



AUSTRALIAN
SCHOOL OF BUSINESS™

THE UNIVERSITY OF NEW SOUTH WALES

www.business.unsw.edu.au

**School of Economics
Australian School of Business
UNSW Sydney NSW 2052 Australia**

<http://www.economics.unsw.edu.au>

**Estimating Sequential-move Games
by a Recursive Conditioning Simulator**

Shiko Maruyama

School of Economics Discussion Paper: 2009/01

The views expressed in this paper are those of the authors and do not necessarily reflect those of the School of Economic at UNSW.

ISSN 1323-8949

ISBN 978 0 7334 2729-9

Estimating Sequential-Move Games by a Recursive Conditioning Simulator

Shiko Maruyama*

PRELIMINARY DRAFT. COMMENTS ARE WELCOME.

February 23, 2009

Abstract

Sequential decision-making is a noticeable feature of strategic interactions among agents. The full estimation of sequential games, however, has been challenging due to the sheer computational burden, especially when the game is large and asymmetric.

In this paper, I propose an estimation method for discrete choice sequential games that is computationally feasible, easy-to-implement, and efficient, by modifying the Geweke-Hajivassiliou-Keane (GHK) simulator, the most widely used probit simulator.

I show that the recursive nature of the GHK simulator is easily dovetailed with the sequential structure of strategic interactions.

*This paper is based on a chapter of my Ph.D. dissertation at Northwestern University. I am thankful to Leemore Dafny, David Dranove, Michael Mazzeo, Aviv Nevo, and Robert Porter for their guidance. School of Economics, University of New South Wales. E-mail: s.maruyama@unsw.edu.au

1 Introduction

The structural estimation of discrete-choice non-cooperative games has rapidly developed since the seminal works by Bjorn and Vuong (1984) and Bresnahan and Reiss (1991). In these models, an econometrician studies the observed discrete decisions of agents (e.g. whether to enter a particular market), assuming latent payoffs of agents and taking strategic interactions into consideration. By imposing a certain game structure, unobservable structural parameters, such as sunk entry costs, are estimated. The econometrician can test strategic behavior predicted by theory, such as preemptive behavior. Counterfactual simulations can be conducted to evaluate the effect of strategic behavior. These analyses, however, are relevant only if the assumed game structure is valid, at least as an approximation.

The most widely studied in this literature is the class of incomplete information simultaneous-move games.¹ Complete information simultaneous-move games, on the other hand, have seen much fewer studies, mainly due to its computational complexity. Recent examples are Bajari, Hong, and Ryan (2008) and Soetevent and Kooreman (2007). Furthermore, research on sequential-move games has been quite limited. In the industrial organization literature, few examples include Berry (1992), Mazzeo (2002), Schmidt-Dengler (2006), and Maruyama (2008). Consequently, there exists very little empirical discussion on whether a researcher should employ simultaneous games or sequential games. The literature of the labor participation of couples also has studies of sequential move games, or Stackelberg games in this

¹There are numerous recent papers on both methodology and empirical application. Much of recent progress occurs on dynamic games. Among others, see Aguirregabiria and Mira (2007), Bajari, Benkard, and Levin (2007), and Su and Judd (2008). For static games, Bajari et al. (2007) is an example.

case (e.g. Kooreman (1994) and Hiedman (1998)). These studies directly compare sequential games with simultaneous or cooperative games, but focus only on simple games with two players. The estimation of the general class of sequential games has suffered from its computational complication.

In this paper, I propose an estimation method for the general class of discrete-choice perfect information sequential games. Relying on the literature of simulation-assisted estimation, the proposed method is efficient, computationally practical, and easy-to-use. By facilitating the estimation of sequential games and, especially, enabling the estimation of large and asymmetric sequential games, this method makes the sequential game a practical option for researchers and reinforces grounds for the selection of the game specification, which is often ad hoc in the current literature.

Moreover, the estimation of sequential games is not only another similar option, but has some distinctive features and advantages over simultaneous games. First, the sequential nature of decision-making among agents is quite often a key aspect in describing real-world strategic interactions. A cartel breaks down following a price cut by a firm. Manufacturers utilize reverse engineering. A monopolist's preemptive behavior deters entry in the market. Not only firms play sequential games. Konrad *et al.* (2002) empirically show that, in a family with parents and two children in Germany, the elder child often chooses to move farther away from the parents so that he or she induces the younger sibling to take care of the parents. Empirical analysis of sequential games is also a developing field in political science and international relations (e.g. Bas *et al.* (2008), and Signorino and Tarar (2006)).

Second, while the exact decision order of agents is rarely observed, a researcher can estimate sequential games imposing different decision order assumptions including randomized orders. In this way, a decision order can be tested. In contrast, the simultaneous-move assumption is an untestable *a priori* assumption, unless in a quite simple game setting.

Third, perfect information sequential games have technical advantages and disadvantages, which are also practically important factors in choosing the game specification. In case of simultaneous games, the assumption of imperfect information often makes the computation much simpler. However, this assumption complicates and limits the inclusion of agent characteristics that are unobserved to researchers but observed to the other agents, which may be crucial factors in the real-world rivalry. Next, whether complete information or incomplete information, simultaneous games need to handle the existence of an equilibrium and the possible multiplicity of equilibria, while perfect information sequential games can utilize the notion of subgame perfection, which guarantees the existence of a unique equilibrium.² While this is an advantage of sequential games, the recent literature has offered various techniques for simultaneous games to handle the multiple equilibrium issue (Ciliberto and Tamer (2007), Sweeting (2005), and Pakes *et al.* (2006)). Nevertheless, which equilibrium occurs is still difficult to rationalize. Sequential games do not have this ambiguity.

On the other hand, the estimation of sequential games is computationally demanding, because it involves multidimensional integrals, which occur because an equilibrium outcome observed to an econometrician has multiple random terms due to multiple agents and mul-

²Theoretically multiple equilibria are possible in tie break cases, but this can be ignored in empirical application with continuous latent variables.

multiple alternatives (as well as multiple time points when a panel data set is employed). This difficulty is analogous to the random-utility multivariate probit model, in which the latent utility of each alternative has a normal random component. In the probit literature, simulation techniques are widely used to numerically approximate multidimensional integrals, but the estimation of sequential-move games is complicated by interdependence among strategies of agents.

Previous studies circumvent or alleviate multidimensional integrals by assuming some degree of symmetry among agents, making assumptions that simplify the game structure, and/or not using identity information of each agent in estimating parameters (Berry (1992), Mazzeo (2002), and Schmidt-Dengler (2006)). However, when a researcher focuses on interactions among heterogeneous agents or when a researcher hopes to capitalize on his or her rich data set that contains detailed agent characteristics, the game becomes a large asymmetric game, making multidimensional integrals more formidable to overcome. This is one of the reasons why the empirical literature of sequential games has been limited compared to its voluminous theoretical counterpart and the well-developed empirical literature of simultaneous games.

To overcome the issue of multidimensional integrals, the literature of limited dependent variable models utilizes the simulation-assisted estimation method. For multinomial probit models, the most widely used simulation technique by far is the Geweke-Hajivassiliou-Keane (GHK) simulator. It is called GHK, after Geweke (1989, 1991), Hajivassiliou (as reported in Hajivassiliou and McFadden, 1998), and Keane (1990, 1994). This simulator is compu-

tationally tractable; the simulated probability is bounded away 0 and 1 and is smooth with respect to parameters. A number of studies have compared numerous probit simulators, and have confirmed its usefulness and relative accuracy (Hajivassiliou, McFadden, and Ruud (1996), Geweke, Keane, and Runkle (1994), Börsch-Supan and Hajivassiliou (1993)). In standard multinomial probit models, each alternative may have "interactions" with others through the variance-covariance matrix of multivariate normal error terms, which the GHK simulator fully incorporates. In the context of discrete games, however, if the estimation of strategic effects is a researcher's main concern, the interaction must be modeled outside the error components based on the structure of the payoff function and equilibrium solution framework, which is beyond scope of the GHK simulator.

This paper develops a modified version of the GHK simulator, "sequential GHK", showing that the recursive nature of the GHK simulator can be dovetailed with perfect information sequential games. The algorithm of GHK relies on the decomposition of a multivariate normal distribution, $f(v_1, \dots, v_J) = f(v_1)f(v_2|v_1)\dots f(v_{J-1}|v_{J-2}, \dots, v_1)f(v_J|v_{J-1}, \dots, v_1)$, and draws recursively from truncated univariate normal distributions. In a sequential-move game, an equilibrium is solved by using backward induction based on its sequential nature. Given that the order of simulation draws in GHK is arbitrary, the GHK simulator can easily fit the sequential game. The modification this paper proposes does not affect the nature of the simulator; the asymptotic properties are the same as the regular GHK simulator. In contrast to the two-step estimation methods that are becoming common in the estimation of discrete games of incomplete information (Aguirregabiria and Mira (2007) and Bajari *et al.* (2007)),

the method presented here is a direct full estimation of the game, so information loss is minimal.³ Estimation is easily accommodated in the standard frameworks of the method of simulated likelihood or the method of simulated scores.

Exploiting the GHK simulator, the method proposed in this paper is efficient, computationally powerful, and easy-to-use. Specifically, sequential games played by many agents with many characteristics variables can be estimated, and counterfactual simulations can be straightforwardly conducted. Maruyama (2008) applies this method to the U.S. health insurance market in Medicare to simulate the effects of subsidy on social welfare. In the estimation of its entry model, where, in each county, at most 16 heterogeneous firms play a sequential-move game with structural profit functions, the computational burden is not found to be a significant problem.

After formally presenting the setup in the next section, the use of the GHK simulator is explained in Section 3. Although the proposed method is intuitive and practical, it requires several *a priori* assumptions in the setup. The prerequisite decision order assumption is a particular example. In Section 4, I clarify the relevance and limitations of the assumptions of the model used in the preceding sections and then discuss possible extensions. To show how the proposed method outperforms the previous methods in the literature, a Monte Carlo simulation is conducted in Section 5. Section 6 concludes. Some practical guidelines for actual estimation are provided in the Appendix.

³The method discussed in this paper is not directly comparable to the two-step estimation methods in terms of efficiency, because these methods have been developed for games of incomplete information, while this paper focuses on perfect-information games.

2 The Setup

I present the use of the "sequential GHK," a modified GHK simulator in the setting of a simplistic static binary-choice game. Possible extensions and applications of the model to different settings are discussed in the next section.

Suppose there are N independent repetitions of games, for example, markets ($i = 1, \dots, N$). In each market, a sequential-move binary-choice game is played by heterogeneous agents ($j = 1, \dots, J_i$). The results below naturally generalize to the case where the number of alternatives is greater than two and the case where the number of agents varies across games. I assume a perfect-information game. There is no private information and all preceding decisions are known to everyone. A subgame perfect pure strategy Nash equilibrium (SPNE) in a market is achieved when every agent expects no gain from individually deviating from its equilibrium strategy in its every subgame. In perfect-information sequential-move games, there always exists a unique SPNE.⁴ The equilibrium can be obtained by backward induction, i.e., in a game tree, by deciding the optimal strategies from the most downstream agent to the most upstream agent. An econometrician observes the decisions agents make and *some* agent characteristics. The payoffs are not observable. The econometrician's task is to draw an inference about the parametrically assumed payoff function.

I introduce some notation to simplify the formal presentation. For a vector of indexes $(1, \dots, J)$, the notation " $< j$ " denotes the subvector $(1, \dots, j-1)$, " $\leq j$ " denotes the subvector $(1, \dots, j)$, and " $-j$ " denotes the subvector that excludes component j . Thus, for a vector

⁴Tie break cases occur with probability measure zero, so they are ignored.

ε , $\varepsilon_{<j}$ is the subvector of the first $j - 1$ components, and ε_{-j} is the subvector excluding component j . For a matrix L , $L_{j,<j}$ denotes a vector containing the first $j - 1$ elements of row j , and $L_{-j,-j}$ denotes the subarray excluding row j and column j . When I use a boldface symbol, it denotes a vector, for example, $\boldsymbol{\varepsilon}_i = (\varepsilon_{1i}, \varepsilon_{2i}, \dots, \varepsilon_{J_i i})$. The subscript for each market, i , is dropped for simplicity when no ambiguity would arise.

In a sequential-move game, agent j 's decision in a market is represented by y_j . To simplify the presentation, an agent's choice is assumed to be a binary choice — "move" or "do not move" — and so y_j is an indicator variable that takes one if the agent moves. Without loss of generality, the order of subscripts for agents $(1, 2, \dots, J_i)$ comprises the reverse of the decision order in market i — agent J_i makes a decision first and agent 1 at the end. In the game tree, agent 1 is at the very bottom. Some agent characteristics are observable to the econometrician and are denoted $\mathbf{x}_j = (x_{1j}, \dots, x_{Kj})$. \mathbf{x}_j may also contain observed market characteristics. A $J_i \times K$ matrix, \mathbf{X}_i , summarizes the observed characteristics of market i and the agents in market i . To capture the other unobservable factors that affect an agent's payoff and decision, a single-indexed error component, $\varepsilon_{\iota j i}$, is introduced in the payoff function for each decision of agent j . The payoff of agent j from choosing $y_j = \iota$ is, after dropping index i :

$$\pi_j(y_j = \iota; \mathbf{X}, \varepsilon_{\iota j}, y_{-j}; \gamma) \equiv \bar{\pi}(y_j = \iota; \mathbf{X}, y_{-j}; \gamma) + \varepsilon_{\iota j}. \quad (1)$$

Here, $\bar{\pi}$ is an assumed parametric function of payoffs with parameters to be estimated, γ .

A subgame perfect pure strategy Nash equilibrium (SPNE) is obtained when every player

expects no gain from individually deviating from the equilibrium in *every subgame*. Formally, an SPNE solution, (y^e) , is any solution for the decisions of the agents that satisfies:

$$\pi_j(y_j = 1; \mathbf{X}, \varepsilon_{1j}, y_{-j}^e; \gamma) - \pi_j(y_j = 0; \mathbf{X}, \varepsilon_{0j}, (y_{>j}^e, y_{<j}^*(y_{>j}^e, y_j = 0)); \gamma) \geq 0, \text{ if agent } j \text{ moves in } y^e \quad (2)$$

$$\pi_j(y_j = 1; \mathbf{X}, \varepsilon_{1j}, (y_{>j}^e, y_{<j}^*(y_{>j}^e, y_j = 1)); \gamma) - \pi_j(y_j = 0; \mathbf{X}, \varepsilon_{0j}, y_{-j}^e; \gamma) < 0, \text{ if agent } j \text{ does not move in } y^e \quad (3)$$

for all $j = 1, \dots, J$,

where $y_{\leq j}^*(y_{>j})$ is the unique SPNE solution for the subgame that starts from agent j given the decisions of the preceding agents, $y_{>j}$. This is recursively defined and $y_{\leq 1}^*(y_{>1})$ is agent 1's best response to the given decisions of the other agents, $y_{>1}$.⁵ The equilibrium solution can be calculated by the backward induction algorithm for any given parameters γ , observed characteristics, \mathbf{X} , and unobservable heterogeneity, ε . Thus, the SPNE solution function is denoted as:

$$y^*(\mathbf{X}, \varepsilon; \gamma) \equiv \{y^e \text{ with given } \mathbf{X}, \varepsilon, \gamma\}.$$

In estimation of binary games, I need only one error component for each agent, which governs the relative difference in agent j 's payoff between the two options, because, in general, the payoff from both options are unobservable and what one can infer from observed decisions

⁵ y_j in this paper denotes not a strategy profile but the revealed decision of agent j . In the SPNE concept, a strategy is defined at every decision node of a game tree, so here an agent's strategy cannot be expressed as a binary variable except for agent J . An agent's deviation from y^* holding the others' decisions fixed does not necessarily mean that all the others make the same decision; it means the others follow the same *strategy*.

concerns only the relative comparison of two payoffs. Hence I require normalization, such as $\varepsilon_{2ji} = 0$, without loss of generality, and let $\varepsilon_{ji} \equiv \varepsilon_{1ji}$. It may be likely that ε_{1ji} and ε_{2ji} are correlated, but the unobserved components that affect a firm's profit in the same way can be ignored in comparing the two payoffs.

I assume a normal distribution for ε_{ji} as a requirement for using the GHK simulator. However, the variance structure *within* a market can be very flexible. To demonstrate this point, I add a market-specific effect as a simple extension and the component unobserved to the econometrician is thus specified as:

$$\varepsilon_{ji} = \omega\eta_{ji} + \rho\eta_i. \tag{4}$$

η_{ji} and η_i are assumed to be independent of \mathbf{X}_i , and are independently and identically distributed in standard normal distribution across firms and markets, respectively. η_i measures market-specific factors that make choosing "move" more attractive for all agents in market i . This specification is only for demonstration and not crucial for the use of the following estimation method. In the literature of probit models, applications of the GHK simulator cover more flexible error structures, such as the multi-period Probit, multivariate rank ordered Probit, multi-period Tobit, and so on. In most cases, the payoffs are modeled scale-free.⁶ Unless the level of payoffs is identified, ω and ρ are not separately identified and thus normalization is necessary, for example $\omega^2 + \rho^2 = 1$, so $\omega = \sqrt{1 - \rho^2}$. Normalization

⁶An exception is Maruyama (2008), in which profit functions are modeled such that the level of variable profits is identified from demand-side estimation.

does not affect the argument below.

The estimation relies on the maximum likelihood estimation method. Each game (e.g. each market) is the unit for which the individual likelihood is defined. Denote the observed market configuration as y_i^o . The log-likelihood function can be written as

$$\hat{\theta}_{ML} = \arg \max_{\theta} \left\{ \frac{1}{N} \sum_i^N \ln \Pr [y_i^o = y_i^*(\mathbf{X}_i; \theta)] \right\}, \quad (5)$$

where θ is the vector of model parameters, (γ, ρ) .

However, the probability in the likelihood does not have an analytical solution due to multidimensional integrals. Unless the dimension of the unobservable factors, namely the number of agents and alternative options in a market, is very small, numerical approximation, such as quadrature methods, is impractical. Following the literature, I discuss the use of the maximum simulated likelihood (MSL) method.⁷

3 Applying the GHK Simulator

3.1 Maximum Simulated Likelihood with the Crude Frequency Simulator

The most straightforward simulator for MSL is the crude frequency simulator, also called the accept-reject simulator and first proposed by Lerman and Manski (1981). The simulator

⁷Method of simulated scores (MSS) is another option, because calculating the derivative of the likelihood is analytically straightforward. MSS may improve efficiency by removing simulation bias that results from the logarithm in the log likelihood function (Hajivassiliou and McFadden (1998)).

they explore for discrete choice models is:

$$\widehat{\theta}_{LM} = \arg \max_{\theta} \left\{ \frac{1}{N} \sum_i^N \ln \frac{1}{R} \sum_r \tilde{l}_{ir} \right\} \quad (6)$$

$$\equiv \arg \max_{\theta} \left\{ \frac{1}{N} \sum_i^N \ln \frac{1}{R} \sum_r I[y_i^o = y_i^*(\mathbf{X}_i, \boldsymbol{\varepsilon}_i^r; \gamma)] \right\}, \quad (7)$$

where the simulation procedure takes R sets of random draws from the assumed distribution. In the framework of this paper, for each simulation draw r , a unique simulated equilibrium market configuration y_i^{r*} is generated for each market i based on data for the market, given payoff function parameters, γ , and the values of the random draws from a multivariate normal distribution, $\boldsymbol{\varepsilon}_i^r$.

Although this simulator provides estimates that are consistent with R and N , the probability estimate constructed by this simulator is a discontinuous function of the parameters and is not bounded away from 0 and 1. The use of the indicator function makes its variance quite large. Due to these problems, Lerman and Manski find that their estimator requires a very large number of simulations for satisfactory performance. Moreover, the discontinuity of the likelihood function requires an optimization method that does not require differentiability of the optimand, such as the nonlinear simplex algorithm of Nelder and Mead (1964).

However, the use of backward induction requires the estimation procedure to compare payoffs at every decision node every time the procedure evaluates the likelihood. Thus, a likelihood evaluation is very expensive and simple discontinuous simulators that require many simulation draws are practically infeasible in estimating large asymmetric games.

3.2 The Modified GHK Simulator

The sequential GHK, the simulator I propose, is an extension of the Geweke-Hajivassiliou-Keane (GHK) simulator. The GHK is a smooth recursive conditioning simulator and is useful in many cases where the log-likelihood function involves high dimensional integrals with the multivariate normal distribution. The GHK algorithm draws recursively from truncated univariate normal distributions, and relies on the decomposition:

$$f(v_1, \dots, v_J) = f(v_1)f(v_2|v_1)\dots f(v_{J-1}|v_{J-2}, \dots, v_1)f(v_J|v_{J-1}, \dots, v_1)$$

along with the fact that the conditional normal density can be written as a univariate normal distribution. The GHK simulator produces probability estimates that are bounded away from 0 and 1. The estimates are continuous and differentiable with respect to parameters, because each contribution is continuous and differentiable. It is also an unbiased estimator of individual likelihood, $l(\gamma, \rho; y_i^o, \mathbf{X}_i)$. It has a smaller variance than the crude frequency simulator, because each element is bounded away from 0 and 1. Currently, the GHK simulator appears to be the most accurate simulator available for a given computation time and the most widely used probit simulator. The discussion below explains how this powerful simulator works and how to apply this simulator to the case in which strategic interactions exist in a sequential game.

The GHK simulator relies on the Cholesky triangular decomposition to decompose the multivariate normal distribution into a set of univariate normal distributions. The multi-

variate normal disturbance vector defined above, $\boldsymbol{\varepsilon}_i$, can be rewritten as:

$$\boldsymbol{\varepsilon}_i = \Gamma_i \boldsymbol{\eta}_i$$

where $\boldsymbol{\eta}_i$ is a $(J_i + 1) \times 1$ vector of independent standard normal variables,

$$\boldsymbol{\eta}_i \sim N(0, I_{J_i+1})$$

and Γ_i is a $J_i \times (J_i + 1)$ parametric array,⁸

$$\Gamma_i = \begin{bmatrix} \omega & & 0 & \rho \\ & \ddots & & \rho \\ & & \ddots & \vdots \\ 0 & & \omega & \rho \end{bmatrix}.$$

Thus, $\boldsymbol{\varepsilon}_i$ can be rewritten as:

$$\boldsymbol{\varepsilon}_i \sim N(0, \Omega_i),$$

⁸More flexible variance structures can be addressed by changing Γ_i and the size of $\boldsymbol{\eta}_i$.

where Ω_i is the following positive definite matrix:

$$\begin{aligned} \Omega_i &\equiv \Gamma_i \Gamma_i' \\ &= \begin{bmatrix} \omega^2 + \rho^2 & & & \rho^2 \\ & \ddots & & \\ & & \ddots & \\ \rho^2 & & & \omega^2 + \rho^2 \end{bmatrix} = \begin{bmatrix} 1 & & & \rho^2 \\ & \ddots & & \\ & & \ddots & \\ \rho^2 & & & 1 \end{bmatrix}. \end{aligned}$$

It follows that, using the Cholesky decomposition, $\boldsymbol{\varepsilon}_i$ can be written as:

$$\boldsymbol{\varepsilon}_i = L(\Omega_i) \cdot \mathbf{v}_i, \quad (8)$$

where $L(\Omega)$ is the lower-triangular Cholesky factor of Ω , or $LL' = \Omega$, and \mathbf{v}_i is another multivariate standard normal vector:

$$\mathbf{v}_i \sim N(0, I_{J_i}).$$

The individual likelihood can be written, after dropping index i , as:

$$\begin{aligned} l(\theta; y^o, \mathbf{X}) &= \Pr [y^o = y^*(\mathbf{X}, \boldsymbol{\varepsilon}; \gamma, \rho)] \\ &= \int_{y^o = y^*(\mathbf{X}, \boldsymbol{\varepsilon}; \gamma)} n(\boldsymbol{\varepsilon}, \Omega) d\boldsymbol{\varepsilon}, \end{aligned}$$

where $n(\cdot)$ is the probability density function of the multivariate normal distribution. This

expression involves multiple integrals, which are difficult to compute. The general objective here is to obtain random draws from the distribution of ε_i subject to $y^o = y^*(\mathbf{X}, \varepsilon; \gamma, \rho)$. To do so, first rewrite the probability expression that explicitly expresses the rectangle in which the event, $y^o = y^*(\mathbf{X}, \varepsilon; \gamma, \rho)$, occurs:

$$\Pr [y^o = y^*(\mathbf{X}, \varepsilon; \gamma, \rho)] = \Pr \left[\begin{array}{l} \text{for } \forall j, \pi_j (y_j = 1; \mathbf{X}, \varepsilon_j, (y'_{>j}, y^*_{<j}(y'_{>j}, 1)'); \gamma) \\ \quad - \pi_j (y_j = 0; \mathbf{X}, \varepsilon_j, (y'_{>j}, y^*_{<j}(y'_{>j}, 0)'); \gamma) \end{array} \left\{ \begin{array}{l} > 0 \text{ if } y_j^o = 1 \\ \leq 0 \text{ if } y_j^o = 0 \end{array} \right. \right]. \quad (9)$$

Recall the form of the payoff function, (1).

$$\begin{aligned} \pi_j(y_j = 1; \mathbf{X}, \varepsilon_{ij}, y_{-j}; \gamma) &= \bar{\pi}(y_j = 1; y_{-j}, \mathbf{X}; \gamma) + \varepsilon_j \\ \pi_j(y_j = 0; \mathbf{X}, \varepsilon_{ij}, y_{-j}; \gamma) &= \bar{\pi}(y_j = 0; y_{-j}, \mathbf{X}; \gamma) \end{aligned}$$

The condition for agent j to satisfy $y^o = y^*(\mathbf{X}, \varepsilon; \gamma, \rho)$ is:

$$\begin{aligned} \text{if } y_j^o = 1, \quad \varepsilon_j &> \bar{\pi}(y_j = 1; y_{-j}^o, \mathbf{X}; \gamma) - \bar{\pi}(y_j = 0; (y'_{>j}, y^*_{<j}(y'_{>j}, 0)'), \mathbf{X}; \gamma) \equiv a_j^* \\ \text{if } y_j^o = 0, \quad \varepsilon_j &\leq \bar{\pi}(y_j = 1; (y'_{>j}, y^*_{<j}(y'_{>j}, 1)'), \mathbf{X}; \gamma) - \bar{\pi}(y_j = 0; y_{-j}^o, \mathbf{X}; \gamma) \equiv b_j^* \end{aligned}$$

By defining:

$$\left\{ \begin{array}{l} a_j = a_j^*, b_j = \infty \text{ if } y_j^o = 1 \\ a_j = -\infty, b_j = b_j^* \text{ if } y_j^o = 0 \end{array} \right\},$$

(9) can be rewritten as:

$$\Pr [y^o = y^*(\mathbf{X}; \gamma, \rho)] = \Pr[\text{for } \forall j, a_j(y^o, \mathbf{X}, \varepsilon_{<j}; \gamma) \leq \varepsilon_j \leq b_j(y^o, \mathbf{X}, \varepsilon_{<j}; \gamma)].$$

This expression shows us the rectangle in which the event, $y^o = y^*(\mathbf{X}; \gamma, \rho)$, occurs. The essential difference from the regular multivariate probit case, which involves no strategic interaction, is that the size of the rectangle, which is determined by the values of a and b , depends on the values of ε as well. However, it does not depend on the error terms of all other agents. The idea behind the use of the GHK simulator is that, to obtain the interval of ε_j , we only need $\varepsilon_{<j}$, because the decisions of the upstream agents are given for agent j . When agent j makes a decision, only the random components of the downstream agents are relevant to predicting the downstream responses to each of agent j 's options. By using the Cholesky decomposition, (8), this equation becomes:

$$\begin{aligned} \Pr [y^o = y^*(\mathbf{X}; \gamma, \rho)] = & \hspace{15em} (10) \\ \Pr[\text{for } \forall j, a_j(y^o, \mathbf{X}, \mathbf{v}_{<j}; \gamma, \Omega) \leq [L(\Omega) \cdot \mathbf{v}]_j \leq b_j(y^o, \mathbf{X}, \mathbf{v}_{<j}; \gamma, \Omega)] \end{aligned}$$

Or:

$$\Pr [y^o = y^*(\mathbf{X}; \gamma, \rho)] = \int_{\text{for } \forall j, a_j(y^o, \mathbf{X}, \mathbf{v}_{<j}; \gamma, \Omega) \leq [L(\Omega) \cdot \mathbf{v}]_j \leq b_j(y^o, \mathbf{X}, \mathbf{v}_{<j}; \gamma, \Omega)} \left[\prod_{j=1}^J \phi(v_j) \right] d\mathbf{v},$$

where $\phi()$ is the probability density function of the standard normal distribution.

Now we are ready to apply the GHK simulator. For each simulation, prepare a vector of independent uniform $(0, 1)$ random variables, (u_1, \dots, u_{J_i}) . Define the following function:

$$q(u, a, b) \equiv \Phi^{-1}(\Phi(a) \cdot (1 - u) + \Phi(b) \cdot u), \quad (11)$$

where $0 < u < 1$ and $-\infty \leq a < b \leq \infty$.

This function, $q(\cdot)$, is a mapping that takes a uniform $(0, 1)$ random variable into a truncated standard normal random variate on the interval $[a, b]$.

The property of sequential games that is exploited here is that, when making a decision, an agent only considers the responses of its downstream agents. For each agent, the decisions of its preceding agents are all given. From agent 1's perspective, there is no strategic interaction. For given $y^o, \mathbf{X}, \mathbf{u}, \gamma, L$, recursively define a sequence of simulated v_j so as to satisfy $y^o = y^*(\mathbf{X}, \boldsymbol{\varepsilon}; \gamma, \rho)$ for $j = 1, \dots, J$:

$$\begin{aligned} \tilde{v}_1 &\equiv q\left(u_1, \frac{a_1}{L_{11}}, \frac{b_1}{L_{11}}\right) \\ \tilde{v}_2 &\equiv q\left(u_2, \frac{a_2(\tilde{v}_1) - L_{2,1}\tilde{v}_1}{L_{22}}, \frac{b_2(\tilde{v}_1) - L_{2,1}\tilde{v}_1}{L_{22}}\right) \\ &\dots \\ \tilde{v}_J &\equiv q\left(u_J, \frac{a_J(\tilde{v}_{<J-1}) - L_{J,1}\tilde{v}_1 \dots - L_{J,J-1}\tilde{v}_{J-1}}{L_{JJ}}, \frac{b_J(\tilde{v}_{<J-1}) - L_{J,1}\tilde{v}_1 \dots - L_{J,J-1}\tilde{v}_{J-1}}{L_{JJ}}\right) \end{aligned}$$

After obtaining the simulated $\tilde{\mathbf{v}}$, the probability interval for \mathbf{v} to satisfy $y^o = y^*(\mathbf{X}, \boldsymbol{\varepsilon}; \gamma, \rho)$

is calculated as:

$$\begin{aligned}
Q_1 &\equiv \Pr\left(\frac{a_1}{L_{11}} \leq v_1 \leq \frac{b_1}{L_{11}}\right) \\
Q_2 &\equiv \Pr\left(\frac{a_2(\tilde{v}_1) - L_{2,1}\tilde{v}_1}{L_{22}} \leq v_2 \leq \frac{b_2(\tilde{v}_1) - L_{2,1}\tilde{v}_1}{L_{22}}\right) \\
&\dots \\
Q_J &\equiv \Pr\left(\frac{a_J(\tilde{v}_{<J-1}) - L_{J,1}\tilde{v}_1 \dots - L_{J,J-1}\tilde{v}_{J-1}}{L_{JJ}} \leq v_J \leq \frac{b_J(\tilde{v}_{<J-1}) - L_{J,1}\tilde{v}_1 \dots - L_{J,J-1}\tilde{v}_{J-1}}{L_{JJ}}\right).
\end{aligned}$$

Given a, b, L , and \tilde{v} , each Q_j is a straightforward calculation. For example:

$$Q_1 = \Phi\left(\frac{b_1}{L_{11}}\right) - \Phi\left(\frac{a_1}{L_{11}}\right).$$

Repeat this simulation R times and define the likelihood contribution simulator as:

$$\tilde{l}(\gamma, \Omega; y^o, \mathbf{X}; R, \mathbf{u}) \equiv \frac{1}{R} \sum_{r=1}^R \prod_{j=1}^J Q_j(\tilde{v}_{1r}, \dots, \tilde{v}_{J-1,r}). \quad (12)$$

The model can be estimated by solving the following maximum simulated likelihood problem:

$$\begin{aligned}
\hat{\theta}_{MSL} &= \arg \max_{\theta} \left\{ \frac{1}{N} \sum_i^N \ln \tilde{l}(\gamma, \Omega; y_i^o, \mathbf{X}_i; R, \mathbf{u}_i) \right\} \\
&= \arg \max_{\theta} \left\{ \frac{1}{N} \sum_i^N \ln \frac{1}{R} \sum_{r=1}^R \prod_{j=1}^J Q_j(\tilde{v}_{1r}, \dots, \tilde{v}_{J-1,r}) \right\}.
\end{aligned} \quad (13)$$

The literature has developed a general consensus that the GHK simulator is accurate

in approximating discrete choice probabilities, especially when considering the low computational effort required. The combination of the recursive conditioning approach and the smooth univariate truncated variate generation algorithm produces an unbiased (for any value of R) multivariate probability simulator of $l(\cdot)$ that is smooth, i.e., a continuous and differentiable function of the model parameters. A key intuition behind this excellent feature is that the Cholesky triangularization underlying the GHK method implies an importance-sampling distribution that, while computationally extremely tractable, provides an excellent approximation to the true correlation structure of the unobservable factors. Most of the computational effort involved for this simulator comes from drawing the univariate truncated normal variates. This effort is approximately linear in K , the dimension of the probability integral. This is a remarkable advantage of this simulator. For the detailed performance and properties of this simulator, see, for example, Börsch-Supan and Hajivassiliou (1993) and Hajivassiliou and McFadden (1998).

The GHK simulator has so far been used in applications with no explicit strategic interactions. This is because the original GHK simulator can deal with interactions across j through the disturbance structure, but not strategic interactions across j , which is clearly a limitation to address strategic interactions in general, because not only an agent's random component but also his decision affect the decisions of the others. For example, Chernew, Gowrisankaran, and Fendrick (2002) use the GHK simulator in their entry model of hospitals, but strategic interactions are not taken into account in the empirical model specification. By exploiting the nature of sequential-move games, that an agent only speculates about the

decisions of the downstream agents, the discussion in this section showed that the GHK simulator can be extended to the cases with strategic interactions without losing any virtues of the simulator.

4 Discussion

4.1 The Relevance of the Sequential-Move Setup

The sequential-move game approach presented in this paper is restrictive in the sense that it explicitly requires an exogenous decision order. A much more commonly used setup in empirical studies of discrete games is the incomplete-information simultaneous-move game. In principle, a researcher should choose a game specification according to the environment in which the data are generated. The choice is obvious for some cases. For example, some experiments or well-designed auctions may stipulate a clear decision order. Such cases, however, are rare. In most cases, both simultaneous games and sequential games are merely ad-hoc approximations and the reality might lie somewhere in between.

In theory, the simultaneous approach is more general in the sense that any sequential-move game can be expressed as a simultaneous game. However, constructing and estimating such simultaneous-move games is not practical, because the size of the choice set of each agent rapidly grows as the number of players or alternatives increases. Thus, in practice, the interactions captured by the two approaches are different.

In choosing a specification, the following should be considered. Background details of

the game and data, such as institutional specifics, behavioral models and frequencies of observed points in time (aggregation of decision processes over time) may facilitate the choice. Computational feasibility is also practically important. For large asymmetric games of perfect information, the method proposed in this paper offers significant computational advantages. The subgame perfect equilibrium concept precludes existence and multiplicity issues and the use of the sequential GHK in a one-step direct estimation simplifies the estimation procedure and provides efficient estimates in a reasonable amount of time.

Furthermore, explicit treatment of the sequential nature of interactions is a noticeable feature of this approach. In application, quantifying first mover's advantage is possible. The decision order assumption may be ad hoc, but the computational advantages allow researchers to test different orders, including a randomized order. In this way, researchers can infer whether the sequential nature is relevant to their particular subject, and, if relevant, which decision order fits the data. This is another advantage of the proposed approach.

4.2 Extensions

The framework I present in this paper can be used in more general settings. Increasing the number of alternatives available to each agent is a straightforward extension, at least conceptually. A general approach is to assign an additional error term for each additional alternative. Within the decision turn of each agent, the order of simulation draws for each of his alternatives is arbitrary, as the same as the regular GHK simulator. However, this may make the estimation time-consuming, because adding another alternative to each agent's

decision considerably increases the number of required draws. The same applies for panel data. Even if there is no inter-temporal interaction across an agent's decisions over time, the existence of serial correlation in the error components considerably increases the number of necessary draws, though the same method can still be used.

Relaxing the assumption of perfect information is possible within the scope of this paper, as long as (1) the distribution of each agent's private information is common knowledge and (2) each agent's payoff depends only on his own realized random component and the decisions of the other agents. This incomplete-information game is conceptually simpler, because each agent predicts the downstream moves by using expected values and the agent's decision does not depend on the realization of random components in downstream. This implies that the proposed sequential GHK becomes the regular GHK. Computation may or may not become more cumbersome. While the lack of sequential dependency through random components simplifies the calculation, the backward calculation of expected payoffs of each agent is necessary. Incorporating mixed strategy is computationally impractical and beyond the scope of the proposed approach.

Can the proposed method be applied to dynamic games? The analysis of dynamic decision-making with strategic interactions is a rapidly growing area of research (Aguirregabiria and Mira (2007) and Bajari, Benkard, and Levin (2007)). The estimation of dynamic games is essential in analyzing aspects that the static approach cannot provide identification for. For example, a dynamic game is used to recover the costs of investment or entry in an oligopolistic market. Most of this literature employs incomplete-information simultaneous-

move games.

While the estimation method in this paper captures a certain aspect of dynamic decision-making or inter-temporal dependence, the empirical approach is quite similar to the classical static entry analysis, where a cross-sectional data set is used and each agent makes a decision only once. However, there are several ways to analyze a dynamic feature of multi-period data. First, when the choice of agents is assumed to be a one-time irreversible decision and the order of such possible irreversible decisions is exogenously given, the framework in this paper can be applied to a multi-period data set. This is the idea of Schmidt-Dengler (2006), the work that analyzes the timing of adoption of MRIs in the hospital industry and employed the perfect-information sequential-move game framework. In this setting, the empirical question can be viewed as when each agent makes a move after the preceding agent's move, and the model is estimated using a panel data set. The second type of possible application utilizes the fact that a sequential-move game can be viewed as a series of one single agent's repetitive decision-making. There is not interaction among agents but sequential inter-temporal dependence, i.e. an agent's payoff in each single period is partly determined by his decision in the previous periods. Third, more generally, repetitive decision-making by multiple agents over multiple periods can be modeled in the same manner. This extension does not alter the nature of the framework, but, because the computational burden becomes significant, the number of time points and/or agents has to be small. In all these potential applications, the key requirements are: (1) the order of sequence is exogenous and (2) there is no uncertainty from the agent's perspective.

5 Monte Carlo Simulation

*** TBW

6 Concluding Remarks

In this paper I develop a computationally practical and efficient estimation method for empirical models of discrete-choice sequential-move games. Based on a simulation assisted estimation approach, I propose the sequential GHK, an extension of the GHK simulator. This method allows researchers to empirically study strategic interactions in a large asymmetric game. Specifically, researchers can infer the existence and degree of strategic complementarity and perform counterfactual simulations that explicitly take strategic interactions into account. The decision sequence that best fits the data can also be investigated. The method can be applied to other fields. In industrial organization, the possibilities span entry, product choice, auction, advertisement, investment, and so on. Games played in household settings, politics, and international relations can also be studied with this framework.

Appendix: Computing Miscellanea

Because the modification of the GHK simulator proposed in this paper does not affect the continuity and differentiability of the GHK, standard maximization routines can be used in estimation. When taking random draws, variance reduction techniques, such as antithetics or the Halton procedure, can be used as usual.

Some computing techniques are helpful in speeding up the computation. The unnecessary part of the calculation of the backward induction algorithm can be skipped by using assumptions of payoff functions that are made by economic theory. For example, Maruyama (2007) exploits the non-increasing property of the profit function in the number of entering rival firms and decreases the computation time by more than 95 percent. From the programming perspective, another useful technique is the use of the recursive function command.⁹ When the number of agents varies across observations, this technique is extremely helpful in efficiently executing the backward induction algorithm.

References

- [1] Aguirregabiria, Victor and Pedro Mira (2007), "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75(1): 1-53.
- [2] Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin (2007), "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75(5): 1331-1370.
- [3] Bajari, Patrick, Han Hong, John Krainer, and Denis Nekipelov (2007), "Estimating Static Models of Strategic Interactions," mimeo.
- [4] Bajari, Patrick, Han Hong, and Stephen P. Ryan (2008), "Identification and Estimation of a Discrete Game of Complete Information," mimeo.

⁹I find this function available in FORTRAN90. In solving for equilibrium using backward induction, I define a recursive function at each node of the game tree that uses the results from one level below as arguments.

- [5] Bas, Muhammet Ali, Curtis S. Signorino, and Robert W. Walker (2008), "Statistical Backwards Induction: A Simple Method for Estimating Recursive Strategic Models," *Political Analysis*, 16(1): 21-40.
- [6] Berry, Steven T. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4): 889-917.
- [7] Börsch-Supan, A. and V. Hajivassiliou (1993), "Smooth Unbiased Multivariate Probability Simulation for Maximum Likelihood Estimation of Limited Dependent Variable Models," *Journal of Econometrics*, 58: 347-368.
- [8] Bresnahan, Timothy F. and Peter C. Reiss (1991), "Empirical-Models of Discrete Games," *Journal of Econometrics*, 48(1-2): 57-81.
- [9] Chernew, Michael, Gautam Gowrisankaran, and A. Mark Fendrick (2002), "Payer Type and the Returns to Bypass Surgery: Evidence from Hospital Entry Behavior," *Journal of Health Economics*, 21: 451-474.
- [10] Ciliberto, Federico and Elie Tamer (2007), "Market Structure and Multiple Equilibria in Airline Markets," mimeo.
- [11] Geweke, J. (1989), "Bayesian Inference in Econometric Models Using Monte Carlo Integration," *Econometrica*, 57(): 1317-1339.
- [12] Geweke, J. (1991), "Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to Linear Constraints," in E. M. Keramidas, ed., *Computer Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface*,

pp. 571-578. Fairfax: Interface Foundation of North America, Inc.

- [13] Geweke, J., M. Keane, and D. Runkle (1994), "Alternative Computational Approaches to Inference in the Multinomial Probit Model," *Review of Economics and Statistics*, 76: 609-632.
- [14] Hajivassiliou, V. and D. McFadden (1998), "The Method of Simulated Scores for the Estimation of LDV Models," *Econometrica*, 66(): 863-896.
- [15] Hajivassiliou, V., D. McFadden, and P. Ruud (1996), "Simulation of Multivariate Normal Rectangle Probabilities and Their Derivatives: Theoretical and Computational Results," *Journal of Econometrics*, 72: 85-134.
- [16] Heckman, J. J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46: 931-60.
- [17] Hiedeman, Bridget (1998), "A Stackelberg Model of Social Security Acceptance Decisions in Dual-Career Households," *Journal of Economic Behavior & Organization*, 34: 263-78.
- [18] Keane, M. (1990), "Four Essays in Empirical Macro and Labor Economics," PhD Thesis, Brown University.
- [19] Keane, M. (1994), "A Computationally Practical Simulation Estimator for Panel Data," *Econometrica*, 62(1): 95-116.
- [20] Konrad, Kai A., Harald Künemund, Kjell Erik Lommerud, and Julio R. Robledo (2002), "Geography of the Family," *American Economic Review*, 92(4): 981-998.

- [21] Kooreman, Peter (1994), "Estimation of Econometric Models of Some Discrete Games," *Journal of Applied Econometrics*, 9: 255-268.
- [22] Lerman, S. and C. Manski (1981), "On the Use of Simulated Frequencies to Approximate Choice Probabilities," *Structural Analysis of Discrete Data with Econometric Applications*, edited by C. Manski and D. McFadden, Cambridge, MA: MIT Press:305-319.
- [23] Maruyama, Shiko (2007), "Welfare Analysis Incorporating a Structural Entry-Exit Model: A Case Study of Medicare HMOs," Ph.D. Dissertation, Northwestern University.
- [24] Maruyama, Shiko (2008), "Measuring the Welfare Effect of Entry in Differentiated Product Markets: The Case of Medicare HMOs," *School of Economics Discussion Paper*, 2008/01, University of New South Wales.
- [25] Mazzeo, Michael J. (2002), "Product Choice and Oligopoly Market Structure," *Rand Journal of Economics*, 33(2): 221-242.
- [26] McFadden, Daniel (1981), "Econometric Models of Probabilistic Choice," *Structural Analysis of Discrete Data with Econometric Applications*, edited by Charles F. Manski and Daniel McFadden, Cambridge, Mass.: MIT Press.
- [27] Nelder, J.A. and R. Mead (1964), "A Simplex Method for Function Minimization," *Computer Journal*, 7(4): 308-313.
- [28] Pakes, A., J. Porter, Kate Ho, and Joy Ishii (2006), "Moment Inequalities and Their

Application," Harvard Working Paper.

- [29] Schmidt-Dengler, Philipp (2006), "The Timing of New Technology Adoption: The Case of MRI," mimeo.
- [30] Signorino, Curtis S. and Ahmer R. Tarar (2006), "A Unified Theory and Test of Extended Immediate Deterrence," *American Journal of Political Science*, 50: 586-605.
- [31] Soetevent, Adriaan R. and Peter Kooreman (2007), "A Discrete-Choice Model with Social Interactions: With an Application to High School Teen Behavior," *Journal of Applied Econometrics*, 22: 599-624.
- [32] Su, Che-Lin and Kenneth Judd (2008), "Constrained Optimization Approaches to Estimation of Structural Models," mimeo.
- [33] Sweeting, A. (2004), "Coordination Games, Multiple Equilibria, and the Timing of Radio Commercials," mimeo.