions, e.g. Qin and Lawless (1994), Newey and Smith (2004) or Kitamura, Tripathi, and Ahn (2004). Then we will have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} m_i(\theta_0) \\ \rightsquigarrow \mathcal{N} \left(0, \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \mathbb{E} \frac{\partial m(x, \theta_0)}{\partial \theta} \right).$$

The result is heuristic but not enough in this particular case. In fact, the quadratic term is unknown and cannot be used for the optimum evaluation, furthermore the expansion is simply obtained by the Taylor expansion which is infeasible in non-differentiable case. An analogous condition is assume a probability is differentiable in quadratic mean (DQM). DQM simply says that the root of implied probability $\theta \mapsto \sqrt{\tilde{p}_\theta}$ is differentiable in θ_0 which entails the existence of a vector of measurable functions $S_{\theta_0, n}$ such that

$$(5.2) \quad \int \left[\tilde{g}_{\theta_0 + \tau, i}^{1/2} - \tilde{g}_{\theta_0}^{1/2} - \frac{1}{2} \tau^T S_{\theta_0, n} \tilde{g}_{\theta_0}^{1/2} \right]^2 d\mu = o(\|\tau\|^2),$$

where $\tau \rightarrow 0$ and μ is a dominating measure of \tilde{g}_θ such that $d\tilde{G}_\theta = \tilde{g}_\theta d\mu$. Note that

$$2 \frac{1}{\sqrt{\tilde{g}_\theta}} \frac{\partial}{\partial \theta} \sqrt{\tilde{g}_\theta} = \frac{\partial}{\partial \theta} \log \tilde{g}_\theta.$$

If the Taylor expansion of the root of \tilde{g}_θ is feasible and the remainder term is negligible in $L^2(\mu)$ norm, $S_{\theta_0, n}$ can be considered as the score function of the implied probability \tilde{g}_θ at θ_0 . However, the differentiable assumption on implied probability implies that the first derivative of $m(\theta, x)$ should exist. Therefore, we need to seek for a weaker assumption than DQM, an assumption without assuming any differentiability.

ASSUMPTION 6 For any θ , there are random vector $S_{\theta,n}$ and random matrices $K_{\theta,n}$ such that

$$(5.3) \quad \sum_{i=1}^n \log \frac{\tilde{g}_{\theta+\delta_n\tau_n}(x_i)}{\tilde{g}_\theta} - \left[\tau_n^T S_{\theta,n} - \frac{1}{2} \tau_n^T K_{\theta,n} \tau_n \right]$$

tends to zero in \tilde{G}_θ probability for any bounded sequence $\{\tau_n\}$.

This assumption is to restrict the log-likelihood ratio to a so-called local asymptotic quadratic (LAQ) family. It is weaker than previous assumptions and only assumes that log-likelihood ratios of implied probabilities approximate to a linear-quadratic formula. Lemma 2 in Appendix C gives a formal statement about three assumptions.

PROPOSITION 1 The matrices $K_{\theta,n}$ are almost surely positive definite. If K_θ is a cluster point of $K_{\theta,n}$ is $P_{\theta,n}$ -law, then K_θ is invertible.

The LAQ construction for local likelihood first proposed by Le Cam (1974) to give a smooth local approximation for the log-likelihood ratio function. He suppose that there is no special interest on peculiar features of the likelihood function. The advantage of this construction is that the quadratic term does not depend very much on the value of θ_n^* but the local neighborhood of θ_n^* . Le Cam and Yang (1990) call the priori chosen θ_n^* an auxiliary estimate and show its relation to the Bayesian Gaussian prior. We will use this relationship in the proof.

THEOREM 5.2 Given assumption 5, for every θ that satisfies condition 6, the constructed estimations T_n , S_n and K_n have following properties:

- (i) $K_n^{-1} S_n$ and K_n pointwise converge to $K_{\theta,n}^{-1} S_{\theta,n}$ and $K_{\theta,n}$ respectively in $\tilde{G}_{\theta,n}$ -law.
- (ii) $\delta_n^{-1}(T_n - \theta)$ remains bounded in $\tilde{G}_{\theta,n}$ -law.
- (iii) If equation (5.2) holds, the moment restrictions is just-identified and essentially $K_{\theta,n} \rightsquigarrow K_\theta$ for $\theta \in \Theta$, the sequence of models $\{\tilde{G}_\theta : \theta \in \Theta\}$ is local asymptotic normal (LAN) such that:

For any existence matrices K_θ and random vectors $S_{\theta,n}$, when $n \rightarrow \infty$ and $\tau_n \rightarrow \tau$ there will be

$$\sum_{i=1}^n \log \frac{\tilde{g}_{\theta+\delta_n\tau_n}(x_i)}{\tilde{g}_\theta} = \tau^T S_{\theta,n} - \frac{1}{2} \tau^T K_\theta \tau + o_{\tilde{G}_\theta}(1),$$

where $S_{\theta,n} \rightsquigarrow \mathcal{N}(0, K_\theta)$. Therefore $S_{\theta_0,n} \rightsquigarrow N(0, K_{\theta_0})$ where

$$K_{\theta_0} = \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \mathbb{E} \frac{\partial m(x, \theta_0)}{\partial \theta}.$$

The local asymptotic normality theory (LAN) shows that many statistical models can be approximated by Gaussian experiments¹² after a suitable re-parametrization. In the parametric likelihood framework, when the original models $\theta \mapsto G_\theta$ are smooth in the parameters, namely given a closed set of θ the model G_θ is DQM, the local parameter τ can construct a model $G_{\theta+\tau\delta_n}$ with similar statistical features of normalized random variables. It is possible to asymptotic optimality of LEL estimator can be described by the following theorem.

THEOREM 5.3 *Given assumption 5 and 6, let W be a non-negative bowl shape loss function, if $K_{n,\theta} \xrightarrow{\tilde{G}_\theta} K$, then any sequence of estimates Z_n of τ , one has*

$$\lim_{b \rightarrow \infty} \lim_{c \rightarrow \infty} \liminf_n \sup_{|\tau| \leq c} \mathbb{E} [b \wedge W(Z_n - \tau) | \theta + \delta_n \tau] \geq \mathbb{E}[W(K^{-1/2} \times \mathcal{N}(0, I))]$$

where $\mathcal{N}(0, I)$ is independent of K . The lower bound is achieved by $Z_n = \delta_n^{-1}(T_n - \theta)$. Under the true θ_0 and $W = 1$, $K_\theta = K_{\theta_0}$ induces this lower bound is equivalent to the efficient semi-parametric lower bound.

This is local asymptotic minimax theorem. It bases on the minimax criterion and gives a lower bound for the maximum risk over a small neighborhood of the parameter θ . Because the local EL achieves this lower bound, it is asymptotic optimal estimation.

6. APPLICATIONS AND NUMERICAL RESULTS

We start the section with a simple application to Rust's model of optimal replacement of bus engines (Rust, 1987). This is a single agent dynamic discrete choice model, but it is a useful starting example to illustrates how the semi-parametric algorithm works.

In this model, the maintenance manager i of the bus company has to decide how long to operate a bus before replacing its engine with a new one. The state s_{it} variable is the accumulated miles of the engine at time t . The manager choose whether to replace the engine $a = 1$ or maintain it $a = 0$. When a bus engine is replaced, it is as good as new,

¹²In decision theory, people usually use statistical experiment to refer statistical models.

so the state of the system regenerates to $s_{it} = 0$ when $a_{it} = 1$. The private shock $\varepsilon_{it}(a_{it})$ is assumed to be additively separable. The profit function is given by

$$\pi(a_{it}, s_{it}, \theta_{1i}, \theta_{2i}) = \begin{cases} -\theta_{1i} - c(0, \theta_{2i}) - \varepsilon_{it}(1) & \text{if } a_{it} = 1 \\ -c(s_{it}, \theta_{2i}) - \varepsilon_{it}(0) & \text{if } a_{it} = 0, \end{cases}$$

where $c(\cdot)$ is the cost function of engines' operating and maintenance. The transition probability for s_{it} is:

$$p_i(s_{it+1}|s_{it}; a_{it}) = \begin{cases} g(s_{it+1} - 0) & \text{if } a_{it} = 1 \\ g(s_{it+1} - s_{it}) & \text{if } a_{it} = 0. \end{cases}$$

$g(\cdot)$ is a known probability density function. β in this model is set to 0.98. $g(\cdot) = \theta_{2i} \exp \theta_{2i}(\cdot)$ includes the parameter ρ_{2i} . We use the estimation results in [Rust \(1987\)](#) as the initial values. It is possible to use more realistic distribution such as log-normal distribution with separate parameters for mean and variance, because the explicit solution likelihood is unnecessary. The specification of cost function is given as below:

$$(6.1) \quad \begin{aligned} \text{Quadratic:} & \quad c(s, \theta) = \theta_1 s + \theta_2 s^2, \\ \text{Power:} & \quad c(s, \theta) = \theta_1 s^{\theta_2}, \\ \text{Mixed:} & \quad c(s, \theta) = \theta_1 / (1.1 - s) + \theta_2 s^{1/2}. \end{aligned}$$

The last equation is slightly different from original one where the constant is set to 91. The reason is we scale the state variables to $[0, 1]$ interval rather than discretize them into 90 states. The advantage of scaling is to avoid states in greater numeric ranges dominate those in smaller numeric ranges. Another advantage is to avoid numerical difficulties in kernel calculation. In kernel evaluations, the inner products of basis functions may generate large values which cause problems in the numerical operation.

The kernel matrix evaluated via $\langle \Psi(s_k, s_{-k}), \Psi(s_j, s_{-j}) \rangle$ is shown in [Figure 3](#) for a530875 (monthly data for 1975 GMC model 5308 buses stored in a 117×37 matrix). We use cross validation to select suitable parameters of RBF kernel. Cross validation separates the data to several folds and then carries out parallel grid search in these folds. The kernel matrix has a main diagonal representation with significant sparse pattern on minor diagonal. This matrix captures the phenomena that all the odometers accumulate miles at the beginning

stage but the transitions among states become divergent for large odometer values. Let s' satisfies $\mathbb{E}V(s', a = 1) = \mathbb{E}V(0, a = 0)$, a standard mileage for the engine replacement. When the engine is close to its limit, the manager can either postpone the engine replacement and then ask for a longer running or install a new engine with certain costs. In other words, when s is close to s' , the consideration of effects of Markov process is more important than that at the beginning stage. The feature elements $(s_i, s_{i\pm j})$ for small j s and large i s in the kernel matrix are significant such that these $K(s_i, s_{i\pm j})$ are useful in mapping the sample to a reproducing kernel Hilbert space. Another advantage is the sparse pattern. With large sparse pattern, the computation of inversion and multiplicity are tractable even for every high dimensional matrices.

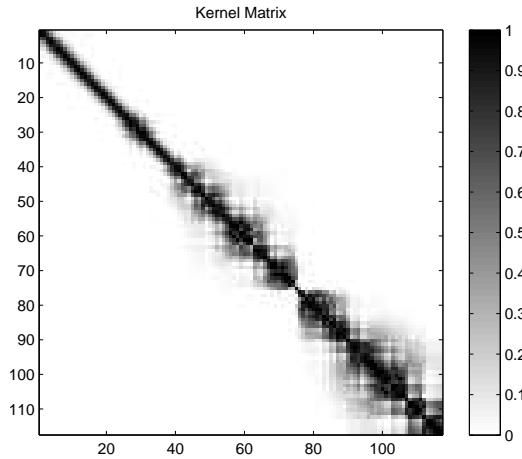


FIGURE 3.— Kernel Matrix: $\Psi(s)^T \Psi(s)$

$\Psi(s_j) = \Phi(s_j) - \beta \sum (\theta^T \Phi(s')) \mathbf{p}(s'|s_j, a)$. The data is from Rust's (a530875) bus data with 4329 observations.

The sensitivity loss ϵ controls the goodness of fit of the kernel approximation and furthermore affects its inferential performance. To analyze the effects of ϵ , we use different ϵ in the power cost function case. We compare approximation and inference ability of this kernel function based on testing and estimating samples. With fixed C , b and θ , Table I and Figure D give the results of goodness of fit and inferential power of kernel functions. The first 50% data is used for estimation while the other is used for testing. The prediction ability is measured in two terms, Mean Squared Error (MSE) and squared correlation coefficient (r^2)

which are $\sum_i^N (\rho^T \Phi(s_i) - \pi_i)/N$ and

$$(6.2) \quad r^2 = \frac{\left(N \sum_i^N \rho^T \Phi(s_i) \pi_i - \sum_i^N \rho^T \Phi(s_i) \sum_i^N \pi_i\right)^2}{\left(N \sum_i^N (\rho^T \Phi(s_i))^2 - (\sum_i^N \rho^T \Phi(s_i))^2\right) \left(N \sum_i^N \pi_i^2 - (\sum_i^N \pi_i)^2\right)}$$

respectively. MSE for testing sample can be considered as the information loss due to the inaccurate prediction. r^2 measures the linear relationship between the approximating value function and the parametric profit function. It is obvious that a too small ϵ make the kernel over-fit the model and thus ask for more information to construct the fitting. The MSE and r^2 are not significant different for $\epsilon = 0.005$ and $\epsilon = 0.05$, the former case nearly utilizes the whole 117 sample vectors while the later case only asks for 63 out of 117 vectors. One can find out there is almost no loss by cutting 50% evaluations of samples. This is important for our nonparametric estimation since the number of basis functions depends on the number of used vectors. Since

$$\rho = \sum_{j \in \mathcal{S}} (\alpha_j^* - \alpha_j) \Psi(s_j, s_{-j}), \quad V(s) = \theta^T \Phi(s)$$

where α and α^* come from equation (3.11). A small number of vectors tells that a small sub-state space of \mathcal{S} is used so as. A complicated basis function leads to untractable kernel computations and optimizations. From $\epsilon = 0.05$ to $\epsilon = 0.5$, the predication accuracy decreases slightly, which, however, is caused by an unpredictable downside shift in the averse direction of the nonparametric trend. From Figure D, we can realize that the curve shape of $\epsilon = 0.5$ is more flexible than that of $\epsilon = 0.05$. In addition, the number of used vector falls to 28, a big gain at sacrifice of an insignificant prediction power. Such benefits does not appear by increasing ϵ to 1. $\epsilon = 1$ is a under-fitting case with poor MSE and r^2 values. Although the basis function number falls to 7, the poor prediction ability indicates it is useless. Figure D shows the values of $(\alpha_i - \alpha_i^*)$ for the sub-state space.

Table II describe the performance of different cost functions. The cost function with power functional form shares similar behavior with that of the quadratic form. Mixed functional form gives unsatisfied outputs for larger structural parameter values. The reason is that power function and quadratic function can be easily approximated via linear basis functions but not for the mixed function. To obtain a satisfied fitting of the mixed function, a bigger number of basis functions is necessary. Therefore the vectors used in mixed function case is significant larger than the other cases. From Figure D, we can find out that the value function

curves do not seem to converge for large parameters. Because large parameter enlarges the fluctuation and thus causes divergent plots.

The cpu time for this computation is less than one second. There are three major reasons. The average of accuracy on predicting the validation sets is the cross validation accuracy. We also test the sensitivity of the choice of loss functions for the fixed point condition. It seems that, a small ϵ will lead to an exact classification policy rule. In this case, each decision is strictly isolated with the others and has very a narrow inference interval. On the other hand, if we allow a flexible ϵ for the fixed point condition, we can use a simple policy rule for classification and have much larger confidence region for the whole data set.

TABLE I

ϵ -LOSS FUNCTION.

ϵ	MSE	r^2	Used vectors
0.005	0.609	0.736	108
0.05	0.614	0.739	63
0.5	0.960	0.690	28
1	3.490	0.322	7

7. CONCLUSION

In this paper, we propose a new two-step inference approach for dynamic discrete choice models with multiple boundedly rational agents. The method does not depend on the parametric assumption of heterogenous beliefs across the agents and relaxes the Bellman's optimal principle condition. In first step, the approach provides a flexible functional form for both estimation and prediction. We also compare its asymptotic behaviors to those of a standard kernel smooth Bellman's operator. In second step, the approach selects optimal structural parameters based on robust semi-parametric likelihood function. Localization transforms the optimization problem involving structural parameters with a nonlinear objective function and nonlinear constraints into an optimization problem involving only local parameters with a linear-quadratic objective function and linear constraints. The local estimator is asymptotically locally optimal. This approach not only preserves standard asymptotic properties but also increases computational speed and avoids peculiar likelihood regions.

APPENDIX A: TECHNICAL ASSUMPTIONS

ASSUMPTION A.1 *There is a λ -measurable weight function ω representing a Radon–Nikodym density of the product measure of transition probability and random shock probability $\omega : (\mathcal{S}, \mathcal{P}, P) \times (\mathcal{E}, \mathcal{G}, G) \mapsto \mathbb{R}$ on the bounded support set \mathcal{S} such that $d(P(s'|s, a) \times G(\varepsilon)) = \omega d(\lambda(s))$.*

ASSUMPTION A.2 *Lipschitz condition for the profit function and the transition probability function. (i). Exists C_π , for any $a \in \mathcal{A}$ and $s, s' \in \mathcal{S}$ such that $|\pi(s, a, \varepsilon) - \pi(s', a, \varepsilon)| \leq C_\pi |s - s'|$. (ii). For any $a \in \mathcal{A}$, $s, s' \in \mathcal{S}$ and $\varepsilon \in \mathcal{E}$, there is $|p(s'|s, a) - p(s'|t, a)| |g(\varepsilon)| \leq C_p(s') ||s - t||$ where $\int C_p^2(s') \omega(ds') < C_p^2 < \infty$.*

ASSUMPTION A.3 (i) *The loss function $c_i(\mathbf{s}_t, \pi_t, \varepsilon_t) = \hat{V}(\mathbf{s}_j) - \beta \int \hat{V}(\mathbf{s}_t) dP(\mathbf{s}_t | \mathbf{s}_j; \mathbf{a}) - \pi_i(\mathbf{a}, \mathbf{s}_t, \varepsilon_{it})$ for each period t is an independent instance. In other words, the joint event $(\mathbf{s}, \varepsilon)_t$ is assumed to be independent from $(\mathbf{s}, \varepsilon)_{t-1}$. (ii) The bound for $||k(\cdot, \cdot)||$ in \mathcal{S} is κ . And $|\hat{V}| = |\rho^T \Phi(\mathbf{s})| < M$ for all $\mathbf{s} \in \mathcal{S}$.*

APPENDIX B: PROOF

PROOF OF THEOREM 3.1: We follow the procedure of Theorem 3.1 and 3.2 in Rust (1997). Instead of using maximal inequality, we apply U -statistics to obtain the normality of $\hat{\Gamma}V - \Gamma V$. With the normality feature, we can obtain an VC-type bound for ε -loss function of $\hat{\Gamma}V$. This is essential to prove Theorem 3.2.

For simplicity, set $d = 1$ for the dimension of state variables. Let any agent i with historic states space \mathcal{S}_i in T periods and set $|\mathcal{S}_i| = J$. First we need to show the approximation error is bounded by the random operators $Z_{a,J}$. In Lemma 3.1 (Rust, 1997) there is

$$\|\hat{\Gamma}V - \Gamma V\| \leq \max_{a \in \mathcal{A}} \|\hat{\Gamma}_a V - \Gamma_a V\| \leq \sum_{a \in \mathcal{A}} \|\hat{\Gamma}_a V - \Gamma_a V\|.$$

So we can restrict attention to $Z_{a,J} = \sqrt{J}[\hat{\Gamma}_a V - \Gamma_a V]$ for given policy rules. In the following proof, we utilize the so-called “kernel track”: Whenever we have a mapping Φ from \mathcal{S} into a dot product space, we obtain a positive definite kernel via

$$(B.1) \quad \sum_{i,j} c_i c_j k(s, s') = \left\langle \sum_i c_i \Phi(s_i), \sum_j c_j \Phi(s_j) \right\rangle = \left\| \sum_i c_i \Phi(s_i) \right\|^2 \geq 0.$$

It gives an equivalent definition of positive definite kernels to functions with the property that there exists a map Φ into a dot product space such that [B.1](#) holds. Therefore, we concentrate on the L^2 norm of the fitting error. By let

$$Z_{a,J}J^{-1/2} = \hat{\Gamma}_a V - \Gamma_a V = \beta \int V(\mathbf{s}|\mathbf{s}')\xi_a(ds'),$$

where

$$\begin{aligned} \xi_a(s') = & \sqrt{J} \left[\sum_i^J k_J(s', s_i)/J - \mathbb{E}k(s', s) \right. \\ & \left. + \mathbb{E}k(s', s) - \omega(d\lambda(s')) \right]. \end{aligned}$$

By assumption [A.1](#), we introduce a \mathcal{S} -measurable function ω such that $\hat{\Gamma}_a V - \Gamma_a V = \beta \int_{\mathcal{S}} V(\mathbf{s}) \cdot \omega(d\lambda(\mathbf{s}))$. Let

$$\|f_J - g_J\|_{2,\omega} = \left\{ \int |f_J(s) - g_J(s)|^2 \omega(s) d\lambda(s) \right\}^{1/2}.$$

In L^2 norm, $\|Z_{a,J}\|_{2,\omega}^2$ can be expressed as $\|Z_{a,J}\|_{2,\omega}^2 = C_{a,J} \times I_J(2)$, where

$$\begin{aligned} (B.2) \quad I_n(2) = & \left\| \frac{1}{J} \sum_i k_J(s', s_i) - \mathbb{E}k(s', s) \right\|_{2,\omega}^2 \\ & + 2 \int \left(\frac{1}{J} \sum_i k_J(s', s_i) - \mathbb{E}k(s', s) \right) \\ & \times \left(\mathbb{E}k(s', s) - \frac{1}{J} \sum_i k(s', s_i) \right) \omega(s) d\lambda(s) \\ & + \|\mathbb{E}k(s', s) - \omega(d\lambda(s))\|_{2,\omega}^2. \end{aligned}$$

The last term is non random and can be analyzed by assumption [A.2](#) (iii). Given assumption [A.1](#) (ii), the second term is readily described by a Central Limited Theorem because kernel function has finite covariance function and it can be written as a sum of i.i.d. random variable by ‘‘symmetrization method’’ or one can simply use Donsker theorem (Lemma 2.3.1 or Theorem 2.5.2 [van der Vaart and Wellner, 1996](#)).

The first term can be expressed as

$$\begin{aligned} (B.3) \quad \left\| \frac{1}{J} \sum_i k_J(s', s_i) - \mathbb{E}k(s', s) \right\|_{2,\omega,g}^2 = & 2 \sum_{1 \leq p < q \leq J} H_J(s_p, s_q) \\ & + \sum_{p=1}^J \int \left[\frac{1}{J} \sum_i k_J(s_p, s_i) \right. \\ & \left. - \mathbb{E}k(s_p, s) \right]^2 \omega(s) d\lambda(s) \end{aligned}$$

The second is a sum of i.i.d. random variables like the second term in (B.2). The first term is

$$H_n(s_p, s_q) = \int_{\mathbb{R}} \left[\frac{1}{J} \sum_i k_J(s_p, s_i) - \mathbb{E}k(s, s_1) \right] \times \left[\frac{1}{J} \sum_i^n k_J(s_q, s_i) - \mathbb{E}k(s, s_1) \right] \omega(u) du.$$

Theorem 1 in Hall (1984) shows that $\sum_{1 \leq p < q \leq J} H_J(s_p, s_q)$ is a centered U -statistic with a variable kernel function $\mathbb{E}(H_n(s_1, s_2)) = 0$. Therefore $n^{-1} \sum_{1 \leq p < q \leq J} H_J(s_p, s_q)$ is asymptotically normally distributed with zero mean and variance $\mathbb{E}H_J^2(s_p, s_q)/2$.

Finally, $I_J(2)$ can be written as the sum of a deterministic term and of three random terms. By CLT and U -statistics property, we have a joint asymptotic normality for $I_J(2)$ so as $Z_{a,J}$. Note that the fact that the deterministic term $\|\mathbb{E}k(s', s) - \omega(d\lambda(s))\|_{2,\omega}^2$ in equation (B.2) drops out results from our regularity Assumption A.2 (ii). In fact, one can “tune” the assumption and allow for ϵ error and in addition, one can also let one of the normal random variable dominate the others. Lemma A.2.2 in van der Vaart and Wellner (1996) states that if $\hat{\Gamma}_a V$ is Lipschitz function and $|\hat{\Gamma}_a V| < M$, there is

$$\Pr \left\{ \hat{\Gamma}_a V - \Gamma_a V > \epsilon \right\} \leq \exp \left(-\frac{\epsilon^2}{2M^2} \right).$$

By Assumption A.2, it is not difficult to have a constant M such that $|\hat{\Gamma}_a V| < M$. Thus the above inequality holds. In the proof of Theorem 3.2, we need above exponential bound. *Q.E.D.*

PROOF OF THEOREM 3.2: In the proof, for simplicity, we change the notation $\pi_i(\mathbf{a}_j, \mathbf{s}_j, \varepsilon_{ij})$ to π_j and set $d = 1$ for the dimension of state variables. We call the fitting error function (3.10)

$$\hat{V}(\mathbf{s}_j) - \beta \int \hat{V}(\mathbf{s}_t) dP(\mathbf{s}_t | \mathbf{s}_j; \mathbf{a}) - \pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i) = \sum_{j \in \mathcal{S}_i} (\alpha_j^* - a_j) \langle \Psi(\mathbf{s}_j), \Psi(\mathbf{s}_j) \rangle - \pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_i)$$

a loss function $c(\mathbf{s}_j, \pi_j, \hat{V})$. Changing m -th observation generates a new set $\mathcal{S}^m := (\mathcal{S} \setminus \{s_m\}) \cup \{s_{new}\}$. A new approximation function in \mathcal{S}^m denoted as $\hat{V}_{\mathcal{S}^m}$. Let any agent i with historic states space \mathcal{S}_i in T periods and set $|\mathcal{S}_i| = J$.

The purpose of the prove is to show the fixed point condition with regularized Bellman operator $\Gamma\hat{V}$ converges uniformly to a neighborhood of the fixed point condition with $\hat{\Gamma}V$. A straightforward idea is to check the validity of the following argument:

$$\left| \sum (\hat{V}_j - \Gamma\hat{V}_j) / J - \mathbb{E} (V - \hat{\Gamma}V) \right|_\epsilon \rightarrow 0.$$

If empirical process theorem (e.g. Glivenko-Cantelli) can be applied to LHS of the argument, then it would be simple. However, there are two difficulties of using empirical process theorem to prove the uniform convergence here. The first problem is the ϵ -loss function (ϵ -fixed point constraints) make $\hat{V}_j - \Gamma\hat{V}_j$ not exactly zero and the second is the effect from regularization. Fortunately, Theorem 12 from [Bousquet and Elisseeff \(2002\)](#) (BE theorem) states that if changing observations in a certain loss function only has a bounded variation then there is a Hoeffding type bound¹³ for this loss function. In our setting, the theorem can be stated in this way:

If Assumption [A.3](#) (i) holds and

$$(B.4) \quad \left| c(\mathbf{s}, \pi, \hat{V}) - c(\mathbf{s}, \pi, \hat{V}_{S^m}) \right| \leq \varsigma$$

for all (s, π) , then there is

$$(B.5) \quad \Pr \left\{ \left| \frac{1}{T} \sum_{t=1}^T \left| c(\mathbf{s}_t, \pi_{it}, \hat{V}) \right|_\epsilon - \mathbb{E} \left(c(\mathbf{s}, \pi_i, \hat{V}) \right) \right| > \epsilon + \varsigma \right\} \\ \leq 2 \exp \left(- \frac{J\epsilon^2}{2(J\varsigma + M)^2} \right).$$

When $J \rightarrow \infty$, the exponential bound goes to zero. While the empirical process of $c(\mathbf{s}_t, \pi_{it}, \hat{\Gamma}V)$ has a better bound such that

$$(B.6) \quad \Pr \left\{ \left| \frac{1}{T} \sum_{t=1}^T c(\mathbf{s}_t, \pi_{it}, \hat{\Gamma}V) - \mathbb{E} \left(c(\mathbf{s}, \pi_i, \hat{\Gamma}V) \right) \right| > \epsilon \right\} \\ \leq N(\epsilon) \exp \left(- \frac{2J\epsilon^2}{M^2} \right),$$

where $N(\epsilon)$ is the covering number for the RBF kernel smoothing $\hat{\Gamma}V$. When J increases $\exp(-2J\epsilon^2/M^2)$ will decrease to zero at a faster rate. Therefore, we can impose equation

¹³The bound slightly differs from original Hoeffding's bound but has similar convergence rate.

(B.6) into (B.5) by the probability rule $\Pr(A) + \Pr(B) \geq \Pr(A \cup B)$ and we don't need to change the bound. The new inequality becomes:

$$\begin{aligned}
& \Pr \left\{ \left| \frac{1}{J} \sum_{j=1}^J |c(\mathbf{s}_j, \pi_j, \hat{V})|_\epsilon - \mathbb{E} \left(c(\mathbf{s}, \pi, \hat{V}) \right) \right| > \epsilon + \varsigma \right\} \\
& \geq \Pr \left\{ \left| \frac{1}{J} \sum_{j=1}^J |c(\mathbf{s}_j, \pi_j, \hat{V})|_\epsilon - \frac{1}{J} \sum_{j=1}^J c(\mathbf{s}_j, \pi_j, \hat{\Gamma}V) \right| - \right. \\
\text{(B.7)} \quad & \left. \left| \mathbb{E} \left(c(\mathbf{s}, \pi, \hat{V}) \right) - \mathbb{E} \left(c(\mathbf{s}, \pi, \hat{\Gamma}V) \right) \right| > 2\epsilon + \xi \right\}
\end{aligned}$$

From RHS of equation (B.7), we can conclude

$$\begin{aligned}
\text{(B.8)} \quad \Pr \left\{ \left| \frac{1}{J} \sum_{j=1}^J |c(\mathbf{s}_j, \pi_j, \hat{V})|_\epsilon - \frac{1}{J} \sum_{j=1}^J c(\mathbf{s}_j, \pi_j, \hat{\Gamma}V) \right| > 4\epsilon + 2\varsigma \right\} \\
\leq 2 \exp \left(-\frac{J\epsilon^2}{2(J\varsigma + M)^2} \right).
\end{aligned}$$

If $\varsigma \rightarrow 0$ as $J \rightarrow \infty$ then the proof ends.

Therefore, now we need to find the bound in (B.4). In Example 1 and Appendix C of [Bousquet and Elisseeff \(2002\)](#), they give an instruction how to use the subgradient of a constructed convex function to derive a stable bound for a support vector regression model. We follow this idea in the following proof. Let's define the specific form of the regularized function:

$$\begin{aligned}
\Upsilon \hat{\Gamma}V &= \frac{1}{J} \sum_{j=1}^J |c(\mathbf{s}_j, \pi_j, \hat{V})|_\epsilon + \frac{1}{2} \|\hat{V}\|^2 \\
&= \frac{1}{J} \sum_{j=1}^J \left| \rho^T \Psi(\mathbf{s}_j) - \pi_j \right|_\epsilon + \frac{1}{2} \|\rho^T \Phi(\mathbf{s})\|^2 \\
&= \frac{1}{J} \sum_{j=1}^J \left| \sum_{i \in \mathcal{S}_i} (\alpha_i^* - \alpha_i) \langle \Psi(\mathbf{s}_i), \Psi(\mathbf{s}_j) \rangle - \pi_i(\mathbf{a}, \mathbf{s}_j, \varepsilon_{ij}) \right|_\epsilon \\
&\quad + \frac{1}{2} \left\| \sum_{i \in \mathcal{S}_i} (\alpha_i^* - \alpha_i) K_{ij} \right\|^2,
\end{aligned}$$

where $\Psi(\mathbf{s}) = \Phi(\mathbf{s}) - \beta \sum_{i=1}^n \Phi(\mathbf{s}') \mathbf{p}(\mathbf{s}'|\mathbf{s}, a)$. The functional derivative of $|c(\mathbf{s}_j, \pi_j, \hat{V})|_\epsilon$ w.r.t.

V is

$$\begin{aligned} c'(\mathbf{s}_j, \pi_j, \hat{V}) &= \partial \left| \rho^T \Psi(\mathbf{s}_j) - \pi_j \right|_\epsilon \\ &= \begin{cases} 0 & \text{if } \left| \rho^T \Psi(\mathbf{s}_j) - \pi_j \right| \leq \epsilon \\ \frac{(\pi_j - \rho^T \Psi(\mathbf{s}_j))}{\left| \rho^T \Psi(\mathbf{s}_j) - \pi_j \right|} & \text{otherwise.} \end{cases} \end{aligned}$$

It is obvious that $|c'(\mathbf{s}_j, \pi_j, \hat{V})| \leq 1$. The derivative of the regularized function is

$$(B.9) \quad \partial \Upsilon \Gamma \hat{V}(\mathcal{S}) = \frac{1}{J} \sum_{j=1} c'(\mathbf{s}_j, \pi_j, \hat{V}) + \rho^T \Phi(\mathbf{s}) = 0,$$

$$(B.10) \quad \partial \Upsilon \Gamma \hat{V}(\mathcal{S}^m) = \frac{1}{J} \sum_{j=1} c'(\mathbf{s}_j, \pi_j, \hat{V}_{\mathcal{S}^m}) + \rho_{\mathcal{S}^m}^T \Phi(\mathbf{s}_m) = 0.$$

Next, we construct an auxiliary convex function:

$$\begin{aligned} A(f) &= \left\langle \partial \Upsilon \Gamma \hat{V}(\mathcal{S}) - \partial \Upsilon \Gamma \hat{V}(\mathcal{S}^m), f - \rho_{\mathcal{S}^m}^T \Phi(\mathbf{s}) \right\rangle \\ &\quad + \frac{1}{2} \left\| f - \rho_{\mathcal{S}^m}^T \Phi(\mathbf{s}_m) \right\|^2. \end{aligned}$$

It is obvious that $A(\rho_{\mathcal{S}^m}^T \Phi(\mathbf{s}_m)) = 0$. Furthermore, the functional derivative of $A(f)$ w.r.t. f is

$$\begin{aligned} \partial A(f) &= \partial \Upsilon \Gamma \hat{V}(\mathcal{S}) - \partial \Upsilon \Gamma \hat{V}(\mathcal{S}^m) + (f - \rho_{\mathcal{S}^m}^T \Phi(\mathbf{s}_m)) \\ &= \partial \Upsilon \Gamma \hat{V}(\mathcal{S}) + f. \end{aligned}$$

The second equation comes from FOC (B.10). By FOC (B.9), the minimum of $A(f)$ is achieved at $\rho^T \Phi(\mathbf{s})$, because $\partial A(\rho^T \Phi(\mathbf{s}_m)) = 0$ and $A(f)$ is a convex function. Thus $A(f) \leq 0$.

$$\begin{aligned} &J \times \left\langle \partial \Upsilon \Gamma \hat{V}(\mathcal{S}) - \partial \Upsilon \Gamma \hat{V}(\mathcal{S}^m), \rho^T \Phi(\mathbf{s}) - \rho_{\mathcal{S}^m}^T \Phi(\mathbf{s}_m) \right\rangle \\ &= \sum_{j \neq m} \left[c'(\mathbf{s}_j, \pi_j, \hat{V}) - c'(\mathbf{s}_j, \pi_j, \hat{V}_{\mathcal{S}^m}) \right] \times \left[\rho^T \Psi(\mathbf{s}_j) - \rho_{\mathcal{S}^m}^T \Psi(\mathbf{s}_j) \right] \\ (B.11) \quad &+ c'(\mathbf{s}_m, \pi_m, \hat{V}) \times \rho^T \Psi(\mathbf{s}_m) - c'(\mathbf{s}_m, \pi_m, \hat{V}_{\mathcal{S}^m}) \times \rho_{\mathcal{S}^m}^T \Psi(\mathbf{s}_m) \end{aligned}$$

If $\rho^T \Psi(\mathbf{s}_j) > \rho_{\mathcal{S}^m}^T \Psi(\mathbf{s}_j)$ then $c'(\mathbf{s}_j, \pi_j, \hat{V}) > c'(\mathbf{s}_j, \pi_j, \hat{V}_{\mathcal{S}^m})$ by the convex property of ϵ -loss function. For a convex function $f(a) + (b - a)f'(a) \leq f(b)$ and $f(b) + (a - b)f'(b) \leq f(a)$, then one can obtain $(f'(a) - f'(b))(a - b) \geq 0$. Then the first term of equation (B.11) is

positive. Therefore,

$$\begin{aligned} J \times \left\{ 0 - \frac{1}{2} \|\hat{V} - \hat{V}_{S^m}\|^2 \right\} &\geq J \times \left\{ A(\hat{V}) - \frac{1}{2} \|\hat{V} - \hat{V}_{S^m}\|^2 \right\} \\ &\geq c'(\mathbf{s}_m, \pi_m, \hat{V}) \times \rho^T \Psi(\mathbf{s}_m) - c'(\mathbf{s}_m, \pi_m, \hat{V}_{S^m}) \times \rho_{S^m}^T \Psi(\mathbf{s}_m). \end{aligned}$$

By Assumption A.3 (ii) and $|c'(\mathbf{s}, \pi, \hat{V})| \leq 1$,

$$\begin{aligned} \frac{J}{2} \|\hat{V} - \hat{V}_{S^m}\|^2 &\leq c'(\mathbf{s}_m, \pi_m, \hat{V}) \times \rho^T \Psi(\mathbf{s}_m) \\ \text{(B.12)} \quad &\quad - c'(\mathbf{s}_m, \pi_m, \hat{V}_{S^m}) \times \rho_{S^m}^T \Psi(\mathbf{s}_m) \leq 4M \end{aligned}$$

Note that Assumption A.2 implies that the cost function is Lipschitz continuous, so there is

$$|c(\mathbf{s}, \pi, \hat{V}) - c(\mathbf{s}_m, \pi_m, \hat{V}_{S^m})| \leq C_\pi |\hat{V} - \hat{V}_{S^m}|.$$

Now we need to derive a bound for $|\hat{V} - \hat{V}_{S^m}|$ using equation (B.12). By Assumption A.3 (ii), Cauchy-Schwarz inequality and the kernel property, we have

$$|\hat{V} - \hat{V}_{S^m}| \leq \|\hat{V} - \hat{V}_{S^m}\| \|k\| \leq \kappa \sqrt{\frac{8M}{J}}.$$

Therefore, for any $\mathbf{s}_m \in \mathcal{S}$, $|c(\mathbf{s}, \pi, \hat{V}) - c(\mathbf{s}_m, \pi_m, \hat{V}_{S^m})|$ will be bounded by $C_\pi \kappa \sqrt{\frac{8M}{J}}$. The factor decreases to zero when $J \rightarrow \infty$. By BE theorem, equation (B.8) turns to

$$\begin{aligned} \Pr \left\{ \left| \frac{1}{J} \sum_{j=1}^J |c(\mathbf{s}_j, \pi_j, \hat{V})|_\epsilon - \frac{1}{J} \sum_{j=1}^J c(\mathbf{s}_j, \pi_j, \hat{V}) \right| > 4\epsilon + 2\varsigma \right\} \\ \leq 2 \exp \left(-\frac{J}{2} \left(\frac{\epsilon}{M} \right)^2 \left(1 + \frac{1}{M} (C_\pi \kappa) \right)^{-2} \right), \end{aligned}$$

where $\varsigma = C_\pi \kappa \sqrt{\frac{8M}{J}}$.

Q.E.D.

PROOF OF THEOREM 3.3: We first derive kernel-based approximation in the case of non-soft margin, then show the kernel-based policy iteration algorithm coincides with the exact policy iteration in the limit.

Set ξ and ξ^* equal to zero in (3.8), the optimization problem becomes

$$\begin{aligned} \text{(B.13)} \quad &\min_{\rho} \frac{1}{2} \|\rho\|^2 \\ &\text{s.t. } \pi_i = \rho^T \Psi(\mathbf{s}_i) \end{aligned}$$

where $\Psi = \Phi - \beta \sum (\rho^T \Phi(s')) \mathbf{p}(s'|s_j, s_{-j}, a)$. Taking the partial derivative of the Lagrangian with respect to ρ , we have $\rho = \sum_i \alpha_i \Psi(s_i)$. Similarly, we can set up the dual problem such that

$$(B.14) \quad \max_{\alpha} -\frac{1}{2} \sum \alpha_i \alpha_j K_{ij}^* + \sum \alpha_i \pi_i,$$

with $\alpha \geq 0$. This is a standard linear quadratic objective function. $K_{ij}^* = \langle \Psi(s_i), \Psi(s_j) \rangle$. The derivative of (B.14) with respect to α gives linear system $K^T \alpha_i = \pi_i$. α_i can be solved by a simple inversion. If we can prove α_i has a unique solution and does not has a zero element, by dual space property (Slater condition) we know the constraint $\pi_i = \rho^T \Psi(s_i)$ is strictly satisfied. Hence ‘‘Step-5’’ in the iterative policy algorithm is to find the policy rule that exactly approximation the value function which is also the purpose of the inner loop in NFXP.

Firstly we prove K is invertible and positive definite. Φ is based function. $\beta < 1$ and $|\mathbf{p}(s'|s_j, s_{-j}, a)| < 1$ imply that $\Psi^T \Psi$ has positive eigenvalues. Secondly, we prove all element in α_i is non-zero. Substitute $K^T \alpha_i = \pi_i$ into (B.14), equation (B.14) reduces to $\sum \alpha_i K^T \alpha_j / 2$. If we suppose $\alpha_1 = 0$, it implies (B.14) has $\pi_1 \leq 0$. Since $\pi_i = \left[\sum_j \alpha_j \Psi(s_j) \right]^T \Psi(s_i) > 0$, contradiction. *Q.E.D.*

PROOF OF THEOREM 4.1: The EL objective function with Lagrange multipliers is

$$\mathcal{L} = \sum_{i=1}^n \log(n g_i) - n \lambda^T \sum_{i=1}^n g_i m_i(\theta) + \gamma \left(\sum_{i=1}^n g_i - 1 \right),$$

where λ and γ are Lagrange multipliers. The partial derivative of \mathcal{L} w.r.t. g_i gives $\gamma = n$. So we have implied probability $\tilde{g}_i = 1/(n + n \lambda^T m_i(\theta))$. By implicit function theorem, the partial derivative of $\sum_{i=1}^n \log \tilde{g}_i$ w.r.t λ gives a function of θ such that

$$(B.15) \quad \frac{\partial \sum \log \tilde{g}_i}{\partial \lambda} = 0, \\ \implies \frac{1}{n} \sum_{i=1}^n \frac{m_i(\theta)}{1 + \lambda_n(\theta)^T m_i(\theta)} = \sum_{i=1}^n \tilde{g}_{\theta, i} m_i(\theta)$$

where $\lambda_n(\theta)$ is unique for fix n and θ . Note that $(\partial \sum \log \tilde{g}_i / \partial \lambda)(\theta) = 0$ for $\forall \theta \in \Theta$, hence $(\partial \sum \log \tilde{g}_i / \partial \lambda)(\theta)$ is continuous on θ . By the continuity of $m(x, \theta)$, we have $\lambda_n(\theta)$ is also continuous on θ . The proof of the uniqueness of limited $\lambda_n(\theta)$ is straightforward, because the set $\Gamma(\theta) = \lim_{n \rightarrow \infty} \cap_{i=1, \dots, n} \{ \lambda | 1 + \lambda m_i(x, \theta) > 1/n \}$ is convex if it is not vanish, the function

of $\log g$ is strictly concave on λ , so $\lambda(\theta)$ uniquely exists. Equation (B.15) can be re-written as

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \left[1 - \frac{\lambda_n(\theta)^T m_i(\theta)}{1 + \lambda_n(\theta)^T m_i(\theta)} \right] m_i(\theta) &= 0 \\ \implies \frac{1}{n} \sum_{i=1}^n m_i(\theta) &= \frac{1}{n} \sum_{i=1}^n \frac{m_i(\theta) \lambda_n(\theta)^T m_i(\theta)}{1 + \lambda_n(\theta)^T m_i(\theta)} \\ &= \left[\sum_{i=1}^n \tilde{g}_{\theta,i} m_i(\theta) m_i(\theta)^T \right] \lambda_n(\theta). \end{aligned}$$

Assumption 5 (iii) shows that $m_i(\theta) m_i(\theta)^T$ is positive definite, let \mathbf{c} is the maximum eigenvalues of $\mathbb{E}(m(\theta) m(\theta)^T)$ and v is the corresponding eigenvector. The convex combination of $m_i(\theta) m_i(\theta)^T$ is bounded by $v^T \mathbf{c} v$. Let $K = \|v^T \mathbf{c} v\|$. According to Assumption 5 (iv), $m_i(\theta)$ has an envelop function $b(\theta)$ such that $\liminf_{\theta} |m(x, \theta)|/b(\theta) \geq 1$, $\lim_{n \rightarrow \infty} \lambda_n(\theta)/b'(\theta) \geq 1$ for any θ where $b'(\theta) = b(\theta)/K$.

The optimization problem becomes

$$\theta = \arg \max \sum_{i=1}^n \log n \tilde{g}_{\theta,i} / n = \sum_{i=1}^n \Lambda_n(\theta) / n.$$

Let's first consider the existence of the limitation.

$$\begin{aligned} \text{(B.16)} \quad \lim_{n \rightarrow \infty} \mathbb{E} \frac{1}{n} \frac{n}{1 + \lambda_n(\theta)^T m(x, \theta)} &= \mathbb{E} \lim_{n \rightarrow \infty} \frac{1}{1 + \lambda_n(\theta)^T m(x, \theta)} \\ &= \mathbb{E} \frac{1}{1 + \lambda(\theta)^T m(x, \theta)}. \end{aligned}$$

The above equation is obtained by dominated convergence theorem, since $[1 + \lambda_n(\theta)^T m(x, \theta)]^{-1}$ is bounded and the limit of $\lambda_n(\theta)$ exists. Next we consider the continuity of $\lambda(\theta)^T m(x, \theta)$. Given n , we know that $\lambda_n(\theta)^T m(x, \theta)$ is continuous. The envelop functions $b'(\theta)$ and $b(\theta)$ are integrable and continuous (Assumption 5), $\lambda(\theta)^T m(x, \theta)$ is bounded by a continuous function. The continuity of $m(x, \theta)$ implies $\lambda(\theta)$ is also continuous.

Now choose a large enough compact set $\Theta_c \subset \Theta$ such that given δ

$$\mathbb{E}[v(x)] = \sup_{\theta \in \Theta_c} \frac{|\sum \Lambda_n(\theta) / n - \mathbb{E} \Lambda(\theta)|}{b^2(\theta)} \leq 1 - 3\delta$$

and $\inf_{\theta \in \Theta_c} |\Lambda(\theta)|/b(\theta) \geq 1 - \delta$, by Assumption 5. The strong LLN implies

$$\sup_{\theta \in \Theta_c} \frac{|n^{-1} \sum [\Lambda_n(\theta) - \mathbb{E} \Lambda(\theta)]|}{b^2(\theta)} \leq \frac{1}{n} \sum v(x_i) \leq 1 - 2\delta;$$

therefore

$$\begin{aligned} \left| \frac{1}{n} \sum [\Lambda_n(\theta) - \mathbb{E}\Lambda(\theta)] \right| &\leq (1 - 2\delta)b(\theta) \\ &\leq \frac{1 - 2\delta}{1 - \delta} |\mathbb{E}l_\theta| \leq (1 - \delta)\mathbb{E}l_\theta \end{aligned}$$

for $\forall \theta \neq \Theta_C$. This equation implies, for any δ

$$\left| \frac{1}{n} \sum \Lambda_n(\theta) \right| \geq \delta |\mathbb{E}\Lambda(\theta)| \geq \delta(1 - \delta)b_0.$$

The last inequality indicates that the any sequence T_n satisfying $\sum \Lambda_n(T_n)/n \rightarrow 0$ will locate in the compact set ultimately.

Now the rest of the proof follows Wald's method. The continuous $\lambda(\theta)m(x, \theta)$ in any x gives

$$\lim_{\theta' \rightarrow \theta} |\Lambda(\theta')(x) - \Lambda(\theta)(x)| = 0.$$

For any decreasing sequence U_l of open neighborhoods around θ of diameter l converging to zero, $\sup_{\theta' \in U_l} \Lambda(\theta')$ decrease for every l . So when l goes to 0, $\sup_{\theta' \in U_l} \Lambda(\theta')$ converges to $\Lambda(\theta)$ almost surely. Monotone convergence theorem implies

$$(B.17) \quad \lim_{l \rightarrow 0} \mathbb{E} \sup_{\theta' \in U_l} \Lambda(\theta')(x) = \mathbb{E} \lim_{l \rightarrow 0} \sup_{\theta' \in U_l} \Lambda(\theta')(x) = \mathbb{E}\Lambda(\theta)(x).$$

Let the open neighborhoods around θ_0 denote Θ_0 For any $\theta \in \Theta_C \setminus \Theta_0$, there exists open neighborhood U_θ around θ . The compactness of U_θ gives a finite subcover $U_{\theta_1}, \dots, U_{\theta_s}$ with sufficient small diameter. For sufficient large n , there is

$$\begin{aligned} \sup_{\theta \in U_{\theta_1}, \dots, U_{\theta_s}} \frac{1}{n} \sum \Lambda_n(\theta) &\leq \\ \sup_{1 \leq j \leq s} \frac{1}{n} \sum \sup_{\theta \in U_{\theta_j}} \Lambda_n(\theta) &\longrightarrow \sup_{1 \leq j \leq s} \mathbb{E} \sup_{\theta \in U_{\theta_j}} \Lambda_n(\theta) < 0 \end{aligned}$$

by (B.17) and LLN. If $T_n \subset U_\theta$, then $\sup_{\theta \in U_\theta} \sum \Lambda_n(\theta)/n$ is at least $\sum \Lambda_n(T_n)/n$, which by definition of T_n such that $\sum \Lambda_n(T_n)/n + o_p(1) = 0$ by LLN. Thus

$$\{T_n \subset U_\theta\} \subset \left\{ \sup_{\theta \in U_\theta} \frac{1}{n} \sum \Lambda_n(T_n) \geq \mathbb{E}\Lambda(\theta_0) - o_p(1) \right\}.$$

The probability of the event on the right side converges to zero as $n \rightarrow \infty$.

Q.E.D.

PROOF OF PROPOSITION 1: To prove K_θ is invertible, we need to prove K_θ is almost surely positive definite. If K_θ is positive semi-definite but not invertible, then there must be a non-zero vector τ such that $K_\theta\tau = 0$. Therefore $\tau^T K_\theta\tau = 0$ which contradicts our assumption about K_θ being positive definite.

$K_{\theta,n}$ and $S_{\theta,n}$ form a relatively compact sequence. Thus a subsequence $(K_{\theta,n_k}, S_{\theta,n_k}) \rightarrow (K_\theta, S_\theta)$. Le Cam's first lemma implies

$$(B.18) \quad \mathbb{E} \exp \left[\tau^T S_\theta - \frac{1}{2} \tau^T K_\theta \tau \right] = 1.$$

Because (B.18) holds for all τ , we can use a symmetrized method to simplify (B.18). For certain value τ and $-\tau$, we have

$$\mathbb{E} \left\{ \exp \left[\tau^T S_\theta - \frac{1}{2} \tau^T K_\theta \tau \right] + \exp \left[-\tau^T S_\theta - \frac{1}{2} \tau^T K_\theta \tau \right] \right\} = 2.$$

By $\cosh \tau^T S_\theta = (\exp \tau^T S_\theta + \exp(-\tau^T S_\theta))/2$, we have $\mathbb{E}[(\cosh \tau^T S_\theta) \exp(-\tau^T K_\theta \tau/2)] = 1$. Assume some τ_i give $\tau^T K_\theta \tau$ negative values, then

$$(B.19) \quad \mathbb{E} \left[\mathbb{I}_{\{\tau^T K_\theta \tau > 0\}} (\cosh \tau^T S_\theta) \exp(-\tau^T K_\theta \tau/2) \right] \\ \leq \mathbb{E} \left[(\cosh \tau^T S_\theta) \exp(-\tau^T K_\theta \tau/2) \right] = 1$$

where $\mathbb{I}_{\{\cdot\}}$ is an indicator function. However, since $\exp(-\tau^T K_\theta \tau/2) > 1$ when $\tau^T K_\theta \tau$ is negative and $(\cosh \tau^T S_\theta) > 0$, we know above equation can not be valid unless the set $\{\tau^T K_\theta \tau > 0\}$ is null. Therefore, for all τ , K_θ is positive definite. *Q.E.D.*

PROOF OF THEOREM 5.2: (i) When θ is given, by condition 6

$$(B.20) \quad \Lambda_n(\theta + \delta_n \tau_n, \theta) = \tau_n^T S_{\theta,n} - \frac{1}{2} \tau_n^T K_{\theta,n} \tau_n + o_{\tilde{p}_\theta}(1) \\ = -\frac{1}{2} \left[(K_{\theta,n}^{-1} S_{\theta,n} - \tau_n^T)^T K_{\theta,n} (K_{\theta,n}^{-1} S_{\theta,n} - \tau_n^T) \right. \\ \left. - (S_{\theta,n}^T K_{\theta,n}^{-1} S_{\theta,n}) \right] + o_{\tilde{p}_\theta}(1).$$

Similarly,

$$(B.21) \quad \Lambda_n(\theta + \delta_n \tau_n, \theta) = -\frac{1}{2} \left[(\delta_n(T_n - \theta) - \tau_n^T)^T K_n (\delta_n(T_n - \theta) - \tau_n^T) \right. \\ \left. - (\delta_n(T_n - \theta))^T K_n (\delta_n(T_n - \theta)) \right].$$

The difference between (B.20) and (B.21) tends to zero in probability $\tilde{G}_{\theta,n}$. The four quadratic terms and non-negativity of K_n and $K_{\theta,n}$ indicate that each quadratic term must be non-negative. The difference between $\tau_n^T K_{\theta,n} \tau_n - 2\tau_n^T S_{\theta,n}$ and $\tau_n^T K_n \tau_n - 2\tau_n^T \delta_n(T_n - \theta) K_n$ must converge to zero in probability, otherwise the arbitrary choosing of τ_n will make the equality invalid.

$$\tau_n^T K_n \tau_n - 2\tau_n^T \delta_n(T_n - \theta) K_n \sim \tau_n^T K_{\theta,n} \tau_n - 2\tau_n^T S_{\theta,n}$$

implies K_n converges to $K_{\theta,n}$ in probability. Therefore, $\delta_n(T_n - \theta)$ converges to $K_{\theta,n}^{-1} S_{\theta,n}$.

(ii) By proposition 1, we know that clustering points K_θ of $K_{\theta,n}$ are invertible. Since $\delta_n(T_n - \theta)$ converges to $K_{\theta,n}^{-1} S_{\theta,n}$, the limit of $\delta_n(T_n - \theta)$ is $K_\theta^{-1} S_{\theta,n}$ which implies $\delta_n(T_n - \theta)$ is bounded in probability $\tilde{G}_{\theta,n}$.

(iii) By corollary 2, we know the condition DQM can imply 6, thus intuitively the linear-quadratic equation (5.1) may coincide with S_n and K_n . The log-likelihood process can be rewritten as a centered log-likelihood process $X_n(\cdot)$ plus a shift item $b_n(\cdot)$:

$$\begin{aligned} \Lambda_n(\eta, \theta)(x) &= \overbrace{\frac{1}{n} \sum_{i=1}^n \log \frac{\tilde{g}_\eta}{\tilde{g}_\theta}(x_i) - \int \log \frac{\tilde{g}_\eta}{\tilde{g}_\theta}(x) dG_\theta}^{X_n(\eta)} \\ &\quad + \underbrace{\int \log \frac{\tilde{g}_\eta}{\tilde{g}_\theta}(x) dG_\theta}_{b_n(\eta)} + o_p(1). \end{aligned}$$

Let $\eta = \theta + \delta_n \tau$ and $\delta_n = n^{-1/2}$. Given λ value in the constraint of equation (4.12), $\log(\tilde{g}_\eta/\tilde{g}_\theta)(x_i)$ in $X_n(\eta)$ can be replaced by a linear quadratic formula w.r.t. τ , namely $\log(\tilde{g}_\eta/\tilde{g}_\theta)$ belongs to a smooth enough function class. Therefore the process $\eta \mapsto X_n(\eta)$ is an empirical process and $\sqrt{n}X_n(\eta) \rightsquigarrow X(\eta)$ by Donsker theorem (see van der Vaart (Example 19.9 1998)) where $X(\eta)$ is a Gaussian process. Note that $X(\eta)$ has mean $\int X(\eta) dG_\theta = 0$ and covariance kernel $K_\theta = \mathbb{E}_\theta X^2(\eta)$ under distribution G_θ . Le Cam's first lemma (contiguity) implies $\mathbb{E}_\theta \exp[X(\eta) + b(\eta)] = 1$ with the expectation taken under G_θ and $b_n(\eta) \rightarrow b(\eta)$. Gaussian properties give $b(\eta) = -(1/2)\mathbb{E}_\theta X^2(\eta)$. By proposition 1 and equation (5.1), we can show that

$$\begin{aligned} X(\eta) &= \tau^T S_\theta \\ b(\eta) &= -\frac{1}{2} \tau^T K_{\theta_0} \tau, \end{aligned}$$

and when $\theta = \theta_0$

$$\begin{aligned} X(\eta) &= \tau \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \delta_n m_i(\theta_0) \\ b(\eta) &= -\frac{1}{2} \tau \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \mathbb{E} \frac{\partial m(x, \theta_0)}{\partial \theta} \tau. \end{aligned}$$

It means that K_θ varying slowly enough as θ varies, one can locally approximate $(T_n - \theta)^T K_\theta (T_n - \theta)$ by Gaussian shift experiments. *Q.E.D.*

PROOF OF THEOREM 5.3: The proof follows the strategies of [van der Vaart](#) (Proposition 8.6 1998) and [Le Cam and Yang](#) (Theorem 6.1 1990). The difficulty comes from the conditional expectation over local parameter τ . In Local EL the approximation is evaluated for each discretized θ_n^* . This implicitly imposes a priori assumption that true parameter concentrates on set $\{\eta : |\eta - \theta| \leq \delta_n \tau\}$ in every grid. We use a Bayesian argument for “local prior measures”.

$b \wedge W(Z_n - \tau)$ gives a bounded function. We can consider the expectation as $b \wedge \mathbb{E}[W(Z_n - \tau)|\theta + \delta_n \tau]$. By the LAN property of T_n from Theorem 5.2(iii), it is straightforward to set the prior measures with Gaussian densities. Since both “prior” and “posterior” concentrate around θ_0 , the updating information only occurs for covariance matrix. Let τ be the gaussian centered at 0 with inverse covariance Γ . The conjugate property indicates the posterior of τ can be written in terms of LEL representation:

$$Z_n = \delta_n^{-1}(T_n' - \theta) = (K_n + \Gamma)^{-1} K_n \delta_n^{-1}(T_n - \theta),$$

especially when $\Gamma = 0$, $Z_n = \delta_n^{-1}(T_n - \theta)$. Note that by Anderson’s lemma $\mathbb{E}[W(Z_n - \tau)|\theta + \delta_n \tau] \geq \mathbb{E}[W Z_n | \theta + \delta_n \tau]$ for bounded W . Since $K_n \delta_n^{-1}(T_n - \theta) \sim \mathcal{N}(0, I)$, the lower bound of $\mathbb{E}[W(Z_n - \tau)|\theta + \delta_n \tau]$ is

$$\mathbb{E} \left\{ W \left[(K_n + \Gamma)^{-1/2} \times \mathcal{N}(0, I) \right] | K_n \right\}.$$

K_n is independent with $\mathcal{N}(0, I)$. With the condition $K_n \rightsquigarrow K_\theta$ in \tilde{P}_θ law, the limit becomes $\mathbb{E} \left\{ W \left[(K_\theta + \Gamma)^{-1/2} \times \mathcal{N}(0, I) \right] \right\}$.

When c is very large, the probability of normal prior $|\tau| > c$ is small enough thus

$$\liminf_n \sup_{|\tau| \leq c} \mathbb{E} \left\{ W \left[(K_n + \Gamma)^{-1/2} \times \mathcal{N}(0, I) \right] \right\} \geq \mathbb{E} \left\{ W \left[(K_\theta + \Gamma)^{-1/2} \times \mathcal{N}(0, I) \right] \right\} - \Delta$$

for small enough Δ . Let Γ go to zero, $Z_n = \delta_n^{-1}(T_n - \theta)$ obtains the lower bound $\mathbb{E}[W(K_\theta^{-1/2}) \times \mathcal{N}(0, I)]$. If $W = 1$ and $K_\theta = K_{\theta_0}$, by Theorem 5.2(iii) we achieve the efficient bound of semi-parametric estimators. Q.E.D.

APPENDIX C: AUXILIARY LEMMA

LEMMA 1 *If (i) $\partial m(x, \theta)/\partial \theta < \infty$ for any x and the rank of $\mathbb{E}[\partial m(x, \theta)/\partial \theta]|_{\theta_0}$ is $\dim(\theta_0)$. (ii) $\mathbb{E}[\lambda_n(\theta)^T m_i(\theta)]^2 < \infty$, the log-likelihood ratio between \tilde{g}_{θ_0} and $\tilde{g}_{\theta_0 + \delta_n \tau}$ can be approximated by a linear-quadratic expansion.*

PROOF OF LEMMA 1: The proof just directly applies Taylor expansion. Without loss of generality, we assume $\delta_n = n^{-1/2}$. Note that

$$(C.1) \quad m(x, \theta_0 + \delta_n \tau) = m(x, \theta_0) + \frac{1}{2} \delta_n \tau [\partial m(x, \theta)/\partial \theta]|_{\theta_0} + o_p(\delta_n^2).$$

The result $\lambda_n(\theta) = \left(\sum_{i=1}^n [m_i(\theta) m_i(\theta)^T] / n \right)^{-1} \sum_{i=1}^n m_i(\theta) / n + o_p(n^{-1/2})$ from [Qin and Lawless \(1994, Lemma 1\)](#) or [Owen \(2001, theorem 2.2\)](#) implies

$$\begin{aligned} \sum_{i=1}^n \log \frac{\tilde{g}_{\theta_0 + \delta_n \tau}(x_i)}{\tilde{g}_{\theta_0}} &= \sum_{i=1}^n \log \left[\lambda_n(\theta_0 + \delta_n \tau)^T m_i(\theta_0 + \delta_n \tau) - \lambda_n(\theta_0)^T m_i(\theta_0) \right] \\ &= \left\{ \left(\frac{1}{n} \sum_{i=1}^n m_i(\theta_0 + \delta_n \tau) \right)^T \times \right. \\ &\quad \left. \left(\frac{1}{n} \sum_{i=1}^n [m_i(\theta_0 + \delta_n \tau) m_i(\theta_0 + \delta_n \tau)^T] \right)^{-1} \sum_{i=1}^n m_i(\theta_0 + \delta_n \tau) \right\} \\ &\quad - \left(\frac{1}{n} \sum_{i=1}^n m_i(\theta_0) \right)^T \left(\frac{1}{n} \sum_{i=1}^n [m_i(\theta_0) m_i(\theta_0)^T] \right)^{-1} \sum_{i=1}^n m_i(\theta_0). \\ &\quad + n o_p(n^{-1/2}) O_p(1) o_p(n^{-1/2}) \end{aligned}$$

The middle term in the big bracket can be simplified as:

$$\begin{aligned} &\frac{1}{n} \sum_{i=1}^n \left[[m_i(\theta_0) m_i(\theta_0)^T] + \delta_n \tau \frac{\partial m_i(\theta_0)^T}{\partial \theta} m_i(\theta_0) \right. \\ &\quad \left. + \frac{(\delta_n \tau)^2}{4} \frac{\partial m_i(\theta_0)^T}{\partial \theta} \frac{\partial m_i(\theta_0)}{\partial \theta} + o_p(\delta_n^3) \right] \\ &= \frac{1}{n} \sum_{i=1}^n [m_i(\theta_0) m_i(\theta_0)^T] + \frac{1}{n} \delta_n O_p(n^{1/2}) + o_p(n^{-2}) + o_p(n^{-5/2}). \end{aligned}$$

With equation (C.1), the big bracket becomes

$$\begin{aligned}
& n \left[\frac{1}{n} \sum_{i=1}^n m_i(\theta_0) + \frac{1}{2n} \sum_{i=1}^n \delta_n \tau \frac{\partial m_i(\theta_0)^T}{\partial \theta} \right]^T \left(\frac{1}{n} \sum_{i=1}^n [m_i(\theta_0) m_i(\theta_0)^T] \right)^{-1} \\
& \times \left[\frac{1}{n} \sum_{i=1}^n m_i(\theta_0) + \frac{1}{2n} \sum_{i=1}^n \delta_n \tau \frac{\partial m_i(\theta_0)^T}{\partial \theta} \right] \\
& = n \left[\frac{1}{n} \sum_{i=1}^n m_i(\theta_0) + \frac{\delta_n \tau}{2} \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \right. \\
& \quad \left. + \delta_n O(n^{-1/2} (\log \log n)^{1/2}) \right]^T \\
& \quad \times \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^n m_i(\theta_0 + \delta_n \tau) \right].
\end{aligned}$$

Thus the local EL is

$$\begin{aligned}
2 \sum_{i=1}^n \log \frac{\tilde{g}_{\theta_0 + \delta_n \tau}}{\tilde{g}_{\theta_0}}(x_i) & = \tau \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \delta_n \sum_{i=1}^n m_i(\theta_0) \\
& \quad - \frac{1}{2} \tau \mathbb{E} \frac{\partial m(x, \theta_0)^T}{\partial \theta} \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1} \mathbb{E} \frac{\partial m(x, \theta_0)}{\partial \theta} \tau \\
& \quad + o_p(1).
\end{aligned}$$

Note that $O(n^{-1/2} (\log \log n)^{1/2}) \times \delta_n \sum_{i=1}^n m_i(\theta_0) = o_p(1)$ and $\lim_{n \rightarrow \infty} A_n \sum_i [m_i(\theta_0 + \delta_n \tau) - m_i(\theta_0)]/n = o_p(1)$ with $A_n = \sum m_i(\theta_0) \left(\mathbb{E} m(x, \theta_0) m(x, \theta_0)^T \right)^{-1}$ by the continuity of $m_i(\theta)$. *Q.E.D.*

LEMMA 2 *Moment differentiable assumption can induce DQM. DQM can induce condition LAQ. But the converse statements are not true.*

PROOF OF LEMMA 2: Let $\mathbb{E} Y_i^2(\theta) = \mathbb{E} [\lambda_n(\theta) m_i(\theta)]^2 < \infty$, we have $\sum_{i=1}^n \Pr(Y_i^2(\theta) > n) < \infty$. The Borel-Cantelli lemma implies only finite number of n satisfies $\Pr(|Y_i(\theta)| > n^{1/2}) > 1 - \varepsilon$, for small enough ε . Therefore only finite number of $\max_{1 \leq i \leq n} |Y_i(\theta)|$ satisfies $\max_{1 \leq i \leq n} |Y_i(\theta)| > n^{1/2}$. We can introduce a constant A such that

$$\limsup_{n \rightarrow \infty} \max_{1 \leq i \leq n} |Y_i(\theta)| n^{-1/2} \leq A$$

holds with probability 1. Since A is arbitrary, we know $\max_{1 \leq i \leq n} |Y_i(\theta)| = o(n^{1/2})$.

Equation (B.15) implies

$$\frac{1}{n} \sum_{i=1}^n \frac{m_i(\theta)}{1 + Y_i(\theta)} = \frac{1}{n} \sum_{i=1}^n \frac{m_i(\theta + \tau)}{1 + Y_i(\theta + \tau)} = 0.$$

With some modifications, we have

$$\begin{aligned}
\text{(C.2)} \quad 0 &= \frac{1}{n} \sum_{i=1}^n \left\{ \frac{[1 + Y_i(\theta + \tau)]m_i(\theta) - [1 + Y_i(\theta)]m_i(\theta + \tau)}{[1 + Y_i(\theta + \tau)][1 + Y_i(\theta)]} \right\} \\
&\implies \frac{1}{n} \sum_{i=1}^n \frac{m_i(\theta) - m_i(\theta + \tau)}{[1 + Y_i(\theta + \tau)][1 + Y_i(\theta)]} \\
&= \frac{1}{n} \sum_{i=1}^n \frac{Y_i(\theta)m_i(\theta + \tau) - Y_i(\theta + \tau)m_i(\theta)}{[1 + Y_i(\theta + \tau)][1 + Y_i(\theta)]}
\end{aligned}$$

Multiply both side of (C.2) with τ and set $\tau \rightarrow 0$ does not violate the identical equation.

$$\begin{aligned}
\text{(C.3)} \quad &\frac{1}{n} \lim_{\tau \rightarrow 0} \sum_{i=1}^n \frac{1}{\tau} \frac{m_i(\theta) - m_i(\theta + \tau)}{[1 + Y_i(\theta + \tau)][1 + Y_i(\theta)]} = \\
&\frac{1}{n} \lim_{\tau \rightarrow 0} \sum_{i=1}^n \frac{1}{\tau} \frac{Y_i(\theta)m_i(\theta + \tau) - Y_i(\theta + \tau)m_i(\theta)}{[1 + Y_i(\theta + \tau)][1 + Y_i(\theta)]}
\end{aligned}$$

$[1 + Y_i(\theta + \tau)][1 + Y_i(\theta)]$ is bounded by previous argument $\max_{1 \leq i \leq n} |Y_i(\theta)| = o(n^{1/2})$ and $m_i(\theta) - m_i(\theta + \tau)$ is bounded by condition ?? . Dominated convergence theorem implies that the limit operator from LHS of (C.3) can be taken inside such that

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^n \lim_{\tau \rightarrow 0} \frac{1}{\tau} \frac{m_i(\theta) - m_i(\theta + \tau)}{[1 + Y_i(\theta + \tau)][1 + Y_i(\theta)]} &= \frac{1}{n} \sum_{i=1}^n \lim_{\tau \rightarrow 0} \frac{1}{\tau} \frac{m_i(\theta) - m_i(\theta + \tau)}{[1 + o(n^{1/2})]} \\
&= \frac{1}{n} \sum_{i=1}^n \frac{\partial m_i(\theta)}{\partial \theta} \times C_i < \infty.
\end{aligned}$$

$n^{-1} \sum_{i=1}^n m_i(\theta)/[1 + Y_i(\theta)] = 0$ for all θ , this equation is differentiable at θ , thus

$$\frac{1}{n} \sum_{i=1}^n \lim_{\tau \rightarrow 0} \frac{1}{\tau} \left\{ \frac{[1 + Y_i(\theta + \tau)]m_i(\theta) - [1 + Y_i(\theta)]m_i(\theta + \tau)}{[1 + Y_i(\theta + \tau)][1 + Y_i(\theta)]} \right\}$$

exists. This induces that the limit of the term in LHS of (C.3) exists. Because $\lim_{\tau \rightarrow 0} Y_i(\theta)m_i(\theta + \tau)/\tau$ is simply $Y_i(\theta)\partial m_i(\theta)/\partial \theta$, so we can deduce $\partial Y_i(\theta)/\partial \theta$ exists. It is straightforward to see that $\partial \log \tilde{p}_\theta/\partial \theta$ is differentiable.

Therefore $\sqrt{\tilde{g}_\theta}$ is continuously differentiable. Since \tilde{g}_θ is continuous, Lemma 7.6 in [van der Vaart \(1998\)](#) indicates that equation (5.2) is valid.

We already induces equation (5.2) holds.

$$\text{(C.4)} \quad \left\| \tilde{g}_{\theta+\tau}^{1/2} - \tilde{g}_\theta^{1/2} \right\|^2 \longrightarrow \left\| \frac{1}{2} \tau^T S_{\theta,n} \tilde{g}_\theta^{1/2} \right\|^2,$$

when $\tau \rightarrow 0$. Replace the local parameter with shrinking neighborhood δ_n , a sequence of n . Then given $\tau \rightarrow 0$ and $\|\tau^T S_{\theta,n} \tilde{g}_\theta^{1/2}\|$ is a real-value, we have

$$\int [\delta_n^{-1} (\tilde{g}_{\theta+\delta_n\tau}^{1/2} - \tilde{g}_\theta^{1/2})]^2 d\mu = O(1).$$

Since δ_n^{-1} is an increasing sequence, $\tilde{g}_{\theta+\delta_n\tau}^{1/2} - \tilde{g}_\theta^{1/2}$ will converge zero in $L^2(\mu)$. In other words, $\tilde{g}_{\theta+\delta_n\tau}^{1/2}$ converges to $\tilde{g}_\theta^{1/2}$ in L^2 norm. Theorem 7.2 in [van der Vaart \(1998\)](#) shows that for a bounded sequence τ_n ,

$$(C.5) \quad \log \prod_{i=1}^n \frac{\tilde{g}_{\theta+\delta_n\tau_n}(x_i)}{\tilde{g}_\theta} = \tau_n^T S_{\theta,n} - \frac{1}{2} \tau_n^T K_\theta \tau_n + o_{\tilde{p}_\theta}(1).$$

Namely, the differences between $\Lambda_n(\theta + \delta_n\tau_n, \theta)$ and a linear-quadratic equation tend to zero in \tilde{g}_θ probability. *Q.E.D.*

APPENDIX D: COMPUTATION IMPROVEMENT

Sequential Minimal Optimization (SMO)

The optimization solver in problems (3.12) and (4.12) is SMO developed by [Platt \(1998\)](#). The structure of SMO is close to sequential quadratic programming algorithm. SMO decomposes the overall quadratic programming (QP) problem into QP sub-problems. Thus a large QP problem can be broken down into a series of smaller QP problems, whose ultimate goal is to identify all of the non-zero Lagrange multipliers and discard all of the zero Lagrange multipliers.

We use problem (3.12) to illustrate SMO. For Local EL, the procedure is similar except that one need to compute a matrix of implied probability functions instead of kernel functions. Let $\Psi(s)^T \Psi(s) = K$ and $(\alpha_1, \dots, \alpha_j)$, (π_1, \dots, π_j) as vector α and π . Rewrite equation (3.12) as:

$$\begin{aligned} \min_{\alpha, \alpha^*} & -\frac{1}{2} [\alpha^T, (\alpha^*)^T] \begin{bmatrix} K & -K \\ -K & K \end{bmatrix} \begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix} \\ & + [\epsilon I^T + \pi^T, \epsilon I^T - \pi^T] \begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix} \\ \text{s.t.} & 0 \leq \alpha_j^*, \alpha_j \leq C/|\mathcal{S}|. \end{aligned}$$

SMO carries out iteration only on two elements each time. Suppose working set is (i, j) at k -th iteration while α_{-ij} and π_{-ij} are the vectors α and π without i -th and j -th elements. K_{-ij} is the K matrix without i, j -th rows and columns. The sub-problem becomes

$$\begin{aligned} \min_{\alpha_i, \alpha_j^*} & -\frac{1}{2} [\alpha_i, \alpha_j^*] \begin{bmatrix} K_{ii} & -K_{ij} \\ -K_{ij} & K_{jj} \end{bmatrix} \begin{bmatrix} \alpha_i \\ \alpha_j^* \end{bmatrix} \\ & + \left(\begin{bmatrix} \pi_i \\ \pi_j \end{bmatrix} + K_{-ij} \alpha_{-ij}^k \right)^T \begin{bmatrix} \alpha_i \\ \alpha_j^* \end{bmatrix} \\ \text{s.t.} & 0 \leq \alpha_j^*, \alpha_j \leq C/|\mathcal{S}|. \end{aligned}$$

Once solving the sub-QP problem at k -th iteration, select another i, j for $k + 1$ -th sub-QP. Like Gauss-Seidel or Gauss-Jacobi method, the purpose of applying sub-QP is to find a simple matrix pattern. Because in general case, matrix K may not have such a good sparse pattern as Figure 3. The decomposition of a big QP into small pieces QP can avoid solving a big linear system introduced by non-zero matrix K .

REFERENCES

- AGUIRREGABIRIA, V., AND P. MIRA (2007): “Sequential Estimation of Dynamic Discrete Games,” *Econometrica*, 75(1), 1–53.
- (2009): “Dynamic Discrete Choice Structural Models: A Survey,” *forthcoming in Journal of Econometrics*.
- ARIFOVIC, J. (1994): “Genetic Algorithm Learning and the Cobweb Model,” *Journal of Economic Dynamics and Control*, 18(1), 3–28.
- BAJARI, P., L. BENKARD, AND J. LEVIN (2007): “Estimating Dynamic Models of Imperfect Competition,” *Econometrica*, 75(3), 1331–1370.
- BAJARI, P., V. CHERNOZHUKOV, H. HONG, AND D. NEKIPELOV (2009): “Semiparametric Estimation of Dynamic Games of Incomplete Information,” .
- BOUSQUET, O., AND A. ELISSEEFF (2002): “Stability and Generalization,” *Journal of Machine Learning Research*, (2), 499–526.
- CAI, Y., AND K. JUDD (2009): “Stable and Efficient Computational Methods for Dynamic Programming,” *Working Paper*.
- DORASZELSKI, U., AND A. PAKES (2007): “A Framework for Applied Dynamic Analysis in IO,” *Handbook of Industrial Organization*, pp. 1889–1962.
- ERICSON, R., AND A. PAKES (1995): “Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *The Review of Economic Studies*, 62(1), 53–82.

- FERNANDEZ-VILLAVERDE, J., J. F. RUBIO-RAMREZ, AND M. S. SANTOS (2006): “Convergence Properties of the Likelihood of Computed Dynamic Models,” *Econometrica*, 74(1), 93–119.
- HALL, P. (1984): “Central limit theorem for integrated square error of multivariate nonparametric density estimators,” *Journal of Multivariate Analysis*, 14(1), 1–16.
- HOTZ, J., AND R. MILLER (1993): “Estimating Dynamic Models of Imperfect Competition,” *The Review of Economic Studies*, 60(3), 497–529.
- HUBER, P. (1967): “The behavior of maximum likelihood estimates under nonstandard conditions,” *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 221–233.
- HUBER, P. (1981): *Robust Statistics*. Wiley, New York.
- KEANE, M. P., AND K. I. WOLPIN (1994): “The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence,” *The Review of Economics and Statistics*, 76(4), 648–672.
- KITAMURA, Y., AND M. STUTZER (1997): “An Information-Theoretic Alternative to Generalized Method of Moments Estimation,” *Econometrica*, 65(4), 861–874.
- KITAMURA, Y., G. TRIPATHI, AND H. AHN (2004): “Empirical Likelihood-Based Inference in Conditional Moment Restriction Models,” *Econometrica*, 72(6), 1667–1714.
- LE CAM, L. (1974): *Notes on Asymptotic Methods in Statistical Decision Theory*. Universit de Montral, Centre de Recherches Mathmatiques.
- LE CAM, L., AND G. YANG (1990): *Asymptotics in Statistics: Some Basic Concepts (Springer Series in Statistics)*. Springer-Verlag, New York.
- MAGNAC, T., AND D. THESMAR (2002): “Identifying Dynamic Discrete Decision Processes,” *Econometrica*, 70(2), 801–816.
- NEWKEY, W., AND R. J. SMITH (2004): “Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators,” *Econometrica*, 72(1), 219–255.
- ORMONEIT, D., AND S. SEN (2002): “Kernel-based reinforcement learning,” *Machine Learning*, 49(2-3), 161–178.
- OWEN, A. (1988): “Empirical likelihood ratio confidence intervals for a single functional,” *Biometrika*, 75(2), 237–249.
- (1990): “Empirical Likelihood Ratio Confidence Regions,” *The Annals of Statistics*, 18(1), 90–120.
- (2001): *Empirical Likelihood*. Chapman & Hall/CRC, Florida.
- PAKES, A. (1986): “Patents as Options: Some Estimates of the Value of Holding European Patent Stocks,” *Econometrica*, 54(4), 755–784.
- PAKES, A., M. OSTROVSKY, AND S. BERRY (2008): “Simple Estimators for the Parameters of Discrete Dynamic Games (with entry/exit examples),” *The RAND Journal of Economics*, 38(2), 373–399.
- PAKES, A., AND D. POLLARD (1989): “Simulation and the Asymptotics of Optimization Estimators,” *Econometrica*, 57(5), 1027–1057.
- PESENDORFER, M., AND P. SCHIMIDT-DENGLER (2003): “Identification and Estimation of Dynamic

- Games,” *Working Paper*.
- PLATT, J. (1998): *Fast training of support vector machines using sequential minimal optimization*, Advances in Kernel Methods - Support Vector Learning. MIT, Cambridge.
- QIN, J., AND J. LAWLESS (1994): “Empirical Likelihood and General Estimating Equations,” *The Annals of Statistics*, 22(1), 300–325.
- RADNER, R. (1981): “Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship,” *Econometrica*, 49(5), 1127–1148.
- (1986): “Can Bounded Rationality Resolve the Prisoners’ Dilemma?,” *Contributions to Mathematical Economics*.
- RUST, J. (1987): “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher,” *Econometrica*, 55(5), 999–1033.
- (1997): “Using Randomization to Break the Curse of Dimensionality,” *Econometrica*, 65(3), 487–516.
- (2008): “Comments on: “Structural vs. Atheoretic Approaches to Econometrics, by Michael Keane”,” *forthcoming in Journal of Econometrics*.
- RUST, J., J. TRAUB, AND H. WOZNAKOWSKI (2002): “Is There a Curse of Dimensionality for Contraction Fixed Points in the Worst Case?,” *Econometrica*, 70(1), 285–329.
- SCHENNACH, S. M. (2007): “Point estimation with exponentially tilted empirical likelihood,” *The Annals of Statistics*, 35(2), 634–672.
- SU, C.-L., AND K. JUDD (2008): “Constrained Optimization Approaches to Estimation of Structural Models,” *Working Paper*.
- VAN DER VAART, A. (1998): *Asymptotic Statistics*. Cambridge University Press, Cambridge.
- VAN DER VAART, A., AND J. WELLNER (1996): *Weak Convergence and Empirical Process: With Applications to Statistics*. Springer-Verlag, New York.
- VAPNIK, V. (1998): *Statistical Learning Theory*. Wiley, New York.
- WALD, A. (1949): “Note on the consistency of the maximum likelihood estimate,” *Annals of Mathematical Statistics*, 20(4), 595–601.
- WELLFORD, C. (1989): “A Laboratory Analysis of Price Dynamics and Expectations in the Cobweb Model,” *Discussion Paper*.
- WOLPIN, K. I. (1984): “An Estimable Dynamic Stochastic Model of Fertility and Child Mortality,” *The Journal of Political Economy*, 92(5), 852–874.

TABLE II

DIFFERENT COST FUNCTIONS

Type of Function	Structural Parameters	Kernel Parameter	MSE and r^2	Vector No.	Local EL
Power	$\theta_1 = 0.10, \theta_2 = 0.10$	$C = 1, \epsilon = 0.1$	0.031, 0.840	21	-51.908
	$\theta_1 = 0.32, \theta_2 = 0.50$	$C = 0.2, \epsilon = 0.5$	0.778, 0.665	23	-20.957
	$\theta_1 = 2.50, \theta_2 = 1.10$	$C = 5, \epsilon = 0.1$	29.576, 0.6457	101	-142.909
Quadratic	$\theta_1 = 0.10, \theta_2 = 0.10$	$C = 1, \epsilon = 0.1$	0.007, 0.401	22	-34.521
	$\theta_1 = 0.32, \theta_2 = 0.50$	$C = 0.2, \epsilon = 0.5$	0.151, 0.636	27	-15.039
	$\theta_1 = 2.50, \theta_2 = 1.10$	$C = 5, \epsilon = 0.1$	10.620, 0.710	70	-137.577
Mixed	$\theta_1 = 0.10, \theta_2 = 0.10$	$C = 1, \epsilon = 0.1$	1.281, 0.596	58	-39.274
	$\theta_1 = 0.32, \theta_2 = 0.50$	$C = 0.2, \epsilon = 0.5$	13.987, 0.621	83	-142.697
	$\theta_1 = 2.50, \theta_2 = 1.10$	$C = 5, \epsilon = 0.1$	1049.497, 0.603	112	-400.521

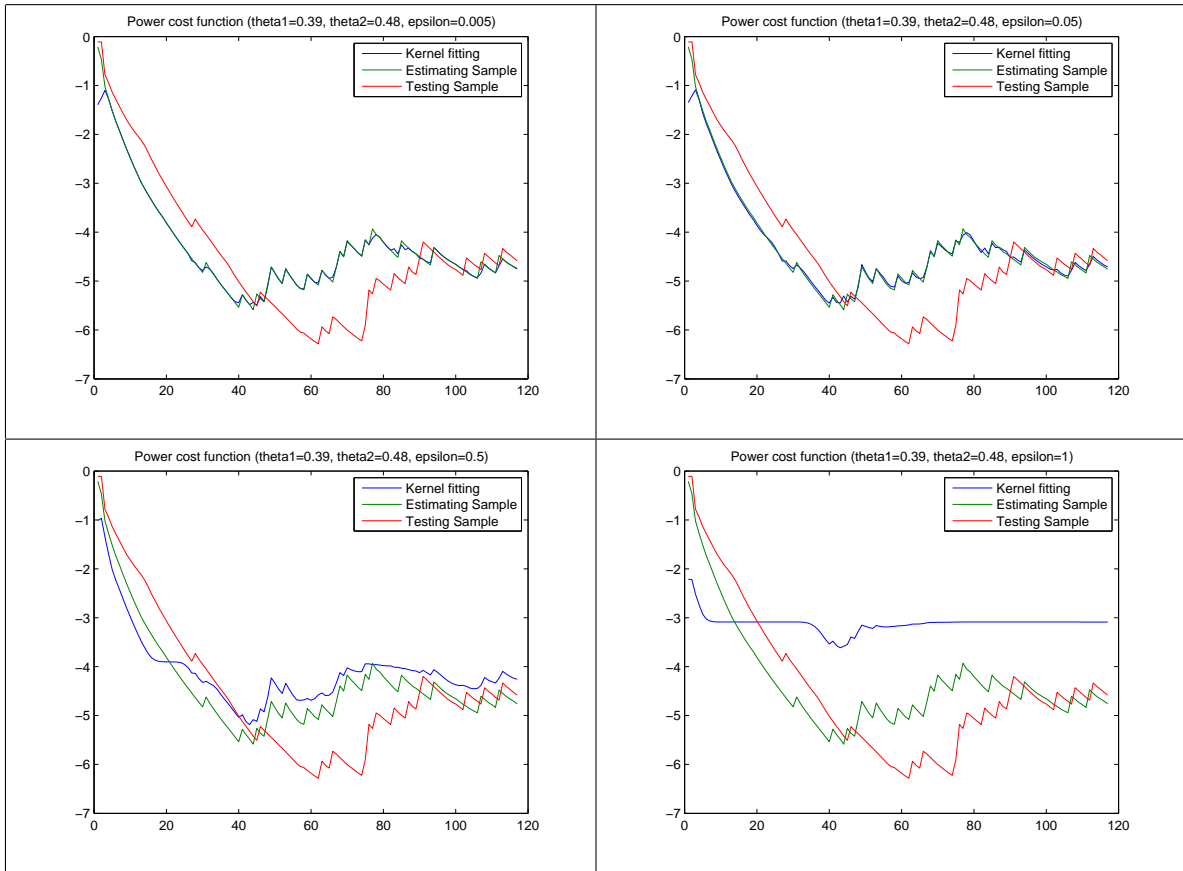


FIGURE 4.— ϵ -effects on value functions estimation.

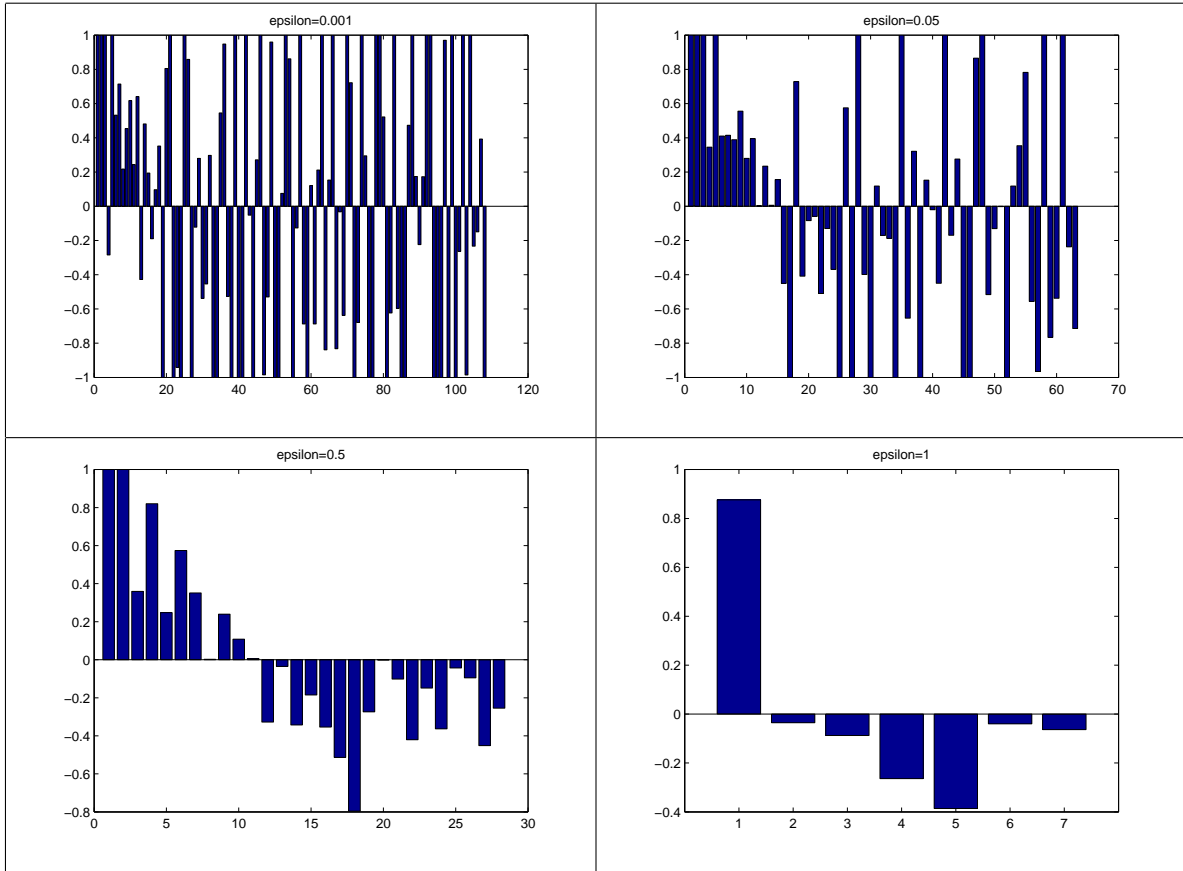


FIGURE 5.— Parameter

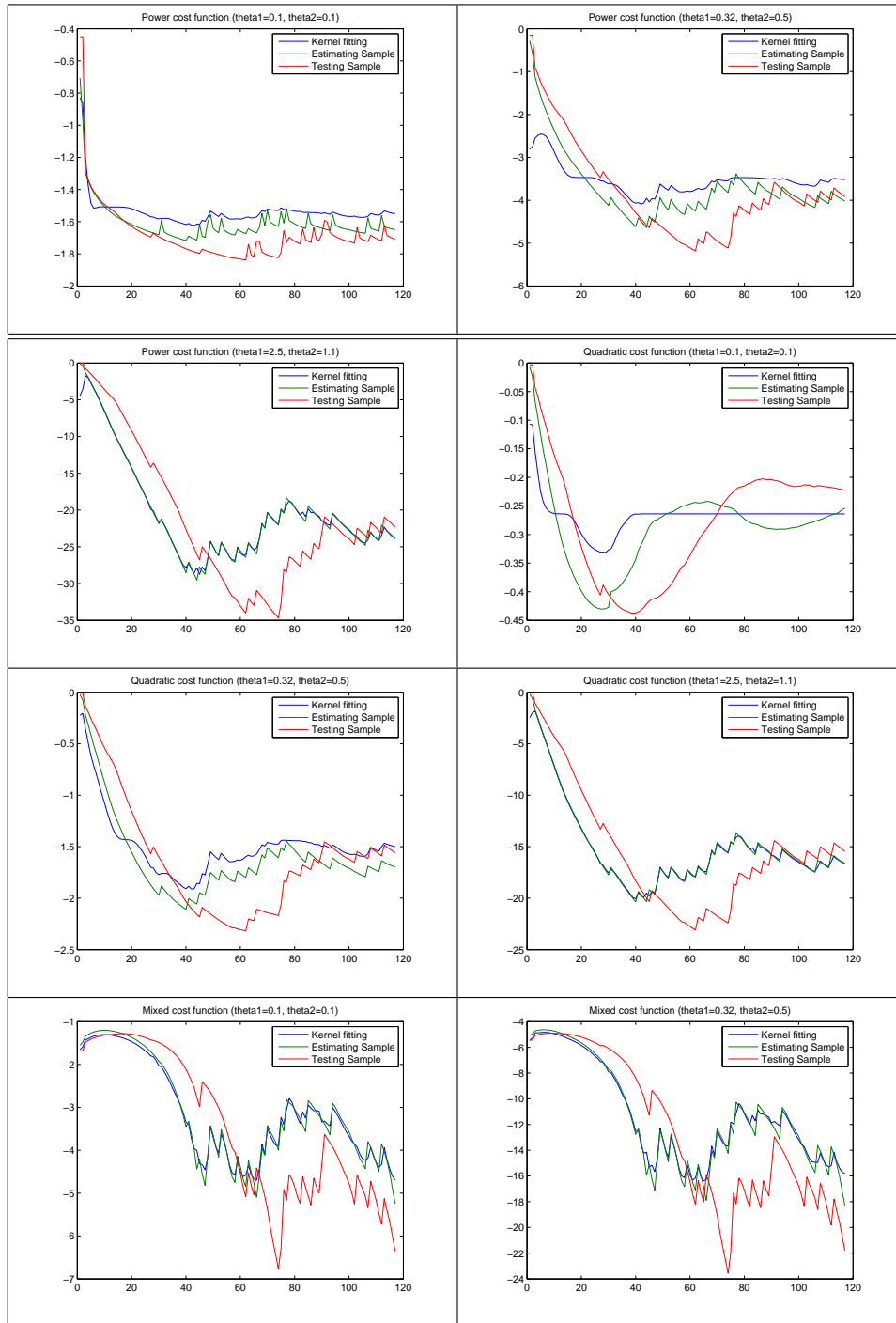


FIGURE 6.— Different cost functions