Generalized Maximum Entropy estimation of discrete sequential move games of perfect information

Wang, Yafeng and Graham , Brett

The Wang Yanan Institute for Studies in Economics (WISE)

20. December 2009

Generalized Maximum Entropy Estimation of Discrete Sequential Move Games of Perfect Information*

Yafeng Wang† Brett Graham‡

Wang Yanan Institute for Studies in Economics (WISE)

Xiamen University

December 20, 2009

Abstract

We propose a data-constrained generalized maximum entropy (GME) estimator for discrete sequential move games of perfect information which can be easily implemented on optimization software with high-level interfaces such as GAMS. Unlike most other work on the estimation of complete information games, the method we proposed is data constrained and does not require simulation and normal distribution of random preference shocks. We formulate the GME estimation as a (convex) mixed-integer nonlinear optimization problem (MINLP) which is well developed over the last few years. The model is identified with only weak scale and location normalizations, monte carlo evidence demonstrates that the estimator can perform well in moderately size samples. As an application, we study the social security acceptance decisions in dual career households.

Keywords: Game-Theoretic Econometric Models; Sequential-Move Game; Generalized Maximum Entropy; Mixed-Integer Nonlinear Programming

JEL Classification Numbers: C01, C13, C35, C51, C72.

*Very preliminary version, comments are welcome.

†E-mail: econwang@econwang.co.cc

‡E-mail: brett.d.graham@gmail.com

1 INTRODUCTION

Nash equilibrium is one of the cornerstones of modern economic theory, with substantive application in all major fields in economics, particularly industrial organization. It is the benchmark theoretical model for analyzing strategic interactions among a handful of players. Given the importance of gaming in economic theory, the empirical analysis of games has been the focus of a recent literature in econometrics and industrial organization, such as Golan, Karp and Perloff (1998, 2000) (hereafter GKP), Haile, Hortacsu and Kosenok (2003), Tamer (2003), Seim (2005), Aguirregabiria and Mira (2007), Aradillas-Lopez (2007, 2008), Bajari, Hong, John Krainer and Nekipelov (2009) and Bajari, Hong and Ryan (2009) (hereafter BHR).

Econometrically, a discrete game is a generalization of a standard discrete choice model, such as the conditional logit or multinomial probit. An agent's utility is often assumed to be a linear function of covariates and a random preference shock. However, unlike a discrete choice model, utility is also allowed to depend on the actions of other agents. Such modeling strategy was first suggested by the seminal work of Bresnahan and Reiss (1990, 1991). Although there are numerous studies on both methodology and empirical applications of game-theoretic models, the most widely studies is the class of incomplete information simultaneous-move games (normal form) and dynamic games, see Bajari, Hong, Krainer and Nekipelov (2009) and Aguirregabiria and Mira (2007). The complete information games received fewer studies due to its computational complexity, since it involves multidimensional integrals. More recently, BHR (2009) provides simulation-based estimators (moreover, Method of Simulated Moments (MSM)) for static complete information discrete games based on importance sampling. Furthermore, estimation of sequential-move (extensive form) games has been quite limited, especially on its' general form, Berry (1992), Mazzeo (2002) and Schmidt-Dengler (2006) estimate some simplified sequential-move games with special game structure. The estimation of the general class of sequential games has suffered from its computational complications, for sequential-move games, to the best of our knowledge, Maruyama (2009) was the only existing literature, which provides a simulation-based estimator for the general class of discrete-choice perfect information sequential game with a modified version of the GHK simulator (Geweke (1989, 1991), Hajivassiliou and McFadden (1998) and Keane (1990, 1994)), which he called as "sequential GHK". The estimator provided by Maruyama (2009) essentially is a maximum simulated likelihood (MSL) estimator, As is well known, MSL is biased for any fixed number of simulations, in order to obtain $\sqrt{T}$ consistent estimators, one needs to increase the number of draws S so that $\frac{S}{\sqrt{T}} \to \infty$. Such estimator requires larger scale simulation and also relies on the normal distribution of random preference shocks. Thus, computational burden also exists and makes its application adjective.

In this paper, we propose a data-constrained generalized maximum entropy (GME) estimator for discrete sequential-move games of perfect information which can be easily implemented since it does not require simulation, moreover, it also does not rely on the normality of random preference shocks. In the spirit of GKP (1998, 2000), the first application of GME to the estimation of game-theoretic models, and Su and Judd (2008), which argues that the direct optimization approach, called the MPEC (Mathematical Programming with Equilibrium Constraints) to structural estimation that avoids repetitive solution of the structural model is more powerful than traditional procedures such as the Nest Fixed-Point (NFXP) algorithm (Rust, 1987), we formulate the GME estimation as a mixed-integer nonlinear optimization (MINLP) problem which also be a direct optimization problem. Moreover, when the deterministic part of the payoff function is linear, it will be a convex MINLP, such optimization problems are well developed over the last few years (Grossmann (2002), Nowak and Vigerske (2008), Bonami, Kilinc and Linderoth (2009)) and there are several state-of-the-art solvers incorporated into many software packages (such as Tomlab which you can call from Matlab) and optimization modeling language, such as GAMS and AMPL[1], the user need

[1] Although such modeling languages (softwares) are commercial, you can use a free internet service, NEOS

not make decision about the algorithmic details thus our estimator is easy to use.

Econometrically, the main concern of such estimation problems is to formulate the critical function, possible choices are likelihood function or some distance function such as method of moments. Unfortunately, with a general structural model, such functions involve the multidimensional integrals, most studies alleviate this problem by simplifying the model structure or make use of some simulation-assisted estimation method, such as MSM and MSL, even with the simulaiton-based method which known to obtain many prefer large sample properties, they always need large draws and then computational burden incurred. Furthermore, such simulation methods always not easy to use which also limit its application. Instead of using simulation to deal with the multidimensional integrals, we overcome this problem by using the data-constrained equilibrium conditions which treat all the random shocks not observed by econometricians as endogenous parameters, thus the parametric distribution assumption wiped off. With these data-constrained equilibrium conditions, the nature choice of the critical function is the entropy (Shannon, 1948) and then formulate a GME problem. The GME principle was introduced by Golan et al. (1996), which is based on the classic maximum entropy (ME) approach of Jaynes (1957a, 1957b, 1984), which uses the entropy-information measure of Shannon (1948) to recover the unknown probability distribution of underdetermined problems, and started a new discussion in econometrics (among others, Golan, Judge and Perloff (1997), Mittelhammer and Cardell (1997), Golan, Perloff, and Shen (2000), Golan (2003), and Nunez, G. (2009)). The GME estimators are obtained by a constrained optimization problem which maximized the entropy objective function constrained by the model properties (such as equilibrium conditions), due to the structure of the perfect information sequential move game, the equilibrium (which known as sub-game perfect equilibria) conditions contain logical connections between endogenous variables, as a result, the common constrained optimization problem comes to be a mixed-

Server (http://neos.mcs.anl.gov/neos/), which gives the user acess to several state-of-the-art solvers such as the MINLP solvers, BARON and BONMIN.

integer nonlinear optimization problem, since we can always modify logical statements with integer variable, moreover, *zero* or *one* variables (H.P. Williams, 1985). With the efficient algorithms such as Branch and Bound (BB), Outer-Approximation (OA) and Hybrid OA based Branch-and-Cut (B-Hyb), we can solve this GME problem accurately, as shown below, with a linear payoff function, our GME problem is a convex MINLP, which can be exactly solved by most of the existing algorithms (Bonami, Kilinc and Linderoth (2009)).

Our approach makes several contributions to the literature on estimating game theoretic models, especially the complete (perfect) information case. First, our approach avoids the usual multidimensional integrals by using the data constraints instead of the moment constraints in complete (perfect) information case, the computational burden is acceptable for most applications. Although we focus on the sequential-move game, our approach can be extend to static game of complete information. GKP (1998, 2000) also make use of the GME to estimate the static game, their constraints are moment based since they deal with the incomplete information case. Second, there is no need for the normality of random preference shocks in our approach, this assumption is prerequisite for the existing estimators for general complete information games, such as BHR (2009) for static case and Maruyama (2009) for sequential-move case. Although BHR (2009) only make the assumption that such distribution should be known to any parametric distribution, mostly the choice only can be normal since we've no prior information about that. And for Maruyama (2009), since GHK simulator can only work under normal distribution, it highly relies on the normal assumption. Third, we reformulate the estimation problem as a MINLP since there are logical connections between endogenous variables among the equilibrium conditions, to the best of our knowledge, our estimator is the first one which makes use of MINLP in econometric estimation problems[2], since our monte carlo shows the validity of this estimation procedure, this reformulation can be extended to other estimation problems where logical statements

[2] Jouneau-Siona and Torrès (2006) formulate Maximized Monte Carlo (MMC) test as a Mixed Integer Programming problem.

incurred. The most shortcoming of our approach is that it is very hard to construct the large sample properties if not impossible, then the exactly tests and inference procedures can not be provided. As argued by Su and Judd (2008), to use such MPEC style results to compute standard errors, we need to work through the implicit construction of the critical function to formulate the exactly Hessian of the critical function with respect to the structural parameters, such work seems hard within the MINLP framework. Mittelhammer and Cardell (1997) provide the large sample distributions for GME estimator of general linear models, also, they prove the consistency and asymptotic normality. Since our monte carlo simulations show the consistency and asymptotic normality of the proposed estimator, following the arguments of Horowitz (1995, 1998, 2001), Campbell and CarterHill (2001) and Su and Judd (2008), we use the paired bootstrap methods to construct standard errors and related inference, although the bootstrap may not provide the asymptotic refinements since our estimator essentially is obtained from a nonsmooth optimization. Campbell and CarterHill (2001) also shows how to reformulate linear inequality restrictions in GME framework.

The paper is organized as follows. In section 2 we outline the general discrete sequential-move game to be estimated and formulate its equilibrium conditions. For purposes of exposition, a simple $2 \times 2 \times 2$ sequential entry game also be provided, which will be used extremely in the following sections. A briefly reviews of the maximum entropy, generalized maximum entropy estimation and the (convex) mixed-integer nonlinear programming are presented in section 3. Although there is no exactly identification problem in GME framework (Golan et al., 1996), we discuss the identification issue from the nature of the game structure and equilibrium conditions in section 4, our GME estimation for the discrete sequential-move game of perfect information is also presented. Monte carlo simulations are conducted in section 5. Section 6 contains the empirical application to the social security acceptance decisions in dual career households. Section 7 concludes the paper.

2 THE MODEL

In the model, there are T independent repetitions of a sequential move game of perfect information (extensive form game). In each game there are $i = 1, ..., N_t$ players, each with the finite set of actions A_{it}. Define $A_t = \times_i A_{it}$ and let $a_t = (a_{1t}, ..., a_{it}, ...a_{Nt})$ denote a generic element of A_t. Without loss of generality, the order of subscripts for players $(1, ..., N_t)$ also represents the decision order of the sequential move game in each repetition, that means player 1 makes decision first and player N_t at the end. Player i's von Neumann-Morgenstern (vNM) utility is a map $u_{it} : A_t \to R$, where R is the real line. Since we study the perfect information case, the corresponding equilibrium concept is the subgame perfect equilibria (SPE), this can be achieved when every player expects no gain from individually deviating from its equilibrium strategy in its every subgame, the standard technique for solving the SPE is backward induction, furthermore, the finite sequential move game of perfect information where there is no player is indifference between any two outcomes has a unique SPE. We will sometimes drop the subscript t for simplicity when no ambiguity would arise.

Following Bresnahan and Reiss (1990, 1991), assume that the vNM utility of player i can be written as:

$$u_i(a, x, \epsilon_i; \theta) = f_i(x, a; \theta) + \epsilon_i(a) \tag{1}$$

In Equation (1), player i's vNM utility from action a is the sum of two terms. The first term $f_i(x, a; \theta)$ is a function which depends on a, the vector of actions taken by all of the players, covariates x, the players' characteristics and some other variables which influence the utility, and parameters θ, covariates x are observed to the econometrician. The second term is $\epsilon_i(a)$, a random preference shock which reflects the information about utility that is common knowledge to the players but not observed by the econometrician. Unlike Maruyama (2009), here the preference shocks depend on the entire vector of actions a, not just the actions taken by player i. As argued by BHR (2009), this is a more general setting and seems straightforward within the game framework, think about a simple entry game, the unobserved information of one player to econometrician may be different not only among players but also action vector dependent. $\epsilon_i(a)$ are assumed to be independent or some known dependence, let ϵ_i denote the vector of the individual $\epsilon_i(a)$ and ϵ_i denote the vector of all the shocks. we will discuss more about the structure of ϵ_i in the identification and estimation section.

As noted above, the equilibrium concept corresponding to the sequential move game of perfect information, SPE, is a equilibrium strategy profile which means that every player expects no gain from individually deviating from its equilibrium in every subgame. A strategy of player $i \in N$ is a function that assigns an action in A_i to each nonterminal history, a player's deviation form equilibrium holding other's decisions fixed does not mean that all the others make the same decision, it means the others follow the same strategy. But what can be observed is only the equilibrium actions (i.e. equilibrium outcome). Thus, for deriving the equilibrium conditions in our econometric model, we should make the others' action profile when one player deviating as endogenous variable. Formally, an SPE action profile, $a^{SPE} = (a_1^{SPE}, ...a_i^{SPE}, ...a_N^{SPE})$, is any solution for the decisions of the players that satisfies:

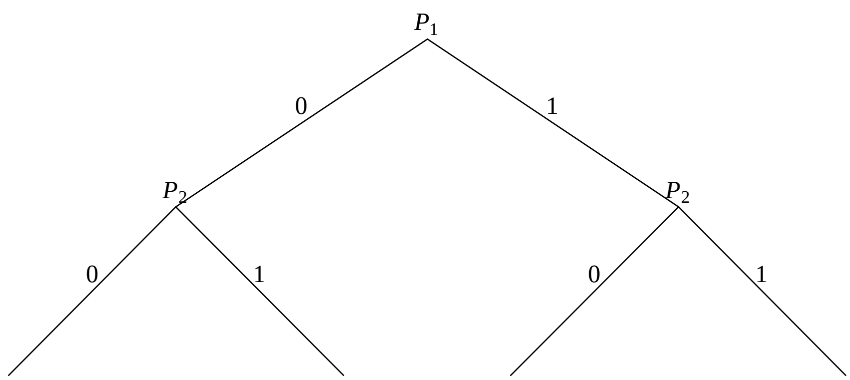
$$u_i(a_i^{SPE}, a_{-i}^{SPE}, x, \epsilon_i; \theta) - u_i(a_i, a_{<i}^{SPE}, a_{>i}^*(a_{<i}^{SPE}, a_i), x, \epsilon_i; \theta) \geq 0 \tag{2}$$
$$\textit{for all } i = 1, ..., N \textit{ and all } a_i \neq a_i^{SPE}.$$

where $a_{>i}^*(a_{<i}^{SPE}, a_i)$ is the unique SPE action profile for the subgame that starts from player $i+1$ given the decisions of the preceding players, $a_{\leq i}$. This equilibrium conditions are defined recursively and the solution can be easily calculated by the backward induction for any given parameters θ, observed covariates, x, and unobservable shocks ϵ.

Given such structure of the discrete choice sequential move game, our task is to estimate and draw an inference about the parameters of payoff functions, θ, with the observation of action profile a^o, some covariates which have effect on the payoffs, x, and an exogenous decision order. Note that the actual payoff levels are unobserved, since in most case, we can not determine what they should be, i.e. they are the latent variables.

For purposes of exposition, here we provide a simple $2 \times 2 \times 2$ sequential entry game as an example, which will be used extremely in our analysis. There are two players who act as the potential entrants in each of the T markets, the structure of this entry game is illustrated in Fig.1 with payoffs u_1 and u_2. The decision rule or the equilibrium conditions corresponding to (2) can be easily formulated, as an example, for action profile $(0,0)$ to be

Fig.1 *A Simple Entry Game*



$$
\begin{array}{cccc}
f_1(x,0,0;\theta)+\epsilon_1(0,0) & f_1(x,0,1;\theta)+\epsilon_1(0,1) & f_1(x,1,0;\theta)+\epsilon_1(1,0) & f_1(x,1,1;\theta)+\epsilon_1(1,1) \\
f_2(x,0,0;\theta)+\epsilon_2(0,0) & f_2(x,0,0;\theta)+\epsilon_2(0,1) & f_2(x,1,0;\theta)+\epsilon_2(1,0) & f_2(x,1,1;\theta)+\epsilon_2(1,1)
\end{array}
$$

an equilibria, equilibrium condition

$$
\begin{array}{lll}
u_1(x,0,0,\epsilon_1(0,0);\theta) > u_1(x,1,0,\epsilon_1(1,0);\theta) & if & u_2(x,1,0,\epsilon_2(1,0);\theta) > u_2(x,1,1,\epsilon_2(1,1);\theta) \\
u_1(x,0,0,\epsilon_1(0,0);\theta) > u_1(x,1,1,\epsilon_1(1,1);\theta) & if & u_2(x,1,0,\epsilon_2(1,0);\theta) \leq u_2(x,1,1,\epsilon_2(1,1);\theta) \\
u_2(x,0,0,\epsilon_2(0,0);\theta) \leq u_2(x,0,1,\epsilon_2(0,1);\theta) & &
\end{array}
\tag{3}
$$

should be satisfied, the equilibrium conditions for other three action profiles to be equilibrium actions can be formulated similarly. As noted above, since we only can observe the equilibrium actions but not the strategies, the equilibrium conditions contain the logical statements due to the off equilibrium path's choice, since the off equilibrium path's choices are unobserved to econometricians. We make use of MINLP to handle such logical statements.

3 PRELIMINARY

Since we use data-constrained GME approach to estimate the perfect information sequential move game in order to avoid the multidimensional integrals. We start by providing some background of how the generalized maximum entropy approach works, furthermore, our GME estimator is obtained via a (convex) MINLP, we also provide a basic review of the MINLP problem.

3.1 A Basic Review of GME

The GME estimation is based on the classic maximum entropy (ME) approach of Jaynes (1957a, 1957b, 1984), which uses the entropy-information measure of Shannon (1948) to recover the unknown probability distribution of underdetermined problems. In the classic ME approach, Shannon's (1948) entropy is used to measure the uncertainty (state of knowledge) we have about the occurrence of a collection of events. Letting x be a random variable with possible outcomes x_s, $s = 1, 2, \ldots, n$, with probabilities α_s such that $\sum_s \alpha_s = 1$, Shannon (1948) defined the entropy of the distribution $\alpha = (\alpha_1, ...\alpha_n)'$, as

$$H \equiv -\sum_s \alpha_s \ln \alpha_s \tag{4}$$

where $0 \ln 0 \equiv 0$. The function H, which Shannon interprets as a measure of the uncertainty in the mind of someone about to receive a message, reaches a maximum when $\alpha_1 = \alpha_2 = ... = \alpha_n = 1/n$. To recover the unknown probabilities α that characterize a given data set, Jaynes (1957a, 1957b) proposed maximizing entropy, subject to available sample-moment information and adding up constraints on the probabilities.

Obviously, within the classic ME framework, the observed moments are assumed to be exact. To extend this approach to the problems with noise, the GME approach (developed by Golan, Judge, and Miller, 1996) generalize the ME approach by using a dual objective (precision and prediction) function. We illustrate the GME approach via a linear model:

$$Y = X\beta + \varepsilon \tag{5}$$

where Y being a $N \times 1$ dependent variable vector, X being a $N \times K$ matrix of explanatory variables, β being $K \times 1$ a vector of parameters, and ε being a $N \times 1$ vector of disturbance terms. The GME rule for defining the estimator of the unknown β in this general linear model formulation is given by $\hat{\beta} = Z\hat{p}$ with $\hat{p} = (\hat{p}_1', ...\hat{p}_K')'$ derived from the following constrained maximum entropy problem:

$$\begin{aligned} \max_{p_k', w_i'} \quad & -\sum_{k=1}^{K} p_k' \ln(p_k) - \sum_{i=1}^{N} \omega_i' \ln(\omega_i) \\ s.t. \quad & Y = XZP + V\omega \\ & 1'p_k = 1, \ \forall k \\ & 1'\omega_i = 1, \ \forall i \\ & p_k > [0], \ \omega_i > 0. \ \forall i, k \end{aligned} \tag{6}$$

where Z and V are $K \times KM$ and $N \times NJ$ matrices of support points for the β and ε vectors, respectively, as:

$$Z = \begin{bmatrix} z_1' & 0 & ... & 0 \\ 0 & z_2' & ... & 0 \\ . & . & ... & . \\ 0 & 0 & ... & z_K' \end{bmatrix} \text{ and } V = \begin{bmatrix} v_1' & 0 & ... & 0 \\ 0 & v_2' & ... & 0 \\ . & . & ... & . \\ 0 & 0 & . & v_N' \end{bmatrix} \tag{7}$$

where $z_k = (z_{k1}, ...z_{kM})'$ is a $M \times 1$ vector such that $z_{k1} < z_{k2} \leq ... \leq z_{kM}$ and $\beta_k \in$

(z_{k1}, z_{kM}),[3] and similarly $v_i = (v_{i1}, ...v_{1J})'$ is a vector such that $v_{i1} < v_{i2} \leq ... \leq v_{1J}$ and $\varepsilon_i \in (v_{i1}, v_{1J})$, typically, v_{i1} and v_{1J} will be uniformly and symmetrically distributed about zero and have the same J dimensions. The actual bounds used for a given problem depend on the observed sample as well as any available conceptual or empirical information[4]. The $M \times 1$ p_k vectors and the $J \times 1$ ω_i vectors are weight vectors having nonnegative elements that sum to unity and are used to represent the β and ε vectors as $\beta = Zp$ and $\varepsilon = V\omega$. Golan, Judge and Miller (1996) has a rigorous discussion of this approach and applies to a rich scopes of econometric problems, such as dynamic model, model selection and discrete choice-consored problems. Mittelhammer and Cardell (1997) establish consistency and asymptotic normality results for the GME estimator under general regularity conditions on the specification of the estimation problem.

3.2 A Basic Review of MINLP

Since our GME estimator for the sequential move game essentially be obtained via a generalized disjunctive programming, which can be reformulated to a MINLP problem, we also provide a basic review of the general structure and feasible algorithms for the MINLP problem. MINLP provides a powerful framework for mathematically modeling optimization problems that involve discrete and continuous variables. Such optimization problems arise in many real world applications. Integer variables are often required to model logical relationships, fixed charges, piecewise linear functions, disjunctive constraints and the non-divisibility of resources. Nonlinear functions are required to accurately reflect physical properties, covariance, and economies of scale. Over the last few years there has been a

[3] This parameter support is based on prior information or economic theory, for example, we might specify boundaries of $z_{k1} = 0$ and $z_{kM} = 1$ when estimating the marginal propensity to consume, without any available prior information, we can specify z_k to be symmetric around zero, with large negative and positive boundaries. For example, $z_{k1} = -z_{kM} = -10^6$.

[4] One viable approach is to use Chebychev's Inequality or the three-sigma rule (Pukelsheim, 1994) and assume the errors are drawn from a uniform distribution with mean zero and variance $(y_{\max} - y_{\min})/12$. (A. Golan et al. 1997)

pronounced increase in the development of these models. The most basic form of an MINLP problem when represented in algebraic form is as follows:

$$\min_{\{x,y\}} \quad Z = f(x,y) \tag{8}$$

$$s.t.\ g_j(x,y) \le 0,\ j \in J$$

$$x \in X, y \in Y$$

where $f(\cdot)$, $g(\cdot)$ are differentiable functions, J is the index set of inequalities, and x and y are the continuous and discrete variables, respectively. The discrete set Y in most applications is restricted to $0-1$ values, $y \in \{0,1\}^m$. When $f(\cdot)$, $g(\cdot)$ both are convex functions, it turns to be a convex MINLP, actually, which can be exactly solved via most of existing algorithms, for the nonconvex case, only a few methods are available.

Methods that have addressed the solution of convex MINLP include the branch and bound method (BB) (Gupta and Ravindran, 1985; Nabar and Schrage, 1991; Borchers and Mitchell, 1994; Stubbs and Mehrotra, 1999; Leyffer, 2001), Generalized Benders Decomposition (GBD) (Geoffrion, 1972), Outer-Approximation (OA) (Duran and Grossmann, 1986; Yuan et al., 1988; Fletcher and Leyffer, 1994), LP/NLP based branch and bound (Quesada and Grossmann, 1992), and Extended Cutting Plane Method (ECP) (Westerlund and Pettersson, 1995). Methods for nonconvex MINLP include LP relaxation (Sherali & Adams, 1990), LP and SDP relaxations (Lov asz & Schrijver, 1991), SDP relaxations (Lasserre, 2001) and Branch-and-Reduce (Tawarmalani & Sahinidis, 2002). Such methods are involved in some optimization software with high-level interfaces such as GAMS, AMPL, and TOMLAB which has a MATLAB interface. In GAMS, the state-of-the-art solvers BARON and BONMIN both can handle the convex MINLP, but only BARON can handle nonconvex MINLPs in general, it implements a spatial branch-and-bound algorithm that is based on a factorable reformulation of the given problem and convexifications of univariate functions. We take the technique of solving MINLP as given.

4 ESTIMATION

Now we propose our GME estimator, in order to make use of the GME estimation, we need further assumptions about the utility functions. Although there is no exactly identification problem in GME framework (Golan et al., 1996), we discuss the identification issue from the nature the game structure and equilibrium conditions, which bring us introduce *Assumption 1*. And since the entropy is additive only for independent source of uncertainty, we also put the *i.i.d* assumption on random shocks for expositional clarity, any known heteroskedasticity and dependence among random shocks all can be handled within the GME framework.

ASSUMPTION 1 (Scale and Location Normalizations). *The payoffs of one action for each player are fixed at a known constant.*

As argued by BHR (2009), this restriction is similar to the argument that we can normalize the mean utility from the outside good equal to a constant, usually zero, in a standard discrete choice model. One clearly find that from the equilibrium condition (2) that adding a constant to all deterministic payoffs does not perturb the set of equilibria, so a location normalization is necessary. A scale normalization is also necessary, as multiplying all deterministic payoffs by a positive constant does not alter the SPE. Actually, without such normalizations, our GME estimator still work, but the level value of each estimated parameter does not make any sense, they are only significative in the ratio term. Thus, the normalizations which act as a prior information can improve our GME estimation, and also reduce the number of parameters, we would like to impose these location and scale normalizations.

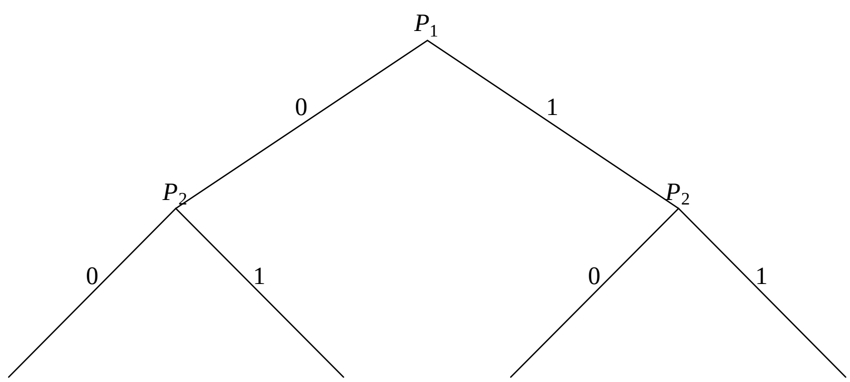
ASSUMPTION 2. (Regularity Conditions of Random Shocks). *The random preference shocks $\epsilon_{it}(a)$ are distributed i.i.d and independent of state variables with zero mean and limit variance, i.e.* $E(\epsilon_{it}(a)) = 0, E(x\epsilon_{it}(a)) = 0, Var(\epsilon_{it}(a)) < \infty$.

Assumption 2 which we need for establishing the GME estimation is more broad than the most other work does, such as Maruyama (2009), the normality of shocks is vital to that estimator relies on simulation. Such *i.i.d* assumption is not strict for our GME estimation, any known heteroskedasticity and dependence among random shocks all can be handled within the our GME framework, we will discuss more about it after presenting our estimator. Our GME estimator is semiparametric in terms of it does not impose any parametric assumption of the random shocks. For purposes of exposition, we use the simple entry game which has been introduced in section 2 to introduce the GME estimation, under *Assumption 1* and the specific utility function:

$$u_i(x, a, \epsilon_i; \theta) = 1(a_i = 1)\{\theta x + \delta g(a) + \epsilon_i(a)\} \tag{9}$$

The entry game turns to be which lists in Fig.2.

Fig.2 *A Reformulated Entry Game*



0	0	$\theta x + g(1,0) + \epsilon_1(1,0)$	$\theta x + g(1,1) + \epsilon_1(1,1)$
0	$\theta x + g(0,1) + \epsilon_2(0,1)$	0	$\theta x + g(1,1) + \epsilon_2(1,1)$

Obviously, in terms of the scale and location normalizations, we set the utility of out the market normalized to 0 and the parameter δ to 1. The equilibrium conditions for action profile $(0,0)$ to be SPE outcomes are:

$$\begin{aligned}
\text{Player1:} \quad & 0 > \theta x + g(1,0) + \epsilon_1(1,0) \quad if \quad 0 > \theta x + g(1,1) + \epsilon_2(1,1) \\
& 0 > \theta x + g(1,1) + \epsilon_1(1,1) \quad if \quad 0 \leq \theta x + g(1,1) + \epsilon_2(1,1) \\
\text{Player2:} \quad & 0 > \theta x + g(0,1) + \epsilon_2(0,1)
\end{aligned} \tag{10}$$

Similarly, the equilibrium conditions for $(0,1)$ to be SPE outcomes are:

$$\begin{aligned}
\text{Player1:} \quad & 0 > \theta x + g(1,0) + \epsilon_1(1,0) \quad if \quad 0 > \theta x + g(1,1) + \epsilon_2(1,1) \\
& 0 > \theta x + g(1,1) + \epsilon_1(1,1) \quad if \quad 0 \leq \theta x + g(1,1) + \epsilon_2(1,1) \\
\text{Player2:} \quad & \theta x + g(0,1) + \epsilon_2(0,1) \geq 0
\end{aligned} \tag{11}$$

for $(1,0)$ are:

$$\begin{aligned}
\text{Player1:} \quad & \theta x + g(1,0) + \epsilon_1(1,0) \geq 0 \quad if \quad 0 > \theta x + g(0,1) + \epsilon_2(0,1) \\
& \theta x + g(1,0) + \epsilon_1(1,0) \geq 0 \quad if \quad 0 \leq \theta x + g(0,1) + \epsilon_2(0,1) \\
\text{Player2:} \quad & 0 > \theta x + g(1,1) + \epsilon_2(1,1)
\end{aligned} \tag{12}$$

finally, for $(1,1)$ are:

$$\begin{aligned}
\text{Player1:} \quad & \theta x + g(1,1) + \epsilon_1(1,1) \geq 0 \quad if \quad 0 > \theta x + g(0,1) + \epsilon_2(0,1) \\
& \theta x + g(1,1) + \epsilon_1(1,1) \geq 0 \quad if \quad 0 \leq \theta x + g(0,1) + \epsilon_2(0,1) \\
\text{Player2:} \quad & \theta x + g(1,1) + \epsilon_2(1,1) \geq 0
\end{aligned} \tag{13}$$

In order to use the GME framework, we need to specify the support space for θ and $\epsilon(a)$, which we define as z, v_1, v_2, v_3, v_4 for θ, $\epsilon_{1t}(1,0)$, $\epsilon_{1t}(1,1)$, $\epsilon_{2t}(0,1)$ and $\epsilon_{2t}(1,1)$ respectively, without loss of generality, each of them are $M \times 1$ vector and $v_1 = v_2 = v_3 = v_4 = v$, the corresponding probabilities are defined as $p^\theta, \omega_t^1, \omega_t^2, \omega_t^3, \omega_t^4$ such that:

$$\theta = \sum_{m=1}^{M} p_m^\theta z_m \tag{14}$$

$$\epsilon_{1t}(1,0) = \sum_{m=1}^{M} \omega_{tm}^1 v \tag{15}$$

$$\epsilon_{1t}(1,1) = \sum_{m=1}^{M} \omega_{tm}^2 v \tag{16}$$

$$\epsilon_{2t}(0,1) = \sum_{m=1}^{M} \omega_{tm}^3 v \tag{17}$$

$$\epsilon_{2t}(1,1) = \sum_{m=1}^{M} \omega_{tm}^4 v \tag{18}$$

Our GME estimator is obtained from the estimated probabilities which are the solution of problem:

$$\begin{aligned} \max_{\{p_m^\theta, \omega_{tm}^1, \omega_{tm}^2, \omega_{tm}^3, \omega_{tm}^4\}} H \quad = \quad & -\sum_{m=1}^{M} p_m^\theta \ln(p_m^\theta) - \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{tm}^1 \ln(\omega_{tm}^1) - \\ & \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{tm}^2 \ln(\omega_{tm}^2) - \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{tm}^3 \ln(\omega_{tm}^3) - \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{tm}^4 \ln(\omega_{tm}^4) \end{aligned} \tag{19}$$

subject to the corresponding constraints which list in equation (10) to (13) with the reparameterized θ and ϵ:

$$\begin{array}{l|l|l} If & 0 > \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,0) + \sum_{m=1}^{M} \omega_{tm}^1 v_m & if \quad 0 > \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m \\ a_t^o = & 0 > \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^2 v_m & if \quad 0 \leq \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m \\ (0,0) & 0 > \sum_{m=1}^{M} p_m^\theta z_m x_t + g(0,1) + \sum_{m=1}^{M} \omega_{tm}^3 v_m & \end{array} \tag{20}$$

$$\begin{array}{l|l|l} If & 0 > \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,0) + \sum_{m=1}^{M} \omega_{tm}^1 v_m & if \quad 0 > \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m \\ a_t^o = & 0 > \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^2 v_m & if \quad 0 \leq \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m \\ (0,1) & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(0,1) + \sum_{m=1}^{M} \omega_{tm}^3 v_m \geq 0 & \end{array} \tag{21}$$

$$\begin{array}{l|l} If \ a_t^o = & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,0) + \sum_{m=1}^{M} \omega_{tm}^1 v_m \geq 0 \\ (1,0) & 0 > \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m \end{array} \tag{22}$$

$$\begin{array}{l|l} If \ a_t^o = & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^2 v_m \geq 0 \\ (1,1) & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m \geq 0 \end{array} \tag{23}$$

and the normalization-additvity constraints:

$$\begin{aligned} \sum_{m=1}^{M} p_m^\theta &= 1 \\ \sum_{m=1}^{M} \omega_{tm}^1 &= 1, \forall t \in T \\ \sum_{m=1}^{M} \omega_{tm}^2 &= 1, \forall t \in T \\ \sum_{m=1}^{M} \omega_{tm}^3 &= 1, \forall t \in T \\ \sum_{m=1}^{M} \omega_{tm}^4 &= 1, \forall t \in T \\ p^\theta, \omega_t^1, \omega_t^2, \omega_t^3, \omega_t^4 &> 0; \forall t \in T \end{aligned} \tag{24}$$

Note that for each market or each repetition of the game, there is unique equilibria, then the constraints for each market are one of the four possible constraints which list in equation (20) to (23). With the estimated $\hat{p}_m^\theta$, our GME estimator of the structure parameter θ will be:

$$\hat{\theta}_{GME} = \sum_{m=1}^{M} \hat{p}_m^\theta z_m \tag{25}$$

For this simple game, except for constraints (22) and (23), the constraints all contain the logical statements such as if... *then*... between endogenous variables, this programming is called disjunctive programming which can be reformulated as MINLP, for example, consider the logical statements:

$$\begin{aligned} y_1 < 0 \quad if \quad x < 0 \\ y_2 < 0 \quad if \quad x \geq 0 \end{aligned} \tag{26}$$

we can reformulate them to the statements with inter variables which can be easily handled in the MINLP problem by introducing a *zero* or *one* variable q, the statement (26) will be:

$$\begin{aligned} x - M(1-q) < 0 \\ x + Mq \geq 0 \\ y_1 < M(1-q) \\ y_2 < Mq \\ q = \{0, 1\} \end{aligned} \tag{27}$$

where M is a *big* positive variable which exceeds the bound of x such as $9.e10$. Such reformulations are discussed severely in H.P. Williams (1985) and Raman and Grossmann (1991). By introducing such a *zero* or *one* variable, our GME programming can be reformulated as the following MINLP problem:

$$\begin{aligned} \max_{\{p_m^\theta, \omega_{tm}^1, \omega_{tm}^2, \omega_{tm}^3, \omega_{tm}^4, q_t\}} H \;=\; & -\sum_{m=1}^{M} p_m^\theta \ln(p_m^\theta) - \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{tm}^1 \ln(\omega_{tm}^1) - \\ & \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{tm}^2 \ln(\omega_{tm}^2) - \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{tm}^3 \ln(\omega_{tm}^3) - \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{tm}^4 \ln(\omega_{tm}^4) \end{aligned} \tag{28}$$

$s.t.$

$$if\ a_t^o = (0,0) \left| \begin{aligned} \textstyle\sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m &\geq -Mq_t \\ \textstyle\sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m &< M(1-q_t) \\ \textstyle\sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,0) + \sum_{m=1}^{M} \omega_{tm}^1 v_m &< M(1-q_t) \\ \textstyle\sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^2 v_m &< Mq_t \\ 0 > \textstyle\sum_{m=1}^{M} p_m^\theta z_m x_t + g(0,1) + \sum_{m=1}^{M} \omega_{tm}^3 v_m & \end{aligned} \right.$$

$$
\begin{array}{l|r}
if\ a_t^o = (0,1) & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m \geq -M q_t \\
 & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^4 v_m < M(1-q_t) \\
 & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,0) + \sum_{m=1}^{M} \omega_{tm}^1 v_m < M(1-q_t) \\
 & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(1,1) + \sum_{m=1}^{M} \omega_{tm}^2 v_m < M q_t \\
 & \sum_{m=1}^{M} p_m^\theta z_m x_t + g(0,1) + \sum_{m=1}^{M} \omega_{tm}^3 v_m \geq 0
\end{array}
$$

$$
\begin{array}{l|r}
and & \text{equation (22), (23)} \\
 & \text{normalization-additvity constraints equation (24)} \\
 & q_t \in (0,1)
\end{array}
$$

Optimization (28) which can be solved via MINLP techniques yields the estimated probability for each unkonwns, which include the probabilities for our structural parameter θ, thus the estimated $\hat{\theta}$ can be recovered from the original reparameterization:

$$
\hat{\theta}_{GME} = \sum_{m=1}^{M} \hat{p}_m^\theta z_m \tag{29}
$$

Note that if the payoff function is linear in all covariates, x, then the optimization problem becomes a convex MINLP, since the objective entropy function is always a concave function. For the game which has more than two players, more than two actions, and more than two stages, the estimation (28) can be straightforwardly extended to involve more constraints, also more *zero* or *one* variables to hold more logical statements. One also can simplify the reformulation of logical conditions by investigating the recursive structure of the sequential move games.

The GME estimator proposed above essentially is a MPEC style estimator, as argued by Su and Judd (2008), implementing asymptotic inference methods is more complex with the MPEC approach. Computing standard errors requires the computation of the Hessian of the objective function with respect to structural parameters θ, such work seems hard within the

MINLP framework since our GME estimation problem is a non-smooth optimization. Although Su and Judd (2008) suggest use the bootstrap methods to construct standard errors which can avoid the finite sample bias that may raise with standard asymptotic methods, the asymptotic refinements may not be obtained in our GME estimation. Following the arguments of Horowitz (1995, 1998, 2001), little is known about the ability of the bootstrap to provide asymptotic refinements for hypothesis tests and confidence intervals based on such non-smooth estimators, but for widely range of non-smooth estimators, such as the least-absolute-deviations (LAD) estimator, bootstrap can provide a consistent approximation to the asymptotic distribution (De Angelis, et al., 1993; Hahn, 1995). In this sense, we also suggest use the bootstrap to get the standard errors for structural parameters and related inferences, and since our model make no parametric assumptions on random shocks, a nonparametric (paired) bootstrap will be the choice. Horowitz (1995, 1998, 2001) also explains how some non-smooth estimators can be smoothed in a way that greatly simplifies the analysis of the their asymptotic distributional properties, the bootstrap provides asymptotic refinements for hypothesis tests and confidence intervals based on the smoothed estimators. Smoothing our GME estimator is also possible since for most cases we can reformulate MINLP problem to nonlinear programming (NLP) with complementarity constraints (MPCC), Chen and Mangasarian (1996) provides a class of smoothing functions for nonlinear and mixed complementarity problems. Furthermore, recently there are some new resampling methods provided in order to deal with such nonregular estimation problems and estimators, Zeng and Lin (2008) based on asymptotic expansion via empirical process arguments suggests some efficient resampling procedures for non-smooth estimators, Andrews and Guggenberger (2009) also provides some efficient Hybrid and Size-Corrected subsampling methods. We will investigate these alternative methods within our GME estimation framework in another paper.

As noted above, our framework can deal with a wide range of random shocks' structures which depart the $i.i.d$ assumption, such as the market specific shocks which considered by

Maruyama (2009). Consider the simple entry game in Fig.2, when introducing a market specific information not observed by econometrician, η_i, and the total random preference shocks of player i in the market t specified as:

$$\epsilon_{it}(a) = \omega_{it}(a) + \eta_i \tag{30}$$

where $\omega_{it}(a)$ and η_i are both independently distributed across players (entrants) and markets, with this specific variance structure, we can deeply treat with variables $\omega_{it}(a)$ and η_i instead of $\epsilon_{it}(a)$, which also means a additive entropy objective function.

5 MONTE CARLO

To demonstrate the performance of our estimator in small samples, we conducted two Monte Carlo experiments using the simple sequential entry game which introduce in section 2 and 3. There are two players and each player has the following profit function:

$$u_i(x, a, \epsilon_i; \theta) = 1(a_i = 1)\{\theta_1 x_1 + \theta_2 x_{i2} - \theta_3 x_{i3} + \epsilon_i(a)\} \tag{31}$$

In the first experiment, we define $x_1 \sim N(10, 1)$, $x_{i2} \sim N(1, 1)$, and $x_{i3} = 9(N(a)-1)$, where $N(a)$ is the number of entrants for a action profile a, and $\epsilon_{it}(a)$, the idiosyncratic error term, are drawn from standard normal distribution. In the second experiment, we define two of the $\epsilon_{it}(a)$ drawn from uniform distribution $[-1, 1]$, others are same as experiment one.

As discussed previously, our model requires both scale and location normalizations, so we assume that $\theta_3 = 1$ and the payoffs of not entering are zero. Thus our game has two unknown parameters: θ_1 and θ_2. The game generates equilibrium conditions for each of the possible equilibrium action profiles which will be the constraints of our GME estimation. We generated 10000 samples of size $t = 25, 50, 100$, and 200 to assess the finite sample properties of our estimator. The true parameter vector was chosen as $\theta_1 = 1$ and $\theta_2 = -1$. The parameter estimates are presented in Table I and II, the empirical distributions of parameter estimates are reported in Fig.3 and Fig.4.

Table I: Monte Carlo Results for Normal Shocks

Parameter	Mean	Median	Standard Deviation	Mean Bias	Median Bias	MSE
			$T = 25$			
θ_1	1.1651	1.1350	0.1387	0.1651	0.1350	0.0465
θ_2	-2.5561	-2.3814	0.8934	-1.5561	-1.3814	3.2194
			$T = 50$			
θ_1	1.0471	1.0388	0.0555	0.0471	0.0388	0.0052
θ_2	-1.5566	-1.4607	0.4868	-0.5566	-0.4607	0.5466
			$T = 100$			
θ_1	0.9934	0.9924	0.0247	-0.0066	-0.0076	0.0006
θ_2	-1.0449	-1.0265	0.2215	-0.0449	-0.0265	0.0511
			$T = 200$			
θ_1	0.9944	0.9941	0.0142	-0.0056	-0.0059	0.0002
θ_2	-1.0403	-1.0344	0.1171	-0.0403	-0.0344	0.0153

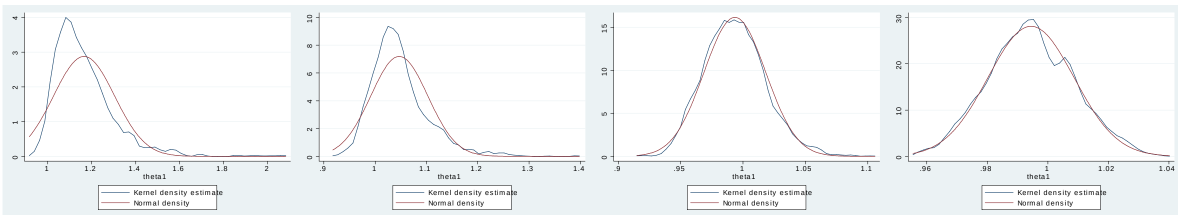
True value: $\theta_1 = 1$, $\theta_2 = -1$; Monte Carlo Times: 10000

Table II: Monte Carlo Results for Normal and Uniform Shocks

Parameter	Mean	Median	Standard Deviation	Mean Bias	Median Bias	MSE
			$T = 25$			
θ_1	1.1660	1.1378	0.1410	0.1660	0.1378	0.0474
θ_2	-2.5643	-2.4009	0.8977	-1.5643	-1.4009	3.2528
			$T = 50$			
θ_1	1.0461	1.0371	0.0548	0.0461	0.0371	0.0051
θ_2	-1.5472	-1.4479	0.4841	-0.5472	-0.4479	0.5338
			$T = 100$			
θ_1	0.9937	0.9931	0.0253	-0.0063	-0.0069	0.0007
θ_2	-1.0460	-1.0283	0.2238	-0.0460	-0.0283	0.0521
			$T = 200$			
θ_1	0.9946	0.9941	0.0141	0.0054	0.0059	0.0002
θ_2	-1.0435	-1.0383	0.1149	0.0435	0.0383	0.0151

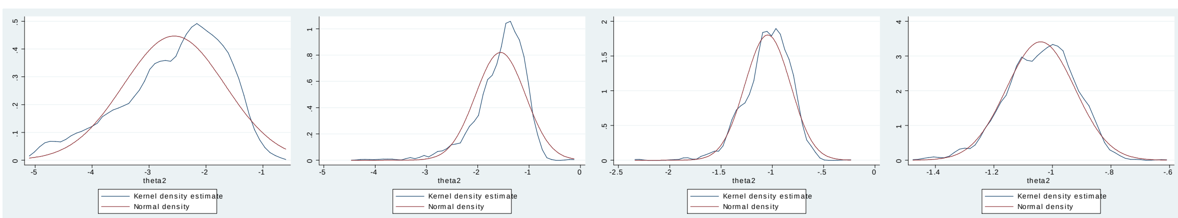
True value: $\theta_1 = 1$, $\theta_2 = -1$; Monte Carlo Times: 10000

Fig.3 Distribution of Estimators with Normal Shocks.



$T = 25$ $T = 50$ $T = 100$ $T = 200$

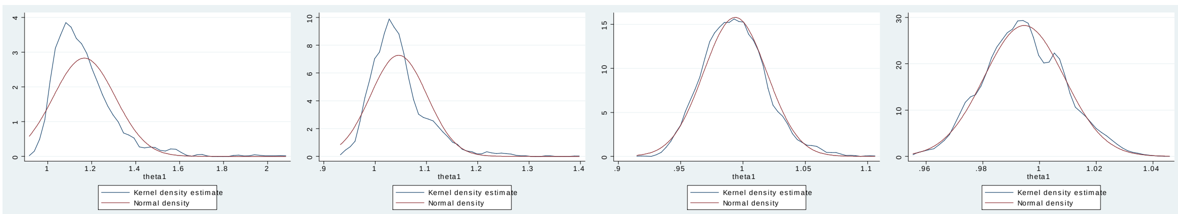
Distribution of Estimated θ_1 with Normal Shocks.



$T = 25$ $T = 50$ $T = 100$ $T = 200$

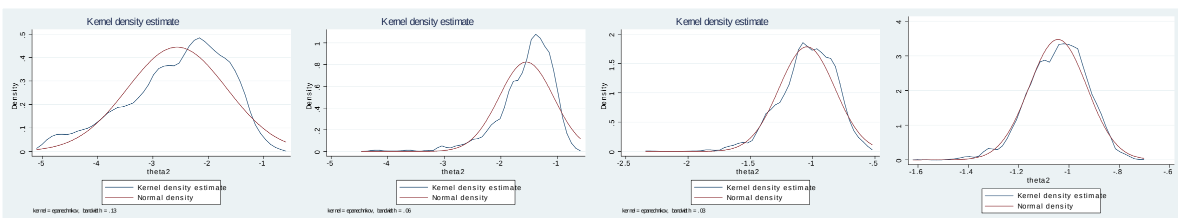
Distribution of Estimated θ_2 with Normal Shocks

Fig.4 Distribution of Estimators with Normal and Uniform Shocks



$T = 25$ $T = 50$ $T = 100$ $T = 200$

Distribution of Estimated θ_1 with Normal and Uniform Shocks



$T = 25$ $T = 50$ $T = 100$ $T = 200$

Distribution of Estimated θ_2 with Normal and Uniform Shocks

The results are encouraging even in the smaller samples sizes, the payoff parameters are estimated near their true values, and as the sample size increase, the estimates become more precisely. One may find that parameter θ_2 is estimated with much less precision, this mostly due to θ_1 has a larger influence over the equilibrium than a change in θ_2, since θ_1 multiplies a covariate with a higher mean than θ_2, even though they have the same average. In a extreme small sample, the change in θ_2 may not change the equilibrium actions, since you can see, with the sample size becomes larger, even θ_2 is estimated precisely.

The little difference between the two simulation outcomes shows that our GME estimator can handle not only the normal distribution. The empirical distributions of parameter estimates which list in Fig.3 and Fig.4 show that our GME estimator is asymptotically normal distributed.

6 APPLICATION

*******************TBW*********************

7 CONCLUSION

In this paper, we developed a data-constrained GME estimator for the discrete sequential move game of perfect information, which can be obtained via a MINLP. By directly using the data-constraints which implied by the equilibrium conditions, we avoid the multidimensional integrals which always make such estimation intractable. Moreover, our GME estimator also does not need the parametric assumption (mostly, normality) of the random shocks, this assumption is prerequisite for the existing estimators for general complete information games. We formulate the GME estimation as a (convex) mixed-integer nonlinear optimization problem (MINLP) which is well developed over the last few years. The estimation can be easily implemented on optimization software with high-level interfaces such as GAMS, AMPL. The model is identified with only weak scale and location normalizations, monte carlo evidence demonstrates that the estimator can perform well in moderately size samples. As an application, we study the social security acceptance decisions in dual career households.

References

[1]. Aguirregabiria, V. and P. Mira (2007). "Sequential Estimation of Dynamic Discrete Games." *Econometrica* 75 : 1-53.

[2]. Andrews, D. W. K. and P. Guggenberger (2009). "Hybrid and Size-Corrected Subsampling Methods." *Econometrica* 77 (3): 721-762.

[3]. De Angelis, D. and P. Hall, et al. (1993). "Analytical and Bootstrap Approximations to Estimator Distributions in L1 Regression." *Journal of the American Statistical Association* 88 : 1310-1316.

[4]. Bajari, P. and H. Hong, et al. (2009). "Estimating Static Models of Strategic Interactions." forthcoming *Journal of Business and Economic Statistics*.

[5]. Bajari, P. Han Hong, and Stephen Ryan (2009). "Identification and Estimation of a Discrete Game of Complete Information." forthcoming *Econometrica*.

[6]. Berry, Steven T. (1992). "Estimation of a Model of Entry in the Airline Industry." *Econometrica* 60 (4): 889-917.

[7]. Pierre Bonami, Mustafa Kilinc and Jeff Linderoth (2009). "Algorithms and Software for Convex Mixed Integer Nonlinear Programs." Technical Report.

[8]. Bresnahan, T. F. and P. C. Reiss (1990). "Entry in Monopoly Markets." *Review of Economic Studies* 57 : 531-553.

[9]. Bresnahan, T. F. and P. C. Reiss (1991). "Empirical models of Discrete Games." *Journal of Econometrics* 48 : 57-81.

[10]. Brooke, A., D. Kendrick, A. Meeraus, and R. Raman (2008). GAMS - A User's Guide, Washington, DC: GAMS Development Corporation. Online at: http://www.gams.com/dd/docs/bigdocs/GAMSUsersGuide.pdf.

[11]. Campbell, R. C. and R. C. Hill (2001). "Maximum Entropy Estimation in Economic Models with Linear Inequality Restrictions." Working Paper.

[12]. Chen and Mangasarian (1996). "A class of smoothing functions for nonlinear and mixed complementarity problems." *Computational Optimization and Applications* 5 (2): 97 - 138.

[13]. Golan, A., George Judge, and J. Miller (1996). *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York: John Wiley & Sons.

[14]. Golan, A. and G. Judge, et al. (1997). "Estimation and Inference with Censored and Ordered Multinomial Response Data." *Journal of Econometrics* 79 (1): 23-51.

[15]. Golan, A. and L. S. Karp, et al. (1998). "Estimating a Mixed Strategy: United and American Airlines." Working Paper.

[16]. Golan, A. and L. S. Karp, et al. (2000). "Estimating Coke's and Pepsi's Price and Advertising Strategies." *Journal of Business and Economic Statistics* 18 (4): 398-409.

[17]. Golan, A. and J. M. Perloff, et al. (2000). "Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand." Working Paper.

[18]. Ignacio E. Grossmann (2002). "Review of Nonlinear Mixed-Integer and Disjunctive Programming Techniques." *Optimization and Engineering* (3): 227–252.

[19]. Hahn, J. (1995). "Bootstrapping Quantile Regression Estimators." *Econometric Theory* 11 (01): 105-121.

[20]. Haile, P., and A. Hortacsu, et al. (2003). "On the Empirical Content of Quantal Response Equilibrium." Working Paper.

[21]. Horowitz, J. L. (1995). "Bootstrap Methods in Econometrics: Theory and Numerical Performance." Prepared for presentation to the Invited Symposium on Bootstrap Methods in Econometrics, 7th World Congress of the Econometric Society, Tokyo.

[22]. Horowitz, J. L. (1998). "Bootstrap Methods for Median Regression Models." *Econometrica* 66 (6): 1327-1352.

[23]. Horowitz, J. L. (2001). "The Bootstrap." In J. Heckman and E. Leamer (Ed.) *Handbook of Econometrics*, Vol.5, pp. 3159-3228, Elsevier Science, The Netherlands.

[24]. Jaynes, E. T. (1957a). "Information Theory and Statistical Mechanics." *Physics Reviews* 106: 620-630.

[25]. Jaynes, E. T. (1957a). "Information Theory and Statistical Mechanics, II." *Physics Reviews* 108: 171-190.

[26]. Jaynes, E. T. (1984). "Prior Information and Ambiguity in Inverse Problems." In McLaughlin (Ed.) *Inverse Problems*, pp. 151-166, SIAM Proceedings, American Mathematical Society, Providence, RI.

[27]. Jouneau-Siona and Torrès. (2006). "MMC techniques for limited dependent variables models: Implementation by the branch-and-bound algorithm." *Journal of Econometrics* 133 : 479–512.

[28]. Aradillas-Lopez (2007). "Pairwise Difference Estimation of Incomplete Information Games." Working Paper.

[29]. Aradillas-Lopez (2008). "Semiparametric Estimation of a Simultaneous Game with Incomplete Information." Working Paper.

[30]. Maruyama, S. (2009). "Estimating Sequential-move Games by a Recursive Conditioning Simulator." Working Paper.

[31]. Mazzeo, Michael J. (2002). "Product Choice and Oligopoly Market Structure." *Rand Journal of Economics* 33 (2): 221-242.

[32]. Mittelhammer, R. C. and N. S. Cardell (1997). "The Data-Constrained GME Estimator of the GLM: Asymptotic Theory and Inference." Working Paper.

[33]. I. Nowak and S. Vigerske (2008), "LaGO - a (heuristic) Branch and Cut algorithm for nonconvex MINLPs." *Central European Journal of Operations Research* 16 (2): 127-138.

[34]. Nunez, G. (2009). "A Maximum Entropy Estimation of Armington Elasticities for Mexico." Working Paper.

[35]. Martin J. Osborne and Ariel Rubinstein. (1994). *A Coure in Game Theory.* The MIT Press.

[36]. R. Raman and I. E. Grossmann. (1991). "Relation between MILP modelling and logical inference for chemical process synthesis." *Computers and Chemical Engineering* 15.

[37]. Klaus Ritzberger (2002). *Foundations of Non-Cooperative Game Theory.* Oxford University Press.

[38]. Rust, J. (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55 (5): 999-1033.

[39]. Schmidt-Dengler, P. (2006). "The Timing of New Technology Adoption: The Case of MRI." Working Paper.

[40]. Seim,. K. (2005). "An Empirical Model of Firm Entry with Endogenous Product-Type Choices." *Rand Journal of Economics* 37 (3): 619-640.

[41]. Shannon, C.E. (1948). "A Mathematical Theory of Communication." *Bell System Technical Journal* 27: 379-423.

[42]. Su, C. and K. L. Judd (2008). "Constrained Optimization Approaches to Estimation of Structural Models." Working Paper.

[43]. Tamer, E. (2003). "Coordination Games, Multiple Equilibria and the Timing of Radio Commercials." Working Paper.

[44]. H.P. Williams. (1985). *Model Building in Mathematical Programming.* New York: John Wiley & Sons.

[45]. Zeng and Lin (2008). "Efficient resampling methods for nonsmooth estimating functions." *Biostatistics* 9 (2): 355–363.