Dynamic Discrete Choice Models with Lagged Social Interactions: with an Application to a Signaling Game Experiment

By Xiaodong Liu, Lung-fei Lee and John Kagel*
Department of Economics, Ohio State University
Columbus, OH 43210

June, 2006

Abstract

This article generalizes Heckman's (1981) dynamic discrete choice panel data models by introducing lagged social interactions, so that the models can accommodate interrelationships of decisions across cross-sectional units. The likelihood function for a general model with dynamic social interactions is derived and simulation methods based on the unbiased GHK simulator are proposed to implement the maximum likelihood estimation. We apply the Markov and Polya models with lagged social interactions to investigate the equilibrium adjustment process in laboratory experiments based on an entry limit pricing game. We find that subjects' decisions are influenced by the past decisions of their peers. Hence the imitation of peers' strategies plays an important role in the learning process of strategic play.

Key Words: Discrete Choice, Dynamic Models, Simulated Maximum Likelihood, Social Interactions, Entry Limit Pricing Games, Experimental Economics.

JEL Classification: C15, C35, C72, C90

---

*Lee appreciates financial support from NSF under grant number SES-0519204. Kagel's research has been supported by NSF grants SES-0136928 and SES-0452911. Any opinions, findings, and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We are responsible for any errors.

1 Introduction

In his seminal work, Heckman (1981) has introduced a rich group of discrete choice stochastic processes that allow each cross-sectional unit's decisions to have complex dynamic economic interrelationships over time. In this article, we generalize the dynamic discrete choice panel data models by introducing time-lagged social interactions, so that the models can accommodate interrelationships of decisions, such as learning from peers, across cross-sectional units. This enriches the class of dynamics in Heckman (1981). As interactions across cross-sectional units carry out with a time lag, the models are well-defined without running into identification or multiple equilibria problems, which occur in some social interaction models (Manski, 1993).

Likelihood functions of dynamic discrete choice models involve multiple integrals, if explanatory variables include lagged latent dependent variables or disturbances allow for serial correlation in addition to that captured by random components. For panel data models, the dimension of integration increases with the number of periods, which makes numerical implementation impractical. To overcome the computational difficulty, simulation estimation methods have been developed. The simulator due to Geweke (1991), Borsch-Supan and Hajivassiliou (1993) and Keane (1994) is known to be practical and accurate to implement the method of simulated maximum likelihood (SML), when the time periods are not too long.

In this paper, we show that the implementation of the Geweke-Hajivassiliou-Keane (GHK) simulator remains tractable for models with social interactions. We investigate the finite sample properties of simulated estimates for model parameters and the effects of misspecification of dynamic structures and disturbances on estimates in the Monte Carlo experiments. As the likelihood function is nonlinear, the SML estimator (SMLE) might have an asymptotic bias if the number of random draws to construct the likelihood simulator does not increase fast enough relative to the sample size. Hence special attention will be given to dominated finite sample bias (relative to standard error) of coefficient estimates due to simulation. We report some Monte Carlo results of a bias-correction procedure proposed by Lee (1995) for the estimation of dynamic models with lagged interactions.

These dynamic social interaction models may have broad applicability, in particular, for experimental economics data. Numerous experiments have been conducted with a discrete choice space, with observations obtained in consecutive rounds. One of the main concerns in experimental games is the effect of a player's learning from other players. As such a dynamic discrete choice model with lagged social interactions may fit well as a possible econometric model for the analysis of experimental data. Specifically, in this paper, we apply our generalized dynamic models with social interactions to investigate the presence and magnitude of peer group effects in experiments based on Milgrom and Roberts' (1982) entry limit pricing game. Similar peer group effects are likely to be present in a variety of experimental designs where subjects receive feedback on their peer's performance. Empirical findings reported here may have broader economic implications. From the statistical inference point of view, the usual limited number of experimental subjects, rounds, and sessions due to feasibility or expense concerns might prevent one from determining whether peer group effects are indeed negligible or overwhelmed by estimation errors caused by insufficient sample size. So our estimation and Monte Carlo experimental results may shed some light on the sample size requirement and sample structures favorable to successfully identifying potential peer group effects in discrete choice games.

The organization of this article is as follows. In Section 2, we introduce a general dynamic discrete choice panel data model with lagged social interactions, derive the likelihood function and illustrate the formulation of simulators and simulated likelihood function for this model. We report Monte Carlo results for the SMLE of the Markov and Polya models with lagged social interactions in Section 3. In Section 4, we formulate empirical dynamic models to investigate the adjustment process of subjects' decisions in laboratory experiments based on an entry limit pricing game. Section 5 briefly concludes.

2 General Dynamic Discrete Choice Models with Social Interactions and SML Estimation

Consider a general dynamic discrete choice panel data model with lagged social interactions

$$y_{it}^* = h_{it}(y_{i,t-1}^*, \cdots, y_{i,-\infty}^*, Y_{n,t-1}, \cdots, Y_{n,-\infty}, X_{nt}, \cdots, X_{n,-\infty}, \xi_i) + v_{it}, \qquad (1)$$

for $i = 1, \cdots, n$, where Y_{nt} is the n-dimensional vector of dichotomous indicators of the latent variables $y_{1t}^*, \cdots, y_{nt}^*$, X_{nt} is the $n \times k$-dimensional matrix of strictly exogenous variables and ξ_i is a random individual component. Suppose that the error components ξ_i are i.i.d. $N(0,\sigma^2)$ for all i and the disturbances v_{it} are i.i.d. $N(0,1)$ for all i and t. This process is assumed to start at $t=1$, and the initial conditions on y^*_{it}, Y_{nt} and X_{nt} for $t \leq 0$ are fixed outside the model and are assumed to be zero. The original specification of the dynamic model in Heckman (1981) does not incorporate lagged social interactions in that $y_{i,s-1}$ and x_{is} appear but not $Y_{n,s-1}$ and X_{ns} $(s \leq t)$. Depending on the specification of the function $h_{it}(\cdot)$ in terms of lagged observed or latent dependent variables, the Heckman discrete dynamic model is known to be sufficiently flexible to accommodate a wide variety of dynamic structures such as Markov models, Polya models, renewal processes, latent Markov models, with rich specifications on disturbances. It allows for unobserved heterogeneity across the n cross-sectional units and serial correlation for the remaining disturbances. The model with social interactions in (1) is generalized to incorporate additional dynamic effects due to peers' influence. We derive the likelihood function for (1) and construct the unbiased GHK simulator to implement the SML estimation for the model.

In addition to Y_{nt}, let $Y^*_{nt} = (y^*_{1t}, \cdots, y^*_{nt})'$ be the n-dimensional vector of the latent dependent variables for all the n cross-sectional units. Let X_t denote the sequence of $X_{nt}, X_{n,t-1}, \cdots$. Conditional on exogenous variables X_T and $\xi = (\xi_1, \cdots, \xi_n)'$, the joint density function of (Y^*_{nt}, Y_{nt}), $t = 1, \cdots, T$, is the product of conditional density of (Y^*_{ns}, Y_{ns}), $s = 1, \cdots, T$, over their past histories, i.e.,

$$\begin{aligned} & f(Y^*_{nT}, Y_{nT}, \cdots, Y^*_{n1}, Y_{n1}|X_T, \xi) \\ = & \left[\prod_{t=2}^{T} f(Y^*_{nt}, Y_{nt}|(Y^*_{ns}, Y_{ns}; s=1,\cdots,t-1), X_t, \xi)\right] f(Y^*_{n1}, Y_{n1}|X_1, \xi). \end{aligned}$$

Because v_{it} are mutually independent for $i = 1, \cdots, n$, each of the conditional densities of (Y^*_{nt}, Y_{nt}) can be further decomposed as

$$\begin{aligned} & f(Y^*_{nt}, Y_{nt}|(Y^*_{ns}, Y_{ns}; s=1,\cdots,t-1), X_t, \xi) \\ = & \prod_{i=1}^{n} f(y^*_{it}, y_{it}|(Y^*_{ns}, Y_{ns}; s=1,\cdots,t-1), X_t, \xi_i) \\ = & \prod_{i=1}^{n} I_{y_{it}}(y^*_{it}) g(y^*_{it}|(Y^*_{ns}, Y_{ns}; s=1,\cdots,t-1), X_t, \xi_i) \end{aligned}$$

for $t = 2, \cdots T$, and

$$f(Y_{n1}^*, Y_{n1}|X_1, \xi) = \prod_{i=1}^{n} I_{y_{i1}}(y_{i1}^*) g(y_{i1}^*|X_1, \xi_i),$$

where $I_{y_{it}}(y_{it}^*)$ is the dichotomous indicator with $I_{y_{it}}(y_{it}^*) = 1$ if the value y_{it}^* determines the observed value y_{it}; $I_{y_{it}}(y_{it}^*) = 0$, otherwise, and g is the conditional density of y_{it}^*. Therefore, the joint probability of $Y_{nT}, \cdots, Y_{n1}$ conditional on X_T and ξ is

$$\begin{aligned} & P(Y_{nT}, \cdots, Y_{n1}|X_T, \xi) \\ = & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(Y_{nT}^*, Y_{nT}, \cdots, Y_{n1}^*, Y_{n1}|X_T, \xi) dvec'(Y_{nT}^*) \cdots dvec'(Y_{n1}^*) \\ = & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{t=2}^{T} \prod_{i=1}^{n} I_{y_{it}}(y_{it}^*) g(y_{it}^*|(Y_{ns}^*, Y_{ns}; s = 1, \cdots, t-1), X_t, \xi_i) \right] \\ & \times \prod_{i=1}^{n} I_{y_{i1}}(y_{i1}^*) g(y_{i1}^*|X_1, \xi_i) dvec'(Y_{nT}^*) \cdots dvec'(Y_{n1}^*). \end{aligned} \tag{2}$$

For (1), $g(y_{it}^*|(Y_{ns}^*, Y_{ns}; s = 1, \cdots, t-1), X_t, \xi_i) = g(y_{it}^*|(y_{is}^*, Y_{ns}; s = 1, \cdots, t-1), X_t, \xi_i)$ as interactions among different units are going through the observed Y_{ns} and X_{ns} but not Y_{ns}^* with $s < t$. Hence we have

$$\begin{aligned} & P(Y_{nT}, \cdots, Y_{n1}|X_T, \xi) \\ = & \prod_{i=1}^{n} \Big\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{t=2}^{T} I_{y_{it}}(y_{it}^*) g(y_{it}^*|(y_{is}^*, Y_{ns}; s = 1, \cdots, t-1), X_t, \xi_i) \right] \\ & \times I_{y_{i1}}(y_{i1}^*) g(y_{i1}^*|X_1, \xi_i) dy_{iT}^* \cdots dy_{i1}^* \Big\}. \end{aligned}$$

Under the distributional assumption that v_{it} is $N(0, 1)$,

$$g(y_{it}^*|(y_{is}^*, Y_{ns}; s = 1, \cdots, t-1), X_t, \xi_i) = \phi(y_{it}^* - h_{it}),$$

where $h_{it} = h_{it}(y_{i,t-1}^*, \cdots, y_{i,-\infty}^*, Y_{n,t-1}, \cdots, Y_{n,-\infty}, X_t, \xi_i)$ for simplicity and ϕ is the standard normal density function. Define the integral limits L_{it} and U_{it}:

$$L_{it} = \begin{cases} -h_{it} & \text{if } y_{it} = 1, \\ -\infty & \text{if } y_{it} = 0, \end{cases} \text{ and } U_{it} = \begin{cases} \infty & \text{if } y_{it} = 1, \\ -h_{it} & \text{if } y_{it} = 0. \end{cases}$$

By transformations of variables, it follows that

$$
\begin{aligned}
& P(Y_{nT}, \cdots, Y_{n1}|X_T, \xi) \\
= & \prod_{i=1}^{n} \left\{ \int_{L_{i1}}^{U_{i1}} \cdots \int_{L_{i,T-1}}^{U_{i,T-1}} \left(\int_{L_{iT}}^{U_{iT}} \phi(v_{iT}) dv_{iT} \right) \phi(v_{i,T-1}) dv_{i,T-1} \cdots \phi(v_{i1}) dv_{i1} \right\} \\
= & \prod_{i=1}^{n} \Big\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\Phi(U_{iT}) - \Phi(L_{iT})) \\
& \times \prod_{s=1}^{T-1} (\Phi(U_{i,T-s}) - \Phi(L_{i,T-s}))\, \phi_{[L_{i,T-s},\ U_{i,T-s}]}(v_{i,T-s}) dv_{i,T-s} \Big\} \\
= & \prod_{i=1}^{n} \Big\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Phi((2y_{iT} - 1) h_{iT}) \\
& \times \prod_{s=1}^{T-1} \Phi((2y_{i,T-s} - 1) h_{i,T-s}) \phi_{[L_{i,T-s},\ U_{i,T-s}]}(v_{i,T-s}) dv_{i,T-s} \Big\},
\end{aligned}
$$

where $\phi_{[L_t,U_t]}$ is a truncated standard normal density function with support $[L_t, U_t]$. The probability $Y_{nT}, \cdots, Y_{n1}$ conditional on exogenous variables X_T is

$$
\begin{aligned}
& P(Y_{nT}, \cdots, Y_{n1}|X_T) \\
= & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P(Y_{nT}, \cdots, Y_{n1}|X_T, \xi_1, \cdots, \xi_n) \phi(\xi_1) \cdots \phi(\xi_n) d\xi_1 \cdots d\xi_n \\
= & \prod_{i=1}^{n} \Big\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Phi((2y_{iT} - 1) h_{iT}) \\
& \times [\prod_{s=1}^{T-1} \Phi((2y_{i,T-s} - 1) h_{i,T-s}) \phi_{[L_{i,T-s},\ U_{i,T-s}]}(v_{i,T-s}) dv_{i,T-s}] \phi(\xi_i) d\xi_i \Big\}.
\end{aligned}
$$

This likelihood suggests that the GHK simulator can be recursively applied to construct a simulated likelihood. Generate u_{it} ($i = 1, \cdots, n$; $t = 1, \cdots, T-1$) independent uniform $[0, 1]$ random variables. Generate ξ_i ($i = 1, \cdots, n$) independent standard normal variables. With initial conditions given, the random variables v_{it} ($i = 1, \cdots, n$; $t = 1, \cdots, T-1$) can be generated from the following steps. For each i, from $t = 1$ to $T-1$:

(1) Compute

$$v_{it} = -(2y_{it} - 1)\Phi^{-1} \left[u_{it} \Phi \left((2y_{it} - 1) h_{it} \right) \right].$$

(2) Generate the latent dependent variable

$$y_{it}^{*} = h_{it} + v_{it}.$$

With m independent simulation runs, the corresponding simulated log likelihood function is

$$\mathcal{L} = \sum_{i=1}^{n} \ln \left\{ \frac{1}{m} \sum_{j=1}^{m} \prod_{t=1}^{T} \Phi((2y_{it} - 1)h_{it}^{(j)}) \right\}, \tag{3}$$

where $h_{it}^{(j)} = h_{it}(y_{i,t-1}^{*(j)}, \cdots, y_{i0}^{*(j)}, Y_{n,t-1}, \cdots, Y_{n0}, X_t, \xi_i^{(j)})$, and the superscript (j) denotes an independent simulation run. Thus, the simulation of the likelihood for the model in (1), is similar to one of the conventional dynamic panel models in Lee (1997).

Asymptotic properties of the SMLE for cross-sectional or short time series panel data have been studied in Hajivassiliou and McFadden (1990), Lee (1992; 1995) and Gourieroux and Monfort (1993), among others. The SMLE can be asymptotically efficient when m increases at a rate faster than $n^{1/2}$. However, when m increases at a rate of $n^{1/2}$, as shown in Lee (1995), an asymptotic bias exists in the limiting distribution. The asymptotic bias will dominate the variance when m increases at a rate slower than $n^{1/2}$. Lee (1995) has suggested a simple bias-correction procedure to remove the leading bias term due to simulation. The asymptotic efficiency of the bias-adjusted estimator requires only that m goes to infinity at a rate faster than $n^{1/4}$.

For experimental economics, subjects are usually divided into several independent groups (experimental sessions), and games are played in several rounds within each group. Suppose that there are G groups. Within each group, there are n players and the number of rounds is T. With data from such a design, the simulated likelihood function shall be

$$\mathcal{L} = \sum_{g=1}^{G} \sum_{i=1}^{n} \ln \left\{ \frac{1}{m} \sum_{j=1}^{m} \prod_{t=1}^{T} \Phi((2y_{g,it} - 1)h_{g,it}^{(j)}) \right\}, \tag{4}$$

where the subscript (g, it) indicates the observation is from individual i of group g at round t.

The model in (1) can be further generalized to allow social interactions in both observed and latent lagged dependent variables,

$$y_{it}^* = \overline{h}_{it}(Y_{n,t-1}^*, \cdots, Y_{n,-\infty}^*, Y_{n,t-1}, \cdots, Y_{n,-\infty}, X_t, \xi_i) + v_{it}, \tag{5}$$

where Y_{nt}^* is the n-dimensional vector of latent dependent variables and X_t is the sequence of strictly exogenous variables $X_{nt}, X_{n,t-1}, \cdots$. As in (1), ξ_i are i.i.d. $N(0, \sigma^2)$ for all i and v_{it} are i.i.d. $N(0,1)$ for all i and t. The initial values for Y_{nt}^*, Y_{nt} and X_t for $t \leq 0$ are assumed to be zero. From (2) , the joint probability for $Y_{nT}, \cdots, Y_{n1}$ conditional on exogenous variables X_T and ξ is given by

$$\begin{aligned} & P(Y_{nT}, \cdots, Y_{n1}|X_T, \xi) \\ = & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{t=2}^{T} \prod_{i=1}^{n} I_{y_{it}}(y_{it}^*) g(y_{it}^* | (Y_{ns}^*, Y_{ns}; s = 1, \cdots, t-1), X_t, \xi_i) \right] \\ & \times \prod_{i=1}^{n} I_{y_{i1}}(y_{i1}^*) g(y_{i1}^* | X_1, \xi_i) dvec'(Y_{nT}^*) \cdots dvec'(Y_{n1}^*). \end{aligned}$$

Under the distributional assumption of v_{it},

$$g(y_{it}^* | (Y_{ns}^*, Y_{ns}; s = 1, \cdots, t-1), X_t, \xi_i) = \phi(y_{it}^* - \overline{h}_{it}),$$

where $\overline{h}_{it} = \overline{h}_{it}(Y_{i,t-1}^*, \cdots, Y_{i,-\infty}^*, Y_{n,t-1}, \cdots, Y_{n,-\infty}, X_t, \xi_i)$ for simplicity and ϕ is the standard normal density function. Define the integral limits $\overline{L}_{it}$ and $\overline{U}_{it}$:

$$\overline{L}_{it} = \begin{cases} -\overline{h}_{it} & \text{if } y_{it} = 1, \\ -\infty & \text{if } y_{it} = 0, \end{cases} \text{ and } \overline{U}_{it} = \begin{cases} \infty & \text{if } y_{it} = 1, \\ -\overline{h}_{it} & \text{if } y_{it} = 0. \end{cases}$$

By transformations of variables, it follows that

$$\begin{aligned} & P(Y_{nT}, \cdots, Y_{n1}|X_T, \xi) \\ = & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\prod_{i=1}^{n} \Phi((2y_{iT} - 1)\overline{h}_{iT}) \right] \\ & \times \prod_{s=1}^{T-1} \left[\prod_{i=1}^{n} \Phi((2y_{i,T-s} - 1)\overline{h}_{i,T-s}) \phi_{[\overline{L}_{i,T-s},\ \overline{U}_{i,T-s}]}(v_{i,T-s}) dv_{i,T-s} \right]. \end{aligned}$$

And the probability $Y_{nT}, \cdots, Y_{n1}$ conditional on exogenous variables X_T is $P(Y_{nT}, \cdots, Y_{n1}|X_T) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P(Y_{nT}, \cdots, Y_{n1}|X_T, \xi) \left[\prod_{i=1}^{n} \phi(\xi_i) d\xi_i\right]$.

In this case, with u_{it} and ξ_i generated as before, the random variables v_{it} ($i = 1, \cdots, n$; $t = 1, \cdots, T-1$) can be generated from the following steps, from $t = 1$ to $T-1$:

(1) Compute for $i = 1, \cdots, n$

$$v_{it} = -(2y_{it} - 1)\Phi^{-1}\left[u_{it}\Phi\left((2y_{it} - 1)\overline{h}_{it}\right)\right].$$

(2) Generate the latent dependent variable

$$y_{it}^* = \overline{h}_{it} + v_{it}.$$

With m independent runs, the corresponding simulated log likelihood function shall be

$$\mathcal{L} = \ln\left\{\frac{1}{m}\sum_{j=1}^{m}\prod_{t=1}^{T}\prod_{i=1}^{n}\Phi((2y_{it} - 1)\overline{h}_{it}^{(j)})\right\}, \tag{6}$$

where $\overline{h}_{it}^{(j)} = \overline{h}_{it}(Y_{n,t-1}^{*(j)}, \cdots, Y_{n0}^{*(j)}, Y_{n,t-1}, \cdots, Y_{n0}, X_t, \xi_i^{(j)})$ for the jth simulation run.

There can be some numerical difficulties in implementing the SML estimation procedure if T and n are large, as the simulated log likelihood functions (3) and (6) involve the product consisting of many terms of small numbers that might be impossible to evaluate with computers without underflow errors. The problem is more severe in (6), where the simulated likelihood involves the product of cumulative probabilities of the entire history for all members in a group. Lee (2000) has suggested an algorithm that can overcome the numerical problem by interchanging the summation and product operators behind the logarithmic transformation. Here we illustrate this algorithm for (6). For simplicity, let $k = (t-1)n + i$, $i = 1, \cdots, n$ for each t with $t = 1, \cdots, T$, and rewrite (6) as

$$\mathcal{L} = \ln\left\{\frac{1}{m}\sum_{j=1}^{m}\prod_{k=1}^{T\times n}\Phi((2y_k - 1)\overline{h}_k^{(j)})\right\}.$$

Let $a_{kj} = \Phi((2y_k - 1)\overline{h}_k^{(j)})$ and let ω_{kj} be weights for $k \geq 1$, which can be computed recursively as

$$\omega_{kj} = a_{kj}\omega_{k-1,j}/\sum_{s=1}^{m}a_{ks}\omega_{k-1,s},$$

starting with $\omega_{0j} = 1/m$ for $j = 1, \cdots, m$. Then following Lee (2000), (6) can be rewritten as

$$\mathcal{L} = \ln\left\{\prod_{k=1}^{T\times n}\sum_{j=1}^{m} a_{kj}\omega_{kj}\right\} = \sum_{k=1}^{T\times n}\ln\left\{\sum_{j=1}^{m}\Phi((2y_k - 1)\overline{h}_k^{(j)})\omega_{kj}\right\}, \tag{7}$$

where the product of cumulative probabilities behind the logarithmic transformation is replaced by the weighted sum of cumulative probabilities.

Social interactions in latent lagged dependent variables are likely to appear if cross-sectional units are allowed to discuss their past preferences and choices. As we plan to apply the model to the estimation of data from lab experiments where, as is typically the case, subjects make independent decisions without communication, we focus on models that conform to (1) in the rest of this paper.

3 Some Monte Carlo Results on SMLEs

3.1 A Markov Model with Lagged Social Interactions

Suppose we have observations of G independent groups, with n subjects in each group. The Markov dynamic choice model for the Monte Carlo study in this section is

$$y_{it}^{*} = \beta x_{i,t-1} + \lambda_1 y_{i,t-1} + \lambda_2 z_{i,t-1} + \sigma\xi_i + \varepsilon_{it}, \tag{8}$$

where $z_{i,t-1} = \sum_{j=1,j\neq i}^{n} y_{j,t-1}/\left(n-1\right)$, $\varepsilon_{it} = \rho\varepsilon_{i,t-1} + v_{it}$, and ξ_i and v_{it} are i.i.d. $N\left(0,1\right)$. The group subscript g has been suppressed for simplicity. By replacing ε_{it} with $\rho(y_{i,t-1}^{*} - (\beta x_{i,t-2} + \lambda_1 y_{i,t-2} + \lambda_2 z_{i,t-2} + \sigma\xi_i)) + v_{it}$, i.e., by a quasi-difference transformation for (8), it is easy to see that (8) conforms to the general model (1).

The x_{it} are generated as $x_{it} = (1/\sqrt{2})r_{it} + \sqrt{6}s_i$ where r_{it} are independent truncated standard normal variables on $[-2, 2]$ and s_i is a uniform variable on $[-0.5, 0.5]$, so that the variance of x_{it} is about 1 and its correlation coefficient over time is about 0.5. This process of generating exogenous variables is to allow the exogenous variables to correlate over time. It is used for all the models in this article. The initial values of all variables for $t \leq 0$ are given as 0. Sample data are generated with $\beta = 1$, $\lambda_1 = 0.2$, $\lambda_2 = 0.4$, $\sigma^2 = 0.5$, and $\rho = 0.4$. The serial correlation of the total disturbance $\sigma\xi_i + \epsilon_{it}$ of two adjacent periods has a correlation coefficient about 0.6

and the fraction of variance due to the individual effect is about 0.3. The sample size is 200, with $G = 50$ and $n = 4$. We have experimented with small, moderate and large numbers of random draws, namely $m = 15$, $m = 50$ and $m = 100$, for the construction of the GHK simulator. The number of periods for the panel data varies from 8 to 30. For each case, the number of replications is 200. For each replication, in addition to random disturbances in the model, the set of exogenous variables is also redrawn. The maximization algorithm used is a conjugate gradient method. For all cases and replications reported here, the algorithm converges without running into numerical problems. The initial estimate of σ is set to 1, and the initial estimates of the other parameters are set to 0. We have also tried some other starting values, with which the algorithm converges to similar solutions.

Table 10 in Appendix C reports the empirical means (Means), standard deviations (SDs) and root mean square errors (RMSEs) for both the bias unadjusted SMLE and the bias-adjusted SMLE. For all panels with periods from 8 to 30, the bias unadjusted SMLEs of β are biased downward. There are upward biases in the SMLEs of λ_1 and downward biases in the SMLEs of λ_2, σ and ρ, so the dynamic effect can be over stated, but the lagged peer group effect and the serial correlation of disturbances can be underestimated. The magnitude of bias increases with panel length, as the dimension of integration and the total number of choice alternatives are proportional to the number of periods. On the other hand, SDs of all the SMLEs decrease as panels become longer, since longer panel data provide more sample information about the stochastic process. If periods are not too long, RMSEs decrease. Biases of estimates are all substantially reduced when the number of simulated random variables m increases from 15 to 50. By increasing m to 100, biases become rather small and RMSEs can further be reduced, but the time cost is double. The issue of selecting m in practice has been addressed by Lee (1997). For small m, bias correction is valuable. Although SDs of bias-adjusted estimates are slightly larger, RMSEs of bias-adjusted estimates are smaller in general. The additional CPU cost for bias correction is negligible. However, as biases of estimates, especially for longer panels, are relatively large to begin with in this model, larger m is desirable for better improvement.[1]

Table 11 reports Means, SDs and RMSEs for alternative group sizes. For a given

[1] Results for the bias-adjusted estimates are omitted in subsequent tables to save space. The bias correction procedure for all the models in this article reduces bias and RMSE. The improvement is comparable with the gains from the bias correction procedure reported in *Table 10*.

sample size $G \times n = 200$, biases, SDs and RMSEs of all the SMLEs increase when the group size n increases from 4 to 8 (by comparing results in *Tables 10* and *11*). As the group size becomes even larger, biases, SDs and RMSEs of the SMLEs of λ_1, λ_2 and ρ further increase, while the estimates of β and σ are not much affected. As such, other things equal, more sessions with fewer subjects are preferred to fewer sessions with more subjects in each session.

To illustrate effects of ignoring potential lagged social interactions on SMLEs, we report the restricted SMLEs under $\lambda_2 = 0$ in *Table 12*. When positive social interactions are ignored, the SMLEs of β, σ and ρ are biased downward, and the SMLEs of λ_1 are biased upward. The estimated values of λ_1 are more than double in magnitude and the estimated values of ρ are reduced almost by half, so true state dependence can be over stated but spurious state dependence can be underestimated.

Misspecified disturbances, in general, would cause parameter estimates to be inconsistent. We investigate effects of misspecification in disturbances by the following Monte Carlo experiments. First, we estimate the random component model with $\sigma\xi_i + v_{it}$, where v_{it} are serially uncorrelated, with the data samples generated by the model specified as in (8). For random component models, multivariate probability functions involve only single integrals, which can be effectively implemented by the Gaussian Quadrature method as suggested by Butler and Moffitt (1982). However, for the sake of easy comparison, here we report the SMLE of the random component model. The simulated log likelihood function for the random component model is

$$\sum_{i=1}^{G\times n} \ln \left\{ \frac{1}{m} \sum_{j=1}^{m} \prod_{t=1}^{T} \Phi \left[(2y_{it} - 1) \left(\beta x_{i,t-1} + \lambda_1 y_{i,t-1} + \lambda_2 z_{i,t-1} + \sigma \xi_i^{(j)} \right) \right] \right\}.$$

The SMLEs are reported in the upper block of *Table 13*. There are substantial downward biases in the SMLEs of β and λ_2 and upward biases in the SMLEs of λ_1. Biases are more severe for longer panels. Even with $m = 100$, the estimated values of λ_1 are three times larger than the true value; and the estimated magnitudes of λ_2 are reduced by 2/3. Hence, true state dependence tends to be overestimated and lagged social interactions tend to be underestimated when serial correlation in ϵ_{it} is ignored. Biases in the SMLEs of σ are not uniform. The lower block of *Table 13* reports the restricted SMLEs under $\sigma = 0$, i.e., random component ξ were ignored. With this error specification, serially correlated disturbances $\varepsilon_{it} = \rho\varepsilon_{i,t-1} + v_{it}$ capture all the spurious state dependence. Ignoring random individual component biases the SMLEs of β, λ_1 downward and λ_2, ρ upward. Biases in λ_1 and λ_2 are more severe for longer panels. The magnitudes of upward bias of λ_2 are not really large. The biases of ρ are upward by 50%. But the biases of λ_1 towards zero are relatively much more severe.

3.2 A Polya Model with Lagged Social Interactions

In the Polya model, the entire history of the dynamic process is relevant to current decision making. The Polya model with a depreciation factor δ is specified as follows[2]:

$$y_{it}^* = \beta x_{i,t-1} + \lambda_1 \sum_{s=1}^{t} \delta^{s-1} y_{i,t-s} + \frac{\lambda_2}{\sum_{s=1}^{t} \delta^{s-1}} \sum_{s=1}^{t} \delta^{s-1} z_{i,t-s} + \sigma \xi_i + \varepsilon_{it}, \qquad (9)$$

where $z_{i,t-s} = \sum_{j=1,j\neq i}^{n} y_{j,t-s} / (n-1)$ and $\varepsilon_{it} = \rho\varepsilon_{i,t-1} + v_{it}$ with ξ_i and v_{it} i.i.d. $N(0,1)$. The group subscript g has been suppressed for simplicity. The initial values of all variables for $t \leq 0$ are given as 0. Substitution of $\varepsilon_{it} = \rho(y_{i,t-1}^* - (\beta x_{i,t-2} + \lambda_1 \sum_{s=1}^{t-1} \delta^{s-1} y_{i,t-s-1} + \lambda_2 \sum_{s=1}^{t-1} \delta^{s-1} z_{i,t-s-1} / \sum_{s=1}^{t-1} \delta^{s-1} + \sigma\xi_i)) + v_{it}$ in (9) conforms it to the general model (1). For comparison purpose, the discount factor δ is assumed to be a known constant and is set at 0.7. Sample data are generated with $\beta = 1$, $\lambda_1 = 0.2$, $\lambda_2 = 0.4$, $\sigma^2 = 0.5$, and $\rho = 0.4$.

The SMLEs are reported in *Table 14*. There are some downward biases in the SMLEs of β, λ_2, σ and ρ and upward bias in λ_1. Compared to estimates of the Markov model in *Table 10*, λ_1 and ρ in the Polya model can be estimated more accurately. They not only have small biases but also have much smaller SDs, due to an apparently stronger state dependence property of the Polya model. On the other hand, since we specified lagged social interactions as a weighted average of the past history instead of a weighted sum, variation in this term is reduced. So with such specification, λ_2 in the Polya model is much more difficult to estimate than in the Markov model. For small m and long panels, biases in the SMLEs of λ_2 is quite severe. By increasing m, biases in the estimates of λ_2 can be substantially reduced. For $T = 8$ or 15, the biases are smaller with $m = 50$ or 100. By comparison with the Markov model, SDs and RMSEs of the estimates of λ_2 here are two times larger.

Monte Carlo experiments are also performed to investigate effects of misspecifi-

[2]Here we specify the lagged social interactions term as the (weighted) average for observed laggd choices of peers over the entire history, so that it is not affected by the number of total observations.

cation in dynamic structures on SMLEs. *Table 15* reports the SMLE of the Markov model with lagged social interactions when data samples are generated by the Polya model (9). The SMLEs of λ_1 and σ are biased upward. And the SMLEs of λ_2 and ρ are biased downward. Hence, when the Polya dynamic structures are misspecified to be Markov, the true state dependence and the serial correlation due to unobserved heterogeneity tends to be overestimated but the lagged social interactions and the serial correlation of the remaining disturbance tend to be underestimated. The SMLEs of β are not affected very much by dynamic misspecification and their biases are not large.

4 An Application: Estimating Peer Group Effects in Experiments on Signaling Games

There is a large volume of literature on measuring peer group effects in field settings, while little attention has been paid to evaluating the influence of peer group effects on subjects' performance in experiments. Measuring the peer group influence in experiments is important as it affects our understanding of the evolution of subjects' behavior over time. Ignoring peer group effects potentially confounds any "sophisticated" learning process (e.g. adaptive learning) where subjects update beliefs, with the less "sophisticated" social learning where subjects simply replicate the strategy generating a better outcome. Furthermore, experimental results across diverse subject pools are much less likely to be consistent in the presence of strong peer group effects, as subjects' performance depends on the overall performance of the experimental session they were in. This section adopts dynamic discrete choice models with lagged social interactions to investigate the presence and magnitude of peer group effects in experiments on signaling games.

Following Manski (1993), similar behavior of individuals belonging to the same reference group may be due to *endogenous effects*, wherein "the propensity of an individual to behave in some way varies with the behavior of the group"; *exogenous effects*, wherein "the propensity of an individual to behave in some way varies with the exogenous characteristics of the group"; and *correlated effects*, wherein "individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments". In experimental settings, exogenous effects and correlated effects can be controlled through recruiting procedures and careful experimental designs, while endogenous effects are relatively hard to control by experimenters. We focus on measuring endogenous peer group effects in experiments in this section.

The plan of this section is as follows. Subsection 4.1 presents the theoretical predictions of the model of entry limit pricing. Subsection 4.2 outlines the experimental procedures and provides a general description of the data. Subsection 4.3 develops the empirical econometric models and interprets the estimation results.

4.1 Theoretical Considerations

Milgrom and Roberts (1982) propose a model of entry limit pricing as follows. There are two firms, an incumbent (established monopolist M) and a potential entrant (E), in a two-stage market producing a homogeneous good. Nature decides M's cost of production along with the distribution of these costs. M's cost is his/her private information throughout the game, with the prior distribution of the cost being common knowledge. In the first stage, M chooses an output (or price) level. In the second stage, E chooses to enter or stay out in response to the observed output (or price) level. If entry occurs, Cournot duopoly profits are realized by both M and E. There is a predetermined opportunity cost to E for entering the market. If there is no entry, M receives the single period monopoly profit. Entry is profitable against M with high cost but not against M with low cost. M may have an incentive to limit pricing, which involves producing greater output (charging lower price) in the first stage than the single period profit maximizing level in order to make entry appear unattractive.

In this game, the information sets are defined by the realized costs of M and E (c_M and c_E) and a choice of Q (quantity) by M. A (pure) strategy for M is a map s from its possible cost levels into the possible choices of Q and a (pure) strategy for E is a map t from R^2 into $\{0,1\}$ giving its decision for each possible pair (c_E, Q), where 1 is interpreted as "enter" and 0 as "stay out". An equilibrium consists of a pair of strategies (s^*, t^*) and a pair of conjectures $\left(\overline{s}, \overline{t}\right)$ such that (i) M's pricing policy s^* is a best response to its conjectures $\overline{t}$ about E's entry rule, (ii) the strategy t^* is a best response for E to its conjecture $\overline{s}$, and (iii) the actual and conjectured strategies coincide (Milgrom and Roberts, 1982, p. 446).

With two cost levels (types) for M, namely, $\underline{c}_M < \overline{c}_M$, if $s^*(\underline{c}_M) = s^*(\overline{c}_M)$, an equilibrium is called *pooling*; and if $s^*(\underline{c}_M) \neq s^*(\overline{c}_M)$ the equilibrium is *separating*. Partial pooling is a mixed strategy equilibrium that $s^*(\underline{c}_M) \neq s^*(\overline{c}_M)$ with a certain probability. In a pooling equilibrium, E can infer nothing from observing Q and so enters if the expected profit is positive. In a separating equilibrium, the observation of Q allows the value of c_M to be inferred exactly (Milgrom and Roberts, 1982, pp. 447-448). Depending on the cost structure, its distribution, and the market demand function, pooling equilibria and/or separating equilibria can occur. One may consider limit pricing as the outcome of competition between the types of established firms, with high cost types attempting to mimic low cost ones and low cost firms attempting to distinguish themselves from the high costs ones. Whether a pooling or a separating equilibrium is established is a matter of whether it is the high or low cost type which is successful. This competition could be purely a conjectural one in the mind of the entrant (Milgrom and Roberts, 1982, pp. 449-450).

Milgrom and Roberts' model of entry limit pricing is investigated experimentally by Cooper, Garvin and Kagel (1997*a*; 1997*b*) and Cooper and Kagel (2003*a*; 2003*b*; 2004). In the experiments, M is either a high-cost type (M_H) or a low-cost type (M_L) with equal probability. The model is further simplified by adding the payoffs of the two stages together and providing the subjects with payoff tables. Payoff tables *1-3* are provided in the "quantity game" with M choosing over output levels (1-7).

Table 1: Payoffs for A Player (Incumbent)

A1 (High Cost)			A2 (Low Cost)		
Your Choice	X (In)	Y (Out)	X (In)	Y (Out)	Your Choice
1	150	426	250	542	1
2	168	444	276	568	2
3	150	426	330	606	3
4	132	408	352	628	4
5	56	182	334	610	5
6	−188	−38	316	592	6
7	−292	−126	213	486	7

Source: Cooper, Garvin and Kagel (1997*b*).

In any given play of the game, Es are either all high cost types (E_Hs; payoff table *2*) or all low cost types (E_Ls; payoff table *3*). With E_Hs there exist pure-

Table 2: Payoffs for High Cost B Player (Entrant)

Your Action Choice	A Player's Type: A1 (High Cost) Your Payoff	A Player's Type: A2 (Low Cost) Your Payoff	Expected Value[a]
X (In)	300	74	187
Y (Out)	250	250	250

[a] Based on prior distribution (50% M_H, 50% M_L) of M types.
Source: Cooper, Garvin and Kagel (1997*b*).

Table 3: Payoffs for Low Cost B Player (Entrant)

Your Action Choice	A Player's Type: A1 (High Cost) Your Payoff	A Player's Type: A2 (Low Cost) Your Payoff	Expected Value[a]
X (In)	500	200	350
Y (Out)	250	250	250

[a] Based on prior distribution (50% M_H, 50% M_L) of M types.
Source: Cooper, Garvin and Kagel (1997*b*).

strategy pooling equilibria at output levels 1-5. There also exist two pure-strategy separating equilibria, in which M_Hs always choose 2 and are always entered on, M_Ls always choose 6 or 7 and are never entered on. Among them, only pooling at 4 or 5, and separating with M_Ls choosing 6 survive Cho-Kreps' (1987) intuitive criteria for equilibrium refinement. With E_Ls no pure-strategy pooling equilibrium exists, while the two pure-strategy separating equilibria still exist. There also exist a number of mixed-strategy equilibria. One that is of particular relevance is the partial pooling equilibrium in which M_Ls always select 5 while M_Hs mix between 2 (with probability 0.80) and 5 (with probability 0.2), and Es always enter on output levels other than 5, enter on 5 with probability 0.11. In simulations using a stochastic fictitious play learning model, this partial pooling equilibrium emerges with high frequency in the presence of E_Ls (Cooper, Garvin and Kagel, 1997*b*). Further, in practice M_Ls choose 5 with relatively high frequency as a separating equilibrium emerges (especially early on) and there is very little entry in response to it (Cooper, Garvin and Kagel, 1997*b*).

The payoffs in the "price game"(shown in Appendix A), where subjects are choosing over price, are a linear transformation of payoff tables *1* and *3*, (with table presentation changed as well). Hence the price game is theoretically identical to the quantity game with analogue equilibrium predictions.

4.2 Experimental Procedures and Data

Detailed description of the experimental procedures are in Cooper, Garvin and Kagel (1997*b*). The following lists some elements that are especially noteworthy, as they will be taken into account when empirically modeling the game.

1. Each experimental session employed between 12 and 16 subjects who were randomly assigned to computer terminals. Sessions typically lasted 36 periods, with the number of periods announced in advance. Subjects switched roles after every six plays, with incumbents (M) becoming entrants (E) and vice versa. Ms' types are generated each play randomly.

2. Following each play of the game the outcomes from all pairings (Ms' choice, Es' choice, and Ms' type) were revealed to all subjects. This made learning across individuals feasible, and provides the basis for potential peer group effects.

3. Subjects were randomly paired with each other for each play of the game, and subject identification numbers were suppressed when the game results were revealed. Hence there was no opportunity for reputation effects to develop. Learning, to the extent that it occurred, had to be based on own outcomes and observations of peer's choices.

Experimental treatments are summarized in *Table 9* in Appendix B. The "Experienced Subjects" treatment recruited subjects who had participated in earlier experimental sessions with exactly the same parameter values. The treatment "Meaningful Context" uses natural language for the instructions, and was introduced to explore the effects of context on subjects' reasoning process in signaling games (Cooper and Kagel, 2003*a*). The treatment "Crossovers from the E_H to E_L game" employed subjects with experience in the quantity game with payoff tables *1* and *2* to play the quantity game with payoff tables *1* and *3*, and was devoted to investigating subjects' ability to generalize learning in one game to related games (Cooper and Kagel, 2003*b*; 2004).

4.3 Empirical Models and Estimation Results

According to payoff table *1*, with full information, output levels 2 and 4 are optimal for M_Hs and M_Ls respectively. Pooling equilibria at output levels 3-5 and (partial) separating equilibria with M_Ls selecting output levels 5-7 involve strategic behavior - limit pricing - as Ms produce above (or price below) full-information levels. A "gradual, history-dependent adjustment process", starting with Ms "at their myopia maxima, followed by an attempt to pool, and then (if no pooling equilibrium exists) separation", has been observed by Cooper, Garvin and Kagel (1997*b*). Here we adopt our dynamic discrete choice models with lagged social interactions to characterize the evolution of subjects' behavior in the experiment.

We consider the estimation of equations with two different samples: One from the experimental sessions with E_Hs (using payoff tables *1* and *2*) and the other from the sessions with E_Ls (using payoff tables *1* and *3*). With E_Hs, play reliably converges to a pure strategy pooling equilibrium in which M_Hs learn to imitate M_Ls. As such we model the learning dynamic of M_Hs in this situation, treating choices of output levels 3-5 as strategic play, $y = 1$, and $y = 0$ otherwise[3]. For games with E_Ls, pure strategy pooling equilibria no longer exist, and we focus on the strategic play M_Ls, with $y = 1$ if the M_L selects outputs level 5-7, and $y = 0$ otherwise[4].

4.3.1 A Markov Model with Lagged Social Interactions

One way to justify the Markov process is to assume that the subject's current decision only depends on his/her last decision and feedback from the previous period besides individual characteristics. We will relax this restrictive assumption and consider the estimation of a more general dynamic model later.

The M_Hs have incentives to limit price in games with E_Hs. In a generic experimental session with n M_Hs, we assume that y^*, the unobservable incentives for M_Hs to limit price, can be characterized by the Markov dynamic discrete choice model with lagged social interactions. By design, subjects are randomly assigned turns as an M_H in different plays of the game within an experimental session. As such, a decision pe-

[3] Note that high-level outputs $6, 7$ are strictly dominated by other outputs for M_Hs, according to payoff table *1*. Among the 4576 observations in the actual experimental sample with E_Hs, 7 choices of output 6 or 7 made by M_Hs are observed.

[4] We have also tried to estimate with alternative criterion for limit pricing. For example, we treated output levels $4, 5$ by M_Hs in the case with E_Hs, and output levels $6, 7$ by M_Ls in the case with E_Ls as limit pricing. The estimation results are similar to those reported here.

riod in which a subject plays as an M_H differs from a (consecutive) calendar period. However, all subjects at his/her τth turn as an M_H observe all the outcomes from all of the preceding calendar periods. We distinguish between his/her own experience as an M_H and that of the peer group information that informs his/her choices. For individual subject i, let $\tau = 1, \cdots, T_i$ be the number of periods in which he/she has played as an M_H. Corresponding to each turn τ, there is a calender period. Let $t_i(\tau)$ be the calendar period when the subject i plays as an M_H. The Markov dynamic discrete choice model with lagged social interactions for i can be specified as

$$y_{i\tau}^* = \alpha + x_{i,t_i(\tau)-1}\beta + \lambda_1 y_{i,t_i(\tau-1)} + \lambda_2 w_{in} Y_{n,t_i(\tau)-1} + \gamma \ln \tau + \sigma \xi_i + \varepsilon_{it_i(\tau)}, \qquad (10)$$

for $\tau = 1, 2, \cdots, T_i$, where $Y_{n,t_i(\tau)-1}$ is an n-dimensional vector with the ith element being $y_{i,t_i(\tau)-1}$, $i = 1, \cdots, n$ and w_{in} is an n row-normalized weighting vector for all M_H peers. We assume that $\varepsilon_{it_i(\tau)} = \rho\varepsilon_{i,t_i(\tau-1)} + v_{it_i(\tau)}$ and ξ_i and $v_{it_i(\tau)}$ are i.i.d.$N(0,1)$. Since the stochastic process started at the first sampling period in the experiment, the initial conditions on all variables for $t \leq 0$ are set to zero.

If the latent dependent variable $y_{i\tau}^* > 0$, incumbent i limits price in his/her τth turn as an M_H, and the corresponding observed dependent variable $y_{i\tau}$ is 1; $y_{i\tau}$ is 0 otherwise. Explanatory variables are on the right hand side of (10). α is a constant. $x_{i,t_i(\tau)-1}$ is the perceived entry rate differential, which is between output levels 3-5 and 1-2 in the case with E_Hs, and output levels 1-4 and 5-7 in the case with E_Ls[5]. In the case with E_Hs, let $d_{is}^L(IN)$ (respectively, $d_{is}^O(IN)$) be a dummy variable indicating that incumbent i chooses output level 3, 4 or 5 (output level 1 or 2) and is entered on in calendar period s. Let $d_{-is}^L(IN)$ ($d_{-is}^O(IN)$) be the number of times in calendar period s that Ms *other than* i choose output levels 3, 4 or 5 (output levels 1 or 2) and observe the response IN. Define $d_{is}^L(OUT)$, $d_{is}^O(OUT)$, $d_{-is}^L(OUT)$, and $d_{-is}^O(OUT)$ in an analogous manner, where OUT involves potential Es staying out. Denote the weight a player put on the experience of other players relative to his/her own in calculating entry rate differential by ω. The perceived entry rate differential is given

[5] The entry rates are calculated conditional on the output level selected, not the type of M which selects the output. As Es can not observe Ms' type when making decisions of entry, the entry rate calculated here can be used to approximate M's beliefs on Es' responses.

by

$$x_{i,t_i(\tau)-1} = \frac{d^O_{i,t_i(\tau)-1}(IN) + \omega d^O_{-i,t_i(\tau)-1}(IN)}{d^O_{i,t_i(\tau)-1} + \omega d^O_{-i,t_i(\tau)-1}} - \frac{d^L_{i,t_i(\tau)-1}(IN) + \omega d^L_{-i,t_i(\tau)-1}(IN)}{d^L_{i,t_i(\tau)-1} + \omega d^L_{-i,t_i(\tau)-1}},$$

where $d^j_{i,t_i(\tau)-1} = d^j_{i,t_i(\tau)-1}(IN) + d^j_{i,t_i(\tau)-1}(OUT)$ for $j = L, O$.[6] This term serves as a proxy for the unobservable beliefs of Ms. $y_{i,t_i(\tau-1)}$, the time-lagged observed dependent variable, is introduced to measure the true state dependence of incumbent i in the dynamic process. $w_{in}Y_{n,t_i(\tau)-1}$ captures the peer group effects in the experiment, namely the influence of other M_Hs' strategic play in the preceding calender period on i's current choice. Given the anonymous nature of experimental design, we assume that the weighting matrix W_n, where w_{in} is its ith row, is simply $[(1_n \cdot 1_n' - I_n)/(n-1)]$, so that $w_{in}Y_{n,t_i(\tau)-1} = \sum_{j=1,j\neq i}^{n} y_{j,t_i(\tau)-1}/(n-1)$. $\ln\tau$, where τ is the number of times incumbent i has played as an M_H (the current period included), stand for all other effects of experience within an experimental session that are not captured by the other explanatory variables. An individual random component ξ_i is introduced to control unobserved heterogeneity across players. The remaining disturbances are assumed to follow an AR(1) process. We model choice dynamics of M_Ls in experimental sessions with E_Ls in an analogous manner.

As we have shown, for incumbent i, the likelihood function involves $(T_i - 1)$-dimension integrals that are analytically intractable and numerically hard to implement. We circumvent this computational difficulty by the SML method based on the GHK simulator. *Table 4* reports the SMLEs for the Markov model based on a simulator generated from 100 random draws with samples from the experimental sessions with E_Hs and E_Ls respectively.[7]

The significantly positive SMLEs of λ_1 in all cases show that a player's current choice depends heavily on his/her choice in the previous round. That is, one round of strategic play substantially increases the likelihood of strategic play in future rounds. This indicates that subjects do not play strategically just by chance. Rather, once they have learned to play strategically, they are very likely to continue to play strategically, clear evidence of learning. Interaction terms are introduced to account for differences

[6] We assume that $(d^j_{i,t_i(\tau)-1}(IN) + \omega d^j_{-i,t_i(\tau)-1}(IN))/(d^j_{i,t_i(\tau)-1} + \omega d^j_{-i,t_i(\tau)-1}) = 0.5$, in the case that $d^j_{i,t_i(\tau)-1} + \omega d^j_{-i,t_i(\tau)-1} = 0$ $(j = L, O)$.

[7] We have tried to add more interaction terms, or remove some regressors or interaction terms with insignificant coefficients. The estimation results are trivially affected.

in behavior between experienced and inexperienced subjects (with the dummy variable NX representing sessions with inexperienced subjects).[8] The significant negative coefficient value for the interaction term of lagged choice and NX in games with E_Ls indicates that in this case inexperienced subjects were much less confident of their choice to play strategically than their experienced selves.

Learning can come about in one of two ways: (1) social learning in which case subjects simply replicate their peers' strategies and/or (2) adaptive learning that is independent of peers' choices. Estimating significant peer group effects in the Markov model is evidence in favor of social learning, while finding a significant coefficient for the entry rate differential would be evidence for adaptive learning independent of peers' choices.

The SMLE of λ_2 is positive and statistically significant in *Table 4*, in games with E_Ls, indicating the existence of endogenous peer group effects in this case. For the specification without interaction terms, the average marginal impact of the peer group effect on the probability of limit pricing given exogenous variables and lagged choices is 0.054.[9] In contrast, peer group effects in games with E_Hs are not statistically significant.

The SMLE of the coefficient β for entry rate differential is positive, statistically significant, and robust to alternative specifications in games with E_Ls. For the specification without interaction terms, the average marginal effect of entry rate differential on the probability of limit pricing is 0.032. Note that in this case M_Ls place primary weight on own entry rate differential, with very limited weight placed on others' experience, whereas fully rational Ms should put equal weight on own entry and entry

[8] The minus two times log likelihood ratios for testing jointly the significance of interactions terms in the Markov model are, respectively, 6.56 for games with E_Hs, and 13.2 for games with E_Ls. The latter is significant at the 5 percent level with an asymptotic $\chi^2(5)$ distribution.

[9] For the general model (1), $E(y_{it}|(y_{is}^*, Y_{ns}, X_{ns}, s = 1, \cdots, t-1), X_{nt}, \xi_i) = \Phi(h_{it})$. The average marginal effect over time and individual of, say X_{nt} (which is assumed continuous), on the transition probability $P(y_{it} = 1|(Y_{ns}, X_{ns}, s = 1, \cdots, t-1), X_{nt})$, is given by

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\int\cdots\int\phi(h_{it})(\partial h_{it}/\partial X_{nt}) \\ \times f\left(y_{i1}^*, \cdots, y_{i,t-1}^*, \xi_i|Y_{ns}, X_{ns}, s = 1, \cdots, t-1\right)dy_{i1}^*\cdots dy_{i,t-1}^* d\xi_i$$

The multiple integrals here can be approximated by simulations. We simulate $h_{it}^{(j)}$ following the same procedure as in (3). With m independent simulation runs, the corresponding (sample average) simulated marginal effects is $\left(\sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{t=1}^{T}\phi\left(h_{it}^{(j)}\right)(\partial h_{it}/\partial X_{nt})\right)/mnT$. Results reported in this paper are based on a simulator generated from 1000 random draws.

on other Ms. This also seems odd given that λ_2 is positive and statistically significant, indicating that subjects care about peers' past choices, while an insignificant ω indicates that subjects tend to ignore other Ms' experience with respect to entry. However, λ_2 captures social learning that replicates peers' strategies, while ω captures adaptive learning that updates beliefs based on other's experience. Given the sophisticated nature of adaptive learning compared to social learning, this result is not unreasonable.[10] In contrast to all of this the SMLE of β is not statistically significant in experiments with E_Hs. As will be reported in the next section, peer group effects are identifiable in the Polya model (our preferred specification) for games with E_Hs, but the entry rate differential continues to be statistically insignificant.

The proportion of M_Hs attempting to pool by choosing output levels 3 and 4 in the previous round is introduced as an explanatory variable in games with E_Ls because an increase in this proportion makes separation at output levels 5-7 more attractive. Although positive in sign, this variable fails to achieve statistical significance in any of the Markov specifications.

The positively significant estimates of the coefficient on $\ln \tau$ pick up other experience effects that fail to be captured in the Markov model. In experiments with E_Hs, the SMLE of its interaction effect with NX indicates that this positive impact is confined to experimental sessions employing inexperienced subjects.

The dummy variables for experienced players are large, positive and statistically significant in games with E_Ls and E_Hs indicating that in both cases experienced subjects start out with much higher levels of strategic play than inexperienced subjects. This is consistent with the strong history dependence of strategic play within inexperienced subject sessions. In games with E_Ls, dummies for experiments with crossovers are positive and statistically significant, which is consistent with the findings in Cooper and Kagel (2004) that there exists positive transfer of learning across related games. The significant negative estimates for the constants (α) are consistent with the slow emergence of strategic play in all cases. The larger absolute value for α in games with E_Ls captures the fact that strategic play is much slower to emerge in this case.

[10] We also consider alternative specifications where own lagged choices and peers' lagged choices are interacted with entrants' responses. This is discussed briefly in the next section where we report on the polya model specification, our preferred specification.

Table 4: SMLEs for the Markov Model (standard errors in parentheses)

	with High-Cost Type Entrants*		with Low-Cost Type Entrants**	
		w/ interactions		w/ interactions
entry rate differential (β)	0.109 (0.085)	0.179 (0.176)	0.318^c (0.069)	0.383^c (0.089)
entry rate differential $\times NX$	-	-0.151 (0.187)	-	-0.110 (0.101)
weight on others' experience (ω)	0.003 (0.254)	0.523 (2.536)	0.028 (0.061)	0.022 (0.062)
lagged choice (λ_1)	0.730^c (0.184)	0.725^c (0.240)	1.730^c (0.094)	1.961^c (0.113)
lagged choice $\times NX$	-	0.153 (0.205)	-	-0.417^c (0.119)
peer group effects (λ_2)	0.023 (0.125)	-0.109 (0.286)	0.541^c (0.106)	0.516^c (0.131)
peer group effects $\times NX$	-	0.146 (0.320)	-	0.011 (0.214)
% of M_H choosing 3,4	-	-	0.019 (0.089)	-0.053 (0.131)
% of M_H choosing 3,4 $\times NX$	-	-	-	0.171 (0.179)
experience within a session (γ)	0.256^c (0.083)	-0.019 (0.188)	0.099^b (0.045)	0.058 (0.067)
experience within a session $\times NX$	-	0.304^a (0.176)	-	0.039 (0.077)
constant (α)	-0.545^c (0.162)	-0.619^c (0.160)	-1.871^c (0.101)	-1.838^c (0.108)
random effects (σ)	1.164^c (0.103)	1.118^c (0.098)	0.888^c (0.062)	0.867^c (0.063)
serial correlation (ρ)	-0.080 (0.103)	-0.148 (0.100)	-0.244^c (0.048)	-0.236^c (0.052)
Dummies:				
w/ experience of the same game	0.702^c (0.127)	1.026^c (0.224)	0.827^c (0.073)	0.794^c (0.118)
meaningful context	-0.079 (0.168)	-0.078 (0.164)	-0.035 (0.093)	-0.035 (0.092)
crossovers from E_H to E_L games	-	-	0.640^c (0.108)	0.611^c (0.141)
Log Likelihood	-1089.88	-1086.60	-2106.82	-2100.22

NX is a dummy variable for experimental sessions employing subjects with no experience of the same or related games

* 266 subjects, 4576 observations; ** 568 subjects, 11536 observations

[a] significantly different from 0 at the 10 percent level

[b] significantly different from 0 at the 5 percent level

[c] significantly different from 0 at the 1 percent level

Though the overall correlation across the disturbances captured by $\sigma\xi_i + \varepsilon_{it_i(\tau)}$ is positive, the negative sign of ρ suggests the presence of some fluctuations not captured by the dynamic structure. Hence we generalize the Markov model to include more lagged terms in the next subsection.

4.3.2 A Polya Model with Lagged Social Interactions

In this section, we model the influence of all past plays on a subject's current decision by a Polya process with lagged social interactions. Similarly to the Markov model, we assume that the unobservable incentives to limit price can be characterized by

$$\begin{aligned} y^*_{it_i(\tau)} &= \alpha + \overline{x}_{i,t_i(\tau)-1}\beta + \lambda_1 \sum_{s=1}^{\tau} \delta_1^{s-1} y_{i,t_i(\tau-s)} + \lambda_2 \sum_{s=1}^{t_i(\tau)} \frac{\delta_2^{s-1} w_{in} Y_{n,t_i(\tau)-s}}{\sum_{s=1}^{t_i(\tau)} \delta_2^{s-1}} \\ &\quad + \gamma \ln \tau + \sigma\xi_i + \varepsilon_{it_i(\tau)}, \end{aligned} \tag{11}$$

and

$$\varepsilon_{it_i(\tau)} = \rho\varepsilon_{i,t_i(\tau-1)} + v_{it_i(\tau)},$$

where ξ_i, $v_{it_i(\tau)}$ are i.i.d.$N(0,1)$. The initial conditions on all variables for $t \leq 0$ are set to be zero, as we observe the data generating process from the very beginning. Most variables in (11) are defined as in the Markov model, while there are some changes in the specification of the entry rate differential as follows. Let $c^j_{i,t_i(\tau)-1}(R) = \sum_{s=1}^{t_i(\tau)-1} d^j_{is}(R)$ and $c^j_{-i,t_i(\tau)-1}(R) = \sum_{s=1}^{t_i(\tau)-1} d^j_{-is}(R)$ for $j = L, O$ and $R = IN, OUT$, with $d^j_{is}(R)$ given as before. Let the weight a player puts on the experience of other players relative to his/her own in calculating entry rate differential be ω. The perceived cumulative entry rate differential between limit-pricing output levels and the other output levels is given by

$$\overline{x}_{i,t_i(\tau)-1} = \frac{c^O_{i,t_i(\tau)-1}(IN) + \omega c^O_{-i,t_i(\tau)-1}(IN)}{c^O_{i,t_i(\tau)-1} + \omega c^O_{-i,t_i(\tau)-1}} - \frac{c^L_{i,t_i(\tau)-1}(IN) + \omega c^L_{-i,t_i(\tau)-1}(IN)}{c^L_{i,t_i(\tau)-1} + \omega c^L_{-i,t_i(\tau)-1}},$$

where $c^j_{i,t_i(\tau)-1} = c^j_{i,t_i(\tau)-1}(IN) + c^j_{i,t_i(\tau)-1}(OUT)$ for $j = L, O$. Analogous to $x_{i,t_i(\tau)-1}$ in the Markov model (10), $\overline{x}_{i,t_i(\tau)-1}$ here represents the payoff incentive for an M to limit price. The depreciation factors δ_1 and δ_2 measure the influence of past plays on the current choice. The weighted average $\sum_{s=1}^{t_i(\tau)} \delta_2^{s-1} w_{in} Y_{n,t_i(\tau)-s} / \sum_{s=1}^{t_i(\tau)} \delta_2^{s-1}$ captures the cumulative peer group effects on incumbent i's current decision. As in (10), we specify the row-normalized weighting matrix W_n, with its ith row being w_{in}, as $[(1_n \cdot 1_n' - I_n)/(n-1)]$. Thus, in the Polya model, the cumulative peer group effects are specified as the (weighted) average of the peers observed choices over the entire history. Based on the GHK simulator generated with 100 random draws, the SMLEs of the Polya model with samples from the experimental sessions with E_Hs and with E_Ls are reported in *Table 5* and *Table 6* respectively.

In games with E_Hs and E_Ls, the positive, and statistically significant, SMLEs for λ_1 of lagged choices imply that prior strategic play significantly increases the likelihood of future strategic plays for any given M. The interaction term for lagged choice by inexperience subjects is not statistically significant in games with E_Hs but has a negative, statistically significant coefficient for games with E_Ls.[11] This is consistent with the results from the Markov model that inexperienced subjects are less confident in their decision to play strategically as M_Ls than their more experienced counterparts, hence are more likely to revert back to non-strategic play.

In games with E_Hs the cumulative peer group effects captured by the SMLEs of λ_2 are positive and statistically significant for inexperienced subjects. In contrast, the SMLEs for the cumulative entry rate differential are not, and the statistical insignificance of β makes the estimate of ω extremely imprecise.[12] In games with E_Ls, cumulative peer group effects are positive and statistically significant overall, with even stronger peer group effects for inexperienced subjects (indicated by the positive coefficient value for the interaction term between peer group effects and NX, with t-ratio 1.484). Thus, inexperienced subjects are influenced more by peer group effects than experienced subjects in games with E_Ls, but unlike games with E_Hs, experienced subjects continue to be influenced by their peers, consistent with the fact that it takes longer for a separating equilibrium to emerge than a pooling equilibrium. The cumulative entry rate differential is statistically significant in games with E_Ls, but less so for inexperienced than experienced subjects. Further, as in the Markov model, subjects place much *less* weight on others' entry experiences than their own. And the marginal effect of own entry rate differential is smaller than that of the peer group effect in games with E_Ls.[13]

[11]The interaction terms in the Polya model are jointly significant at the 5 percent level with the minus two times log likelihood ratios being 11.76 for the games with E_Hs, and significant at the 1 percent level with the minus two times log likelihood ratios being 24.4 for games with E_Ls.

[12]Note that ω would not be identifiable if the coefficient of x_{it} were zero. The value of this estimate may reflect the insignificance of the coefficient estimate of x_{t-1} in this case.

[13]For the specification without interaction terms, in games with E_Hs, the average marginal cu-

We believe that the learning results reported on above come about for three reasons: (1) Adaptive learning is more demanding than social learning, as it requires that subjects form expectations about opponents' responses from observing opponents play as compared to social learning where subjects simply imitate peers' strategies. As such social learning is likely to be more prominent in the early stages of the learning process. (2) Because Es' choices are less stable in inexperienced subject sessions (especially with respect to entry on 5, 6 or 7 in games with E_Ls) than experienced subject sessions, entry rate differentials serve as a poor proxy for M's beliefs in those sessions. (3) Strategic play of M_Ls requires innovation, whereas strategic play by M_Hs simply requires imitating M_Ls choices. As such there must be some element of adaptive learning in games with E_Ls, while this is likely to be superseded by social learning from M_Ls choices (and responses to same) in games with E_Hs where such innovation is not required.

The SMLEs for both δ_1 and δ_2 are positive and statistically significant in games with E_Hs and E_Ls, indicating that a subject's current decision making is influenced by all past plays of the game. Note, δ_1 and δ_2 are not directly comparable as the depreciation factor for own lagged choices δ_1 is defined on the decision period, while the depreciation factor for peer group effects is defined on the (consecutive) calendar period. On average a subject only has a decision period (a chance to limit price) every 4 calendar periods (because a subject plays as an M only half time, and the type of M is randomly decided with equal probability). Take the games with E_Ls for instance, for the specification with no interaction terms, a generic M_L discounts peers' lagged choices $\delta_1/\delta_2^4 \approx 2$ times as fast as own lagged choices.

As in the Markov model, for games with E_Ls, we introduce the proportion of M_Hs attempting to pool by choosing output levels 3 and 4. Different from the Markov model where this value is calculated based on M_Hs' choice in the previous round only, we calculate its cumulative counterpart in the Polya model. The positively significant estimates of this variable in experimental sessions employing inexperienced subjects is consistent with the observation made by Cooper, Garvin and Kagel (1997*b*) that the "history-dependent adjustment process" starts with incumbents "at their myopia maxima, followed by an attempt to pool, and then ... separation". In both the

mulative peer group effect on the probability of limit pricing conditional on the exogenous variables and lagged choices is 0.058, and the average marginal effect of cumulative entry rate differential is 0.009. In games with E_Ls, the average marginal cumulative peer group effect is 0.113, and the average marginal effect of the cumulative entry rate differential is 0.085.

Markov and Polya models, other experience effects as represented by $\ln \tau$ are not statistically significant at conventional levels. And analogous to what happens in the Monte Carlo study, the weird negative sign of the SMLEs for ρ in the Markov model for games with E_Hs can now be explained by model misspecification.

As the Polya model does not nest the Markov model because of our different specifications of entry rate differential and the proportion of M_Hs attempting to pool, we address the issue of model selection by the well known Akaike information criterion (AIC) given as

$$AIC = -\frac{2}{\#obs} \log L + \frac{2\#p}{\#obs}, \tag{12}$$

where $\#obs$ is the sample size, $\#p$ is the number of parameters and $\log L$ is log likelihood of a model. According to (12), the Polya model is a better model than the Markov model as the former gives smaller value of AIC.[14]

One element that has been left out of the analysis reported on so far involves distinguishing between attempts at limit pricing and opposed to successful attempts at limit pricing. To do this we introduce two new variables into the regressions: Individual M's own success in limit pricing and the percentage of successful limit done by peers. We view these new regressors as, essentially, additional interaction terms, the results of which are reported in Appendix D. Introduction of these variables has essentially no effect on the log likelihood function for games with E_Hs so that the distinction has no impact on the results reported in this case.[15] In games with E_Ls own success at limit pricing plays a statistically significant role in promoting limit pricing (and diminishes the effect of the cumulative entry rate differential). The percentage of limit pricing by peers (as opposed to attempts at limit pricing) plays no statistically significant, independent role in promoting limit pricing in games with E_Ls. This probably comes about because attempts at limit pricing were usually successful so that imitators only needed innovators actions to promote limit pricing in this case.

[14] For the specification without interaction terms, in games with E_Hs, the AIC of the Markov model is 0.4807 and the AIC of the Polya model is 0.4788; and in games with E_Ls, the AIC of the Markov model is 0.3673 and the AIC of the Polya model is 0.3610.

[15] The minus two times log likelihood ratios for testing jointly the significance of new regressors in the Polya model without interactions with NX are, respectively, 1.86 for games with E_Hs, and 41.14 for games with E_Ls. The latter is significant at the 1 percent level with an asymptotic $\chi^2(2)$ distribution.

Table 5: SMLEs for the Polya Model (Experiments with High-Cost Type Entrants)

	Model w/ peer group effects				Model w/o peer group effects	
			w/ interactions			
cumulative entry rate differential (β)	0.069	(0.139)	0.279	(0.218)	0.097	(0.138)
cumulative entry rate differential $\times NX$	-		-0.271	(0.264)	-	
weight on others' experience (ω)	0.552	(6.227)	0.563	(3.865)	0.528	(4.216)
lagged choices (λ_1)	0.510^c	(0.159)	0.579^c	(0.211)	0.490^c	(0.161)
lagged choices $\times NX$	-		0.001	(0.154)	-	
depreciation factor (δ_1)	0.594^c	(0.158)	0.528^c	(0.173)	0.601^c	(0.161)
peer group effects (λ_2)	0.441^a	(0.250)	-0.254	(0.526)	-	
peer group effects $\times NX$	-		1.082^a	(0.620)	-	
depreciation factor (δ_2)	0.932^c	(0.338)	0.992^c	(0.210)	-	
experience within a session (γ)	0.069	(0.113)	-0.161	(0.229)	0.128	(0.110)
experience within a session $\times NX$	-		0.259	(0.202)	-	
constant (α)	-0.542^c	(0.165)	-0.739^c	(0.170)	-0.419^c	(0.150)
random effects (σ)	1.011^c	(0.124)	0.991^c	(0.125)	1.032^c	(0.125)
serial correlation (ρ)	0.102	(0.106)	0.051	(0.119)	0.119	(0.107)
Dummies:						
w/ experience of the same game	0.555^c	(0.149)	1.234^c	(0.313)	0.638^c	(0.135)
meaningful context	-0.093	(0.161)	-0.104	(0.158)	-0.083	(0.161)
Log Likelihood	-1083.41		-1077.53		-1085.18	

standard errors in parentheses

Table 6: SMLEs for the Polya Model (Experiments with Low-Cost Type Entrants)

	Model w/ peer group effects		Model w/o
		w/ interactions	peer group effects
cumulative entry rate differential (β)	0.563^c (0.133)	0.878^c (0.203)	0.885^c (0.122)
cumulative entry rate differential $\times NX$	-	-0.528^b (0.231)	-
weight on others' experience (ω)	0.098 (0.078)	0.120 (0.085)	0.219^b (0.103)
lagged choices (λ_1)	0.870^c (0.134)	1.048^c (0.146)	0.911^c (0.130)
lagged choices $\times NX$	-	-0.325^c (0.096)	-
depreciation factor (δ_1)	0.608^c (0.083)	0.605^c (0.076)	0.624^c (0.076)
peer group effects (λ_2)	1.094^c (0.194)	0.837^c (0.234)	-
peer group effects $\times NX$	-	0.580 (0.391)	-
depreciation factor (δ_2)	0.742^c (0.099)	0.769^c (0.103)	-
% of M_H choosing 3,4 (cumulative)	0.221 (0.176)	-0.052 (0.199)	0.351^b (0.173)
% of M_H choosing 3,4 $\times NX$	-	0.868^b (0.366)	-
experience within a session (γ)	-0.053 (0.055)	-0.120 (0.082)	-0.004 (0.054)
experience within a session $\times NX$	-	0.056 (0.097)	-
constant (α)	-1.754^c (0.113)	-1.886^c (0.129)	-1.731^c (0.111)
random effects (σ)	0.821^c (0.075)	0.785^c (0.077)	0.747^c (0.072)
serial correlation (ρ)	0.241^c (0.088)	0.252^c (0.087)	0.241^b (0.085)
Dummies:			
w/ experience of the same game	0.589^c (0.094)	0.794^c (0.144)	0.740^c (0.089)
meaningful context	-0.098 (0.094)	-0.110 (0.092)	-0.093 (0.087)
crossovers from E_H to E_L games	0.466^c (0.117)	0.705^c (0.165)	0.554^c (0.112)
Log Likelihood	-2068.10	-2055.90	-2088.56

standard errors in parentheses

Finally, the last column in *Tables 5* and *6* look at the impact of neglecting the peer group effects. The reference specification against which to compare these estimates is the first one reported in each case. For games with E_Hs there is little if any effect on any of the coefficient values estimated and only a small change in the log likelihood function. This is not surprising as peer group effects are only significant at the 10% level in this case. For games with E_Ls, the SMLEs for the entry rate differential and for weight on others' entry experiences are most effected by dropping the peer group effects, with both coefficients biased upwards. This is not too surprising since it is the increased entry differential in response to choices 5-7 versus other output levels (especially output levels $3, 4$) that drive M_Ls to limit price in the first place. In this context what the introduction of peer group effects does is to clarify the behavioral mechanism under which these increased entry rates operate. It is only partly related to what individual subjects have experienced themselves. Rather, much of the impact is related to what others have experienced and *their responses to same.* It is the latter that is largely missing by ignoring peer group or session level effects in the data in this case.

5 Concluding Remarks

This article has generalized Heckman's (1981) dynamic discrete choice panel data models by introducing lagged social interactions. The likelihood function for a general model has been derived and simulation method based on the unbiased GHK simulator has been proposed to implement the SML estimation. Monte Carlo experiments have been conducted to investigate the finite sample performance of the SMLEs for the Markov and Polya model with lagged social interactions. Some clear patterns have emerged from the Monte Carlo results.

- Parameters capturing the strength of true state dependence tend to be biased upward, and parameters capturing lagged social interactions tend to be biased downward in the Markov and Polya models when T is long. The biases are small for $T = 8$ and 15, and $m = 50$ or 100. Parameters capturing lagged social interactions are relatively more difficult to estimate precisely in the Polya model than in the Markov model.

- Overall, the SMLEs of the variance of random effects and the autoregressive parameter of disturbances have small downward biases in the Markov and Polya model.

- Given a fixed sample size, biases and SDs of the SMLEs of all parameters increase with group size, given the corresponding reduction in the number of groups. The estimates of parameters of state dependence and lagged social interactions are more sensitive to group size than the other parameters.

- The bias correction procedure reduces bias and RMSE, but the improvements are generally small. For further improvement, a larger number of random draws is desirable.

- In the Markov model, when positive lagged social interactions are ignored in the estimated model, the parameter of true state dependence is more upward biased and serial correlation in disturbances is more downward biased. These biases can be severe.

- In the Markov model, when the true specification on disturbances allows for random components and serial correlation but the estimated model incorporates only random components, the parameter of state dependence is upward biased and the parameter of lagged social interactions is downward biased. These biases are severe. On the other hand, when the estimated model incorporates only serial correlation but the random component is ignored, the parameter of state dependence can be severely downward biased and the parameter of lagged social interactions is moderately upward biased.

- When the data generating process is the Polya model but the estimated model is the Markov model, the parameter of state dependence can be severely biased upward and the parameter of lagged social interactions has some downward biases.

We apply the Markov and Polya models with lagged social interactions to investigate the learning process in laboratory experiments based on Milgrom and Roberts (1982) entry limit pricing game. The Polya model is superior to the Markov model as it has a more natural justification and provides a better fit to the data. We find that subjects' decisions are influenced by the past decisions of their peers in this experiment. These time-lagged peer group effects are more evident in the experiments employing subjects with no experience of the same or related games. These results suggest that the imitation of peers' strategies plays an important role in learning to play strategically. It would be interesting to see if this result continues to hold under an experimental design that does not provide subjects with full information regarding peers' play on their computer screens.

Appendices

A Payoff Schedules for the Price Game

Table 7: Payoffs for A Player (Incumbent)

A1(High Cost)			A2(Low Cost)		
Your Choice	X (In)	Y (Out)	X (In)	Y (Out)	Your Choice
1	−428	−220	204	545	1
2	−298	−110	333	678	2
3	8	165	355	700	3
4	103	448	378	723	4
5	125	470	350	695	5
6	148	493	283	648	6
7	125	470	250	615	7

Source: Cooper and Kagel (2004).

Table 8: Payoffs for B Player (Entrant)

	A Player's Type		
Your Action Choice	A1 (High Cost) Your Payoff	A2 (Low Cost) Your Payoff	Expected Value[a]
X (In)	219	594	406.5
Y (Out)	281	281	281

[a] Based on prior distribution (50% M_H, 50% M_L) of M types.

Source: Cooper and Kagel (2004).

B Summary of Data

Table 9: Experimental Treatments

Payoff Tables (Type of Entrants)	Prior Experience	Number of Sessions GC^a	MC	BNE Prediction
1 & 2 (High cost type E_Hs)	None or same game	7	9 (2)[b]	Pure-strategy pooling & separating equilibria exist
1 & 3 or 7 & 8 (Low cost type E_Ls)	None or same game	15 (9)	12 (7)	No pure-strategy pooling equilibria exist
	Game with high cost Es[c]	5	7 (2)	

[a]GC: generic context; MC: meaningful context

[b]number of inexperienced-subject sessions (number of experienced-subject sessions)

[c]crossover after the 1st 12-period cycle

C Monte Carlo Results

Table 10: Markov model with unobserved heterogeneity and serially correlated disturbances (sample size: G=50, n=4)

T	m	$\beta = 1$	$\lambda_1 = 0.2$	$\lambda_2 = 0.4$	$\sigma = \sqrt{0.5}$	$\rho = 0.4$
Bias unadjusted SMLE						
8	15	0.969 (0.067) [0.073]	0.246 (0.132) [0.139]	0.349 (0.141) [0.150]	0.628 (0.136) [0.157]	0.379 (0.123) [0.125]
15	15	0.965 (0.052) [0.063]	0.296 (0.105) [0.142]	0.318 (0.103) [0.132]	0.672 (0.077) [0.084]	0.323 (0.079) [0.111]
30	15	0.957 (0.035) [0.056]	0.345 (0.071) [0.161]	0.292 (0.068) [0.127]	0.686 (0.058) [0.061]	0.286 (0.045) [0.123]
8	50	0.991 (0.066) [0.067]	0.207 (0.124) [0.123]	0.387 (0.137) [0.137]	0.675 (0.111) [0.115]	0.398 (0.108) [0.108]
15	50	0.988 (0.052) [0.053]	0.236 (0.091) [0.097]	0.371 (0.102) [0.106]	0.692 (0.071) [0.073]	0.369 (0.061) [0.069]
30	50	0.982 (0.034) [0.039]	0.264 (0.068) [0.093]	0.354 (0.066) [0.080]	0.698 (0.053) [0.053]	0.349 (0.042) [0.066]
8	100	0.996 (0.065) [0.065]	0.203 (0.124) [0.124]	0.393 (0.135) [0.135]	0.684 (0.110) [0.112]	0.397 (0.104) [0.104]
15	100	0.994 (0.052) [0.052]	0.219 (0.088) [0.090]	0.384 (0.098) [0.099]	0.697 (0.070) [0.070]	0.382 (0.059) [0.061]
30	100	0.990 (0.035) [0.037]	0.235 (0.065) [0.074]	0.370 (0.065) [0.072]	0.702 (0.052) [0.052]	0.372 (0.040) [0.049]
Bias adjusted SMLE						
8	15	0.991 (0.069) [0.069]	0.217 (0.131) [0.132]	0.380 (0.143) [0.144]	0.681 (0.135) [0.138]	0.385 (0.127) [0.127]
15	15	0.982 (0.053) [0.056]	0.259 (0.107) [0.122]	0.352 (0.106) [0.116]	0.699 (0.078) [0.078]	0.347 (0.082) [0.098]
30	15	0.970 (0.035) [0.046]	0.307 (0.072) [0.129]	0.326 (0.069) [0.101]	0.704 (0.060) [0.060]	0.314 (0.046) [0.097]
8	50	1.000 (0.067) [0.067]	0.194 (0.123) [0.123]	0.400 (0.138) [0.137]	0.692 (0.107) [0.108]	0.403 (0.109) [0.109]
15	50	0.996 (0.052) [0.052]	0.214 (0.092) [0.092]	0.390 (0.103) [0.104]	0.701 (0.072) [0.072]	0.386 (0.062) [0.063]
30	50	0.991 (0.035) [0.036]	0.232 (0.068) [0.075]	0.377 (0.067) [0.071]	0.708 (0.054) [0.054]	0.374 (0.043) [0.050]
8	100	1.001 (0.066) [0.066]	0.195 (0.124) [0.124]	0.400 (0.135) [0.135]	0.692 (0.109) [0.109]	0.401 (0.104) [0.104]
15	100	0.999 (0.052) [0.052]	0.205 (0.088) [0.088]	0.395 (0.098) [0.098]	0.701 (0.070) [0.070]	0.394 (0.059) [0.060]
30	100	0.996 (0.036) [0.036]	0.212 (0.066) [0.067]	0.386 (0.066) [0.067]	0.708 (0.053) [0.053]	0.391 (0.040) [0.041]

The values are given as Means (SDs) [RMSEs]

Table 11: Markov model with unobserved heterogeneity and serially correlated disturbances (alternative group sizes)

T	m	$\beta = 1$	$\lambda_1 = 0.2$	$\lambda_2 = 0.4$	$\sigma = \sqrt{0.5}$	$\rho = 0.4$
$G = 25, n = 8$						
8	15	0.961 (0.067) [0.078]	0.261 (0.157) [0.168]	0.321 (0.180) [0.197]	0.604 (0.174) [0.202]	0.372 (0.140) [0.142]
8	50	0.986 (0.066) [0.067]	0.215 (0.141) [0.142]	0.374 (0.168) [0.169]	0.665 (0.126) [0.133]	0.392 (0.119) [0.119]
8	100	0.993 (0.063) [0.063]	0.195 (0.125) [0.124]	0.388 (0.166) [0.165]	0.675 (0.100) [0.105]	0.398 (0.106) [0.106]
$G = 10, n = 20$						
8	15	0.958 (0.069) [0.080]	0.285 (0.182) [0.200]	0.288 (0.218) [0.245]	0.605 (0.155) [0.185]	0.356 (0.154) [0.160]
8	50	0.983 (0.068) [0.070]	0.227 (0.160) [0.162]	0.357 (0.203) [0.207]	0.663 (0.119) [0.126]	0.384 (0.131) [0.131]
8	100	0.989 (0.067) [0.068]	0.218 (0.160) [0.161]	0.370 (0.199) [0.200]	0.676 (0.107) [0.111]	0.386 (0.126) [0.127]
$G = 4, n = 50$						
8	15	0.957 (0.069) [0.081]	0.300 (0.193) [0.217]	0.271 (0.238) [0.270]	0.606 (0.150) [0.181]	0.345 (0.158) [0.167]
8	50	0.985 (0.068) [0.070]	0.231 (0.161) [0.164]	0.354 (0.209) [0.213]	0.666 (0.113) [0.120]	0.381 (0.130) [0.131]
8	100	0.990 (0.068) [0.068]	0.221 (0.164) [0.165]	0.368 (0.208) [0.210]	0.675 (0.107) [0.112]	0.386 (0.127) [0.127]
$G = 1, n = 200$						
8	15	0.958 (0.071) [0.082]	0.302 (0.200) [0.224]	0.268 (0.252) [0.284]	0.606 (0.152) [0.182]	0.345 (0.163) [0.172]
8	50	0.985 (0.070) [0.071]	0.234 (0.171) [0.174]	0.352 (0.221) [0.226]	0.665 (0.116) [0.123]	0.380 (0.136) [0.137]
8	100	0.991 (0.069) [0.070]	0.220 (0.170) [0.171]	0.370 (0.217) [0.219]	0.674 (0.110) [0.114]	0.388 (0.130) [0.131]

Table 12: Model Misspecification (sample size: G=50, n=4)

T	m	$\beta = 1$	$\lambda_1 = 0.2$	$\lambda_2 = 0.4$	$\sigma = \sqrt{0.5}$	$\rho = 0.4$
8	15	0.939 (0.065)	0.478 (0.116)	-	0.617 (0.118)	0.236 (0.121)
15	15	0.941 (0.054)	0.493 (0.097)	-	0.643 (0.074)	0.214 (0.080)
8	50	0.961 (0.066)	0.469 (0.111)	-	0.662 (0.090)	0.228 (0.109)
15	50	0.962 (0.054)	0.459 (0.097)	-	0.657 (0.068)	0.241 (0.075)
8	100	0.966 (0.064)	0.466 (0.114)	-	0.665 (0.096)	0.229 (0.110)
15	100	0.967 (0.054)	0.444 (0.100)	-	0.661 (0.068)	0.254 (0.079)

Table 13: Error Misspecification (sample size: G=50, n=4)

T	m	$\beta = 1$	$\lambda_1 = 0.2$	$\lambda_2 = 0.4$	$\sigma = \sqrt{0.5}$	$\rho = 0.4$
Random components						
8	15	0.935 (0.066)	0.592 (0.090)	0.091 (0.122)	0.685 (0.085)	-
15	15	0.919 (0.051)	0.639 (0.058)	0.075 (0.090)	0.652 (0.069)	-
8	50	0.961 (0.067)	0.567 (0.091)	0.123 (0.120)	0.713 (0.081)	-
15	50	0.939 (0.051)	0.626 (0.057)	0.101 (0.084)	0.655 (0.063)	-
8	100	0.965 (0.068)	0.561 (0.090)	0.132 (0.118)	0.717 (0.079)	-
15	100	0.943 (0.052)	0.623 (0.057)	0.105 (0.085)	0.654 (0.058)	-
AR(1) correlated errors						
8	15	0.932 (0.061)	0.103 (0.103)	0.434 (0.124)	-	0.624 (0.055)
15	15	0.940 (0.051)	0.102 (0.081)	0.436 (0.095)	-	0.612 (0.043)
8	50	0.945 (0.060)	0.072 (0.100)	0.462 (0.122)	-	0.649 (0.050)
15	50	0.962 (0.052)	0.056 (0.078)	0.474 (0.096)	-	0.649 (0.039)
8	100	0.948 (0.061)	0.065 (0.100)	0.468 (0.122)	-	0.655 (0.049)
15	100	0.968 (0.051)	0.044 (0.076)	0.484 (0.096)	-	0.658 (0.038)

Table 14: Polya model with unobserved heterogeneity and serially correlated disturbances (sample size: G=50, n=4)

T	m	$\beta = 1$	$\lambda_1 = 0.2$	$\lambda_2 = 0.4$	$\sigma = \sqrt{0.5}$	$\rho = 0.4$
8	15	0.981 (0.071) [0.073]	0.218 (0.085) [0.087]	0.337 (0.249) [0.256]	0.647 (0.177) [0.187]	0.399 (0.102) [0.101]
15	15	0.977 (0.056) [0.060]	0.228 (0.050) [0.057]	0.301 (0.179) [0.204]	0.679 (0.088) [0.092]	0.366 (0.057) [0.066]
30	15	0.964 (0.036) [0.051]	0.255 (0.038) [0.067]	0.201 (0.134) [0.240]	0.676 (0.067) [0.074]	0.339 (0.033) [0.069]
8	50	0.995 (0.070) [0.070]	0.200 (0.083) [0.082]	0.394 (0.245) [0.245]	0.690 (0.134) [0.135]	0.399 (0.083) [0.083]
15	50	0.995 (0.055) [0.055]	0.204 (0.051) [0.051]	0.389 (0.181) [0.181]	0.697 (0.079) [0.080]	0.390 (0.054) [0.055]
30	50	0.988 (0.037) [0.038]	0.219 (0.040) [0.044]	0.329 (0.132) [0.149]	0.699 (0.062) [0.062]	0.381 (0.037) [0.041]
8	100	0.997 (0.068) [0.068]	0.198 (0.082) [0.082]	0.400 (0.246) [0.245]	0.691 (0.118) [0.119]	0.402 (0.083) [0.083]
15	100	0.997 (0.056) [0.056]	0.201 (0.052) [0.052]	0.398 (0.185) [0.185]	0.696 (0.081) [0.081]	0.393 (0.053) [0.054]
30	100	0.994 (0.037) [0.037]	0.211 (0.042) [0.043]	0.358 (0.133) [0.139]	0.703 (0.064) [0.064]	0.391 (0.036) [0.037]

Table 15: Model Misspecification (true model: Polya; estimated model: Markov; sample size: G=50, n=4)

T	m	$\beta = 1$	$\lambda_1 = 0.2$	$\lambda_2 = 0.4$	$\sigma = \sqrt{0.5}$	$\rho = 0.4$
8	15	0.978 (0.069)	0.429 (0.198)	0.289 (0.176)	0.710 (0.142)	0.250 (0.179)
15	15	0.968 (0.057)	0.515 (0.149)	0.304 (0.128)	0.733 (0.080)	0.196 (0.116)
8	50	1.004 (0.069)	0.408 (0.188)	0.321 (0.170)	0.767 (0.099)	0.242 (0.161)
15	50	0.993 (0.057)	0.436 (0.157)	0.369 (0.133)	0.751 (0.077)	0.252 (0.120)
8	100	1.009 (0.068)	0.392 (0.192)	0.334 (0.168)	0.770 (0.096)	0.253 (0.166)
15	100	0.998 (0.058)	0.416 (0.169)	0.382 (0.137)	0.754 (0.079)	0.267 (0.131)

D Alternative Specifications for Empirical Models

In the main content of the application, we focus on disentangling the influence on incumbents' current decisions from the entrants (captured by the entry rate differential) and from the other incumbents of the same type (captured by the peer group effects). Here we also consider some interactions between them in the Polya model. Let the dichotomous indicator o_{is} be 1 if incumbent i is not entered on in calendar period s, and 0 otherwise. Let $\overline{Y}_{ns}$ be an n-dimensional vector with the ith element being $\overline{y}_{is} = y_{is}o_{is}$. We consider an alternative specification of the Polya model (11) as follows

$$\begin{aligned} y^*_{it_i(\tau)} &= \alpha + \overline{x}_{i,t_i(\tau)-1}\beta + \sum_{s=1}^{\tau} \delta_1^{s-1}\left(\lambda_1 y_{i,t_i(\tau-s)} + \lambda_1' \overline{y}_{i,t_i(\tau-s)}\right) \\ &\quad + \sum_{s=1}^{t_i(\tau)} \frac{\delta_2^{s-1} w_{in}\left(\lambda_2 Y_{n,t_i(\tau)-s} + \lambda_2' \overline{Y}_{n,t_i(\tau)-s}\right)}{\sum_{s=1}^{t_i(\tau)} \delta_2^{s-1}} + \gamma \ln \tau + \sigma \xi_i + \varepsilon_{it_i(\tau)}, \end{aligned}$$

and

$$\varepsilon_{it_i(\tau)} = \rho \varepsilon_{i,t_i(\tau-1)} + v_{it_i(\tau)},$$

where ξ_i, $v_{it_i(\tau)}$ are i.i.d.$N(0,1)$. The coefficients λ_1' and λ_2' capture the influence of own successful limit pricing and peers' successful limit pricing on M's current choice respectively.

Based on the GHK simulator generated with 100 random draws, the SMLEs of the Polya model with samples from the experimental sessions with E_Hs and with E_Ls are reported in *Table 16* and *Table 17* respectively. In those tables, we also include the SMLEs of the original specification from *Table 5* and *Table 6* for ease of comparison.

The SMLEs of λ_1' are positively significant in games with E_L, indicating that, in the case where strategic play requires innovation, subjects are more confident in their decision to play strategically when limit pricing generates higher payoffs in the previous rounds. However, the SMLEs of λ_1' are insignificant in games with E_H, which is consistent with the insignificant estimates of β in the original specification. The SMLEs of λ_2' are insignificant, which is consistent with the insignificant estimates of ω in the original specification, indicating subjects tend to ignore previous entries on their peers when making current decisions.

Table 16: Alternative Specifications for the Polya Model (Experiments with High-Cost Type Entrants)

	Original Specification		Alternative Specification	
		w/ interactions		w/ interactions
cumulative entry rate differential (β)	0.069 (0.139)	0.279 (0.218)	0.173 (0.182)	0.287 (0.317)
cumulative entry rate differential $\times NX$	-	-0.271 (0.264)	-	-0.144 (0.396)
weight on others' experience (ω)	0.552 (6.227)	0.563 (3.865)	0.947 (5.152)	0.929 (5.027)
lagged choices (λ_1)	0.510^c (0.159)	0.579^c (0.211)	0.452^c (0.175)	0.543^a (0.305)
lagged choices $\times NX$	-	0.001 (0.154)	-	-0.052 (0.329)
successful limit pricing (λ_1')	-	-	0.099 (0.107)	0.057 (0.350)
successful limit pricing $\times NX$	-	-	-	0.098 (0.394)
depreciation factor (δ_1)	0.594^c (0.158)	0.528^c (0.173)	0.580^c (0.161)	0.498^c (0.179)
peer group effects (λ_2)	0.441^a (0.250)	-0.254 (0.526)	0.748^b (0.364)	-0.169 (0.872)
peer group effects $\times NX$	-	1.082^a (0.620)	-	1.383 (1.045)
positive peer group effects (λ_2')	-	-	-0.501 (0.442)	-0.118 (0.923)
positive peer group effects $\times NX$	-	-	-	-0.566 (1.188)
depreciation factor (δ_2)	0.932^c (0.338)	0.992^c (0.210)	0.905^c (0.283)	0.991^c (0.192)
experience within a session (γ)	0.069 (0.113)	-0.161 (0.229)	0.072 (0.114)	-0.150 (0.238)
experience within a session $\times NX$	-	0.259 (0.202)	-	0.268 (0.205)
constant (α)	-0.542^c (0.165)	-0.739^c (0.170)	-0.556^c (0.165)	-0.759^c (0.174)
random effects (σ)	1.011^c (0.124)	0.991^c (0.125)	1.009^c (0.126)	1.002^c (0.130)
serial correlation (ρ)	0.102 (0.106)	0.051 (0.119)	0.099 (0.107)	0.041 (0.124)
Dummies:				
w/ experience of the same game	0.555^c (0.149)	1.234^c (0.313)	0.567^c (0.146)	1.251^c (0.325)
meaningful context	-0.093 (0.161)	-0.104 (0.158)	-0.097 (0.160)	-0.106 (0.161)
Log Likelihood	-1083.41	-1077.53	-1082.48	-1076.23

standard errors in parentheses

Table 17: Alternative Specifications for the Polya Model (Experiments with Low-Cost Type Entrants)

	Original Specification		Alternative Specification	
		w/ interactions		w/ interactions
cumulative entry rate differential (β)	0.563^c (0.133)	0.878^c (0.203)	0.368^b (0.155)	0.824^c (0.250)
cumulative entry rate differential $\times NX$	-	-0.528^b (0.231)	-	-0.797^b (0.320)
weight on others' experience (ω)	0.098 (0.078)	0.120 (0.085)	0.811 (1.515)	0.804 (1.148)
lagged choices (λ_1)	0.870^c (0.134)	1.048^c (0.146)	0.423^c (0.153)	0.541^c (0.187)
lagged choices $\times NX$	-	-0.325^c (0.096)	-	-0.175 (0.164)
successful limit pricing (λ_1')	-	-	0.680^c (0.111)	0.785^c (0.164)
successful limit pricing $\times NX$	-	-	-	-0.207 (0.210)
depreciation factor (δ_1)	0.608^c (0.083)	0.605^c (0.076)	0.603^c (0.085)	0.593^c (0.078)
peer group effects (λ_2)	1.094^c (0.194)	0.837^c (0.234)	0.992^c (0.360)	1.238^b (0.487)
peer group effects $\times NX$	-	0.580 (0.391)	-	-0.623 (0.854)
positive peer group effects (λ_2')	-	-	0.206 (0.418)	-0.464 (0.568)
positive peer group effects $\times NX$	-	-	-	1.649 (1.056)
depreciation factor (δ_2)	0.742^c (0.099)	0.769^c (0.103)	0.746^c (0.095)	0.793^c (0.093)
% of M_H choosing 3,4 (cumulative)	0.221 (0.176)	-0.052 (0.199)	0.213 (0.177)	-0.106 (0.206)
% of M_H choosing 3,4 $\times NX$	-	0.868^b (0.366)	-	0.947^b (0.375)
experience within a session (γ)	-0.053 (0.055)	-0.120 (0.082)	-0.046 (0.055)	-0.122 (0.083)
experience within a session $\times NX$	-	0.056 (0.097)	-	0.066 (0.097)
constant (α)	-1.754^c (0.113)	-1.886^c (0.129)	-1.733^c (0.115)	-1.851^c (0.130)
random effects (σ)	0.821^c (0.075)	0.785^c (0.077)	0.830^c (0.076)	0.792^c (0.078)
serial correlation (ρ)	0.241^c (0.088)	0.252^c (0.087)	0.221^b (0.094)	0.223^b (0.092)
Dummies:				
w/ experience of the same game	0.589^c (0.094)	0.794^c (0.144)	0.597^c (0.095)	0.771^c (0.147)
meaningful context	-0.098 (0.094)	-0.110 (0.092)	-0.062 (0.095)	-0.069 (0.093)
crossovers from E_H to E_L games	0.466^c (0.117)	0.705^c (0.165)	0.482^c (0.118)	0.683^c (0.167)
Log Likelihood	-2068.10	-2055.90	-2047.53	-2034.57

standard errors in parentheses

References

Borsch-Supan, A. and V. Hajivassiliou. 1993. "Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models." *Journal of Econometrics* 58:347–368.

Butler, J. S. and R. Moffitt. 1982. "A computationally efficient quadrature procedure for the one-factor multinomial probit model." *Econometrica* 50:761–764.

Cho, I. and D. Kreps. 1987. "Signaling games and stable equilibria." *Quarterly Journal of Economics* 102:179–221.

Cooper, D. J. and J. H. Kagel. 2003*a*. "The impact of meaningful context on strategic play in signaling games." *Journal of Economic Behavior and Organization* 50:311–337.

Cooper, D. J. and J. H. Kagel. 2003*b*. "Lessons learned: generalizing learning across games." *American Economic Review* 93:202–207.

Cooper, D. J. and J. H. Kagel. 2004. "Learning and transfer in signaling games." Working Paper.

Cooper, D. J., S. Garvin and J. H. Kagel. 1997*a*. "Adaptive learning vs. equilibrium refinements in an entry limit pricing game." *The Economic Journal* 107:553–575.

Cooper, D. J., S. Garvin and J. H. Kagel. 1997*b*. "Signaling and adaptive learning in an entry limit pricing game." *RAND Journal of Economics* 28:662–683.

Fudenberg, D. and D. K. Levine. 1998. *The Theory of Learning in Games*. Cambridge: MIT Press.

Geweke, J. 1991. Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints. In *Computer science and statistics: Proceedings of the Twenty-Third Symposium on the Inference*. American Statistical Association Alexandria, VA: pp. 571–578.

Geweke, J. and M. Keane. 2001. Computationally intensive methods for integration in econometrics. In *Handbook of Econometrics*, ed. J. J. Heckman and E. Leamer. Vol. 5 Amsterdam: North-Holland.

Gourieroux, C. and A. Monfort. 1993. “Simulation-based inference: a survey with special reference to panel data models.” *Journal of Econometrics* 59:5–33.

Hajivassiliou, V. and D. McFadden. 1990. “The method of simulated scores for the estimation of LDV models: with an application to the problem of external debt crises.” Cowles foundation discussion paper No. 967. Yale University, New Haven, CT.

Heckman, J. J. 1981. Statistical models for discrete panel data. In *Structural Analysis of Discrete Data with Econometric Applications*, ed. C. F. Manski and D. McFadden. Cambridge: MIT Press.

Hsiao, C. 2003. *Analysis of Panel Data.* Cambridge: Cambridge University Press.

Keane, M. P. 1994. “A computationally practical simulation estimator for panel data.” *Econometrica* 62:95–116.

Lee, L. F. 1992. “On efficiency of methods of simulated moments and maximum simulated likelihood estimation of discrete response models.” *Econometric Theory* 8:518–552.

Lee, L. F. 1995. “Asymptotic bias in maximum simulated likelihood estimation of discrete choice models.” *Econometric Theory* 11:437–483.

Lee, L. F. 1997. “Simulated maximum likelihood estimation of dynamic discrete choice statistical models: some Monte Carlo results.” *Journal of Econometrics* 82:1–35.

Lee, L. F. 2000. “A numerically stable quadrature procedure for the one-factor random-component discrete choice model.” *Journal of Econometrics* 95:117–129.

Lee, L. F. 2001. “Interpolation, quadrature, and stochastic integration.” *Econometric Theory* 17:933–961.

Manski, C. F. 1993. “Identification of endogenous social effects: the reflection problem.” *The Review of Economic Studies* 60:531–542.

Milgrom, P. and J. Roberts. 1982. “Limit pricing and entry under incomplete information: an equilibrium analysis.” *Econometrica* 50:443–459.