5-2007

Dynamic Stochastic Games with Sequential State-to-State Transitions

Ulrich Doraszelski
University of Pennsylvania

Kenneth L. Judd

Dynamic Stochastic Games with Sequential State-to-State Transitions

Abstract

Discrete-time stochastic games with a finite number of states have been widely applied to study the strategic interactions among forward-looking players in dynamic environments. The model as written down by Ericson & Pakes (1995), Pakes & McGuire (1994, 2001) (hereafter, EP, PM1, and PM2), the subsequent literature (e.g., Gowrisankaran 1999, Fershtman & Pakes 2000, Benkard 2004), and in standard textbook treatments of stochastic games (e.g., Filar & Vrieze 1997, Basar & Olsder 1999) assumes that the states of all players change at exactly the same point in each period (say at the end of the period). That is, the transitions from this period's state to next period's state are simultaneous. As PM2 and Doraszelski & Judd (2004) (hereafter, DJ) point out, these games with simultaneous state-to-state transitions suffer from a "curse of dimensionality" since the cost of computing players' expectations over all possible future states increases exponentially in the number of state variables. However, there are many other ways to formulate dynamic stochastic games, and some of them may be computationally more tractable than others. In particular, we show that there are games with sequential state-to-state transitions that do not su®er from the curse of dimensionality in the expectation over successor states.

Disciplines

Behavioral Economics | Business | Cognitive Psychology | Marketing | Statistics and Probability

Comments

This is an unpublished manuscript.

Dynamic Stochastic Games with Sequential State-to-State Transitions

Ulrich Doraszelski[*]
Harvard University and CEPR

Kenneth L. Judd[†]
Hoover Institution and NBER

May 2007
Preliminary and incomplete.

1 Introduction

Discrete-time stochastic games with a finite number of states have been widely applied to study the strategic interactions among forward-looking players in dynamic environments. The model as written down by Ericson & Pakes (1995), Pakes & McGuire (1994, 2001) (hereafter, EP, PM1, and PM2), the subsequent literature (e.g., Gowrisankaran 1999, Fershtman & Pakes 2000, Benkard 2004), and in standard textbook treatments of stochastic games (e.g., Filar & Vrieze 1997, Basar & Olsder 1999) assumes that the states of all players change at exactly the same point in each period (say at the end of the period). That is, the transitions from this period's state to next period's state are simultaneous. As PM2 and Doraszelski & Judd (2004) (hereafter, DJ) point out, these games with simultaneous state-to-state transitions suffer from a "curse of dimensionality" since the cost of computing players' expectations over all possible future states increases exponentially in the number of state variables.

However, there are many other ways to formulate dynamic stochastic games, and some of them may be computationally more tractable than others. In particular, we show that there are games with sequential state-to-state transitions that do not suffer from the curse of dimensionality in the expectation over successor states.

2 Model

In this section we describe a discrete-time, finite-state stochastic game with sequential state-to-state transitions. Perhaps the simplest model is that each period one player is picked at random to choose an action. The state of the player with the move then changes in response

[*]Cambridge, MA 02138, U.S.A., doraszelski@harvard.edu.
[†]Stanford, CA 94305-6010, U.S.A., judd@hoover.stanford.edu.

to its chosen action. Then another random draw is taken to pick a player, and so on. Our game may thus be thought of as a "random-leadership Stackelberg game:" When a player chooses an action it knows that it is the leader at this point in time and that its rivals are the followers.

The notation is the same as in DJ. As in the discrete-time model with simultaneous state-to-state transitions, the horizon is infinite, the state of the game at period t is $\omega_t \in \Omega$, there are N players, and player i's action in period t is denoted by $x_t^i \in \mathbb{X}^i(\omega_t)$.

We let $\pi^{i,j}(x_t, \omega_t)$ denote the per-period payoff to player i when player j has the move, players' actions are $x_t = \left(x_t^1, \ldots, x_t^N\right)$, and the state is $\omega_t = \left(\omega_t^1, \ldots, \omega_t^N\right)$. We make the assumption that

Assumption 1 *Per-period payoffs $\pi^{i,j}(x_t, \omega_t)$ can be written as*

$$\begin{cases} \pi^{i,i}(x_t^i, \omega_t), & j = i, \\ \pi^{i,j}(\omega_t), & j \neq i. \end{cases} \tag{1}$$

That is, if player i has the move, then its per-period payoff depends on its own action but not those of its rivals as well as on the state, whereas its per-period payoff depends solely on the state if it does not have the move. To simplify the notation we set $\Phi^i(x_t, \omega_t, \omega_{t+1}) = 0$ in what follows.

We further make the assumption that the state-to-state transitions in a period are controlled entirely by the action of the player with the move alone and that a player's state can change only if it has the move. In this case

Assumption 2 *The law of motion is*

$$\Pr\left(\omega' | \omega_t, x_t\right) = \begin{cases} \Pr^i\left(\left(\omega'\right)^i | \omega_t^i, x_t^i\right), & \left(\omega'\right)^{-i} = \omega_t^{-i}, \\ 0, & \text{otherwise}. \end{cases} \tag{2}$$

We note from the outset that Assumptions 1 and 2, while critical for computational tractability, may be questionable in some applications of EP's framework (see Section 4.3 for details).

Two further comments are in order: First, because players move sequentially instead of simultaneously, the expected net present value of future cash flows to a player if it has the move in the current period is in general different than if the player does not have the move. Consequently, the value function of the player depends not only on the state of the game but also on the identity of the player with the move.[1]

Second, in the model with sequential state-to-state transitions, a player's state can change once in every N periods *on average*, whereas it can change once in every period in the model with simultaneous transitions. To make the models comparable in terms of the frequency of changes, we therefore shorten the length of a period in the model with sequential transitions and replace the discount factor β by $\sqrt[N]{\beta}$.

[1]Put differently, the identity of the player with the move becomes part of the state of the game.

Let $V^{i,j}(\omega)$ denote the expected net present value of future cash flows to player i if player j has the move and the state is ω. If player i has the move ($j = i$), then its Bellman equation is

$$V^{i,i}(\omega) = \max_{x^i} \pi^{i,i}(x^i, \omega) + \sqrt[N]{\beta} \mathrm{E}_{k',(\omega')^i} \left\{ V^{i,k'} \left(\left(\omega'\right)^i, \omega^{-i} \right) | \omega^i, x^i \right\}, \tag{3}$$

where k' denotes the next player to move. Since the identity of the next player to move is a random variable, players form expectations over it. Player i's strategy is given by

$$X^i(\omega) = \arg\max_{x^i} \pi^{i,i}(x^i, \omega) + \sqrt[N]{\beta} \mathrm{E}_{k',(\omega')^i} \left\{ V^{i,k'} \left(\left(\omega'\right)^i, \omega^{-i} \right) | \omega^i, x^i \right\}. \tag{4}$$

If player i does not have the move ($j \neq i$), then its Bellman equation is

$$V^{i,j}(\omega) = \pi^{i,j}(\omega) + \sqrt[N]{\beta} \mathrm{E}_{k',(\omega')^j} \left\{ V^{i,k'} \left(\left(\omega'\right)^j, \omega^{-j} \right) | \omega^j, X^j(\omega) \right\}. \tag{5}$$

The main advantage of the model with sequential state-to-state transitions now becomes clear. In the model with simultaneous transitions, there is one Bellman equation for player i and it contains one N-dimensional expectation that consists of K^N terms (assuming that a player can move to one of K states). In comparison, in the model with sequential transitions, there are N Bellman equations for player i (equation (3) plus $N-1$ incarnations of equation (5)). Each of them contains one one-dimensional expectation that consists of K terms. Since this yields a total of KN terms, sequential transitions break the curse of dimensionality.

The remaining problem with the model with sequential state-to-state transitions is that there are N Bellman equations for player i compared to one Bellman equation in the model with simultaneous transitions. Fortunately, this problem is easy to solve. Define

$$\overline{V}^i(\omega) = \frac{1}{N} \sum_{j=1}^{N} V^{i,j}(\omega) \tag{6}$$

to be the expected value function. Substituting equation (6) into equations (3)–(5) yields

$$\begin{aligned}
V^{i,i}(\omega) &= \max_{x^i} \pi^{i,i}(x^i, \omega) + \sqrt[N]{\beta} \mathrm{E}_{(\omega')^i} \left\{ \overline{V}^i \left(\left(\omega'\right)^i, \omega^{-i} \right) | \omega^i, x^i \right\}, & (7) \\
X^i(\omega) &= \arg\max_{x^i} \pi^{i,i}(x^i, \omega) + \sqrt[N]{\beta} \mathrm{E}_{(\omega')^i} \left\{ \overline{V}^i \left(\left(\omega'\right)^i, \omega^{-i} \right) | \omega^i, x^i \right\}, & (8) \\
V^{i,j}(\omega) &= \pi^{i,j}(\omega) + \sqrt[N]{\beta} \mathrm{E}_{(\omega')^j} \left\{ \overline{V}^i \left(\left(\omega'\right)^j, \omega^{-j} \right) | \omega^j, X^j(\omega) \right\}. & (9)
\end{aligned}$$

While $V^{i,j}(\omega)$ and $X^i(\omega)$ are the value and policy functions *after* it is known that player j has the move, $\overline{V}^i(\omega)$ is the value function *before* it is known who is next to move. Adding equation (7) and equation (9) for all $j \neq i$ yields

$$\overline{V}^i(\omega) = \frac{1}{N}\Bigg\{ \max_{x^i} \pi^{i,i}(x^i,\omega) + \sqrt[N]{\beta}\mathrm{E}_{(\omega')^i}\left\{\overline{V}^i\left(\left(\omega'\right)^i, \omega^{-i}\right)|\omega^i, x^i\right\}$$

$$+ \sum_{j\neq i} \pi^{i,j}(\omega) + \sqrt[N]{\beta}\mathrm{E}_{(\omega')^j}\left\{\overline{V}^i\left(\left(\omega'\right)^j, \omega^{-j}\right)|\omega^j, X^j(\omega)\right\}\Bigg\}, \quad (10)$$

$$X^i(\omega) = \arg\max_{x^i} \pi^{i,i}(x^i,\omega) + \sqrt[N]{\beta}\mathrm{E}_{(\omega')^i}\left\{\overline{V}^i\left(\left(\omega'\right)^i, \omega^{-i}\right)|\omega^i, x^i\right\}. \quad (11)$$

Now there is one Bellman equations for player i, just as in the model with simultaneous transitions. Moreover, equation (10) contains N one-dimensional expectations that consist of K terms each. Again this yields a total of KN terms, and sequential state-to-state transitions break the curse of dimensionality.

3 Computational Aspects

This section illustrates the computational advantages of sequential state-to-state transitions using the quality ladder model of Section 4 of DJ as an example.[2] Our algorithm for the model with sequential transitions adapts the block Gauss-Seidel scheme in Section 3.1 of DJ.

For the sake of brevity, we omit a full-fledged comparison of the time per iteration and the number of iterations and instead focus on the time to convergence. The left panel of Table 1 illustrates this comparison and the total gain from sequential state-to-state transitions. In addition, it reports on the continuous-time model in DJ. Note that the results in Table 1 differ somewhat from those in Table 4 in DJ because we used different hardware.

As can be seen from the middle panel of Table 1, the discrete-time model with sequential state-to-state transitions suffers an "iteration penalty." Even though it needs more iterations, the loss in the number of iterations is small when compared to the gain from avoiding the curse of dimensionality. Overall, the discrete-time model with sequential transitions is much faster to solve than the discrete-time model with simultaneous transitions, but not quite as fast as the continuous-time model. In fact, as comparing the middle and right panels of Table 1 shows, the gain from avoiding the curse of dimensionality is smaller and the loss in the number of iterations is larger in the discrete-time model with sequential state-to-state transitions than in the continuous-time model.

In ongoing research we investigate whether extrapolation techniques can be used to alleviate the iteration penalty in the model with sequential state-to-state transitions.

[2]Recall that we shorten the length of a period in the model with sequential transitions by a factor of N. To make the model comparable to the model with simultaneous transitions, we accordingly decrease per-period payoffs by a factor of N.

#firms	discrete time simultaneous transitions mins.	discrete time sequential transitions mins.	continuous time mins.	ratio discrete time simultaneous to sequential transitions: time per iteration	number of iterations	time to convergence	ratio discrete time simultaneous transitions to continuous time: time per iteration	number of iterations	time to convergence
2	1.41(-4)	1.00(-4)	4.84(-5)	2.55	0.55	1.40	3.13	0.93	2.91
3	1.33(-3)	8.11(-4)	3.98(-4)	4.58	0.36	1.64	6.00	0.56	3.34
4	1.24(-2)	4.18(-3)	2.36(-3)	10.72	0.28	2.96	12.60	0.42	5.25
5	1.02(-1)	1.99(-2)	9.68(-3)	22.61	0.23	5.13	29.71	0.36	10.56
6	7.74(-1)	6.89(-2)	3.15(-2)	57.84	0.19	11.23	75.64	0.32	24.52
7	5.19(0)	2.03(-1)	8.45(-2)	150.00	0.17	25.54	200.00	0.31	61.46
8	3.21(1)	5.99(-1)	2.31(-1)	352.17	0.15	53.66	462.86	0.30	139.09
9	1.94(2)	1.48(0)	5.25(-1)	960.78	0.14	131.12	1,256.41	0.29	370.08
10	*1.10(3)*	3.43(0)	1.09(0)	2,642.86	*0.12*	*320.59*	3,468.75	*0.29*	*1,009.25*
11	*5.80(3)*	7.37(0)	2.13(0)	7,134.15	*0.11*	*786.81*	9,435.48	*0.29*	*2,725.30*
12	*2.94(4)*	1.53(1)	3.98(0)	19,008.97	*0.10*	*1,923.53*	25,563.81	*0.29*	*7,383.72*
13	*1.51(5)*	2.97(1)	7.34(0)	54,527.14	*0.09*	*5,101.20*	71,566.92	*0.29*	*20,620.98*
14	*7.08(5)*	5.62(1)	1.28(1)	144,763.30	*0.09*	*12,592.71*	191,294.33	*0.29*	*55,118.70*

Table 1: Time to convergence. (k) is shorthand for $\times 10^k$. Stopping rule is "distance to truth$< 10^{-4}$." Entries in italics are based on an estimated 119 iterations to convergence in discrete time. Quality ladder model with $M = 9$ quality levels per firm and a discount factor of $\beta = 0.925$.

4 Conceptual Aspects

In Section 3 we have emphasized the computational advantages of sequential state-to-state transitions. Several other issues have to be considered in deciding whether to formulate an economic problem as a game with sequential or simultaneous transitions.

4.1 Economic Mechanism

Perhaps the main issue that has to be addressed is what economic mechanism gives rise to sequential state-to-state transitions. That is, is there a plausible story for why players are handed the move one after the other in a random order? The answer will undoubtedly depend on the application at hand.

Note that in some applications of EP's framework a player controls $D > 1$ coordinates of the state (e.g., Benkard 2004, Langohr 2003). If all states of a player are allowed to change at once, then computing the expectation over successor states involves summing over NK^D terms and thus suffers from a curse of dimensionality in D (but not in N). Alternatively one may extend our assumption that each period one player is picked at random to choose an action by postulating that, in addition, one coordinate of the state is picked at random and allowed to change, but it may be hard to motivate this assumption in an economically meaningful fashion.

4.2 Equilibrium and Dynamics

It is not obvious how players' behavior in equilibrium and the dynamics of the game implied by that behavior change as the model is recast with sequential state-to-state transitions. For example, in a preemption race (Fudenberg, Gilbert, Stiglitz & Tirole 1983, Harris & Vickers 1987) an early mover has a head start over a late mover, so that handing one player the move may well be decisive for the outcome of the race. This suggests that the Stackelberg character of the game with sequential transitions may have repercussions for players' behavior.

It turns out that in the quality ladder model the differences between sequential and simultaneous state-to-state transitions are very small. Figure 1 illustrates this point by plotting the equilibrium value and policy functions for the case of $N = 2$ firms and $M = 18$ quality levels per firm.

From the equilibrium policy function we construct the probability distribution over next period's state $\left((\omega')^1, (\omega')^2\right)$ given this period's state (ω^1, ω^2), i.e., the transition probability matrix that characterizes the Markov process of industry dynamics. This allows us to use stochastic process theory to analyze the Markov process of industry dynamics rather than rely on simulation. We compute the transient distribution over states in period t, $\mu^t(\cdot)$, starting from state $(1, 1)$. This tells us how likely each possible industry structure is in period t, given that both firms began the game at the minimal quality level. In addition, we

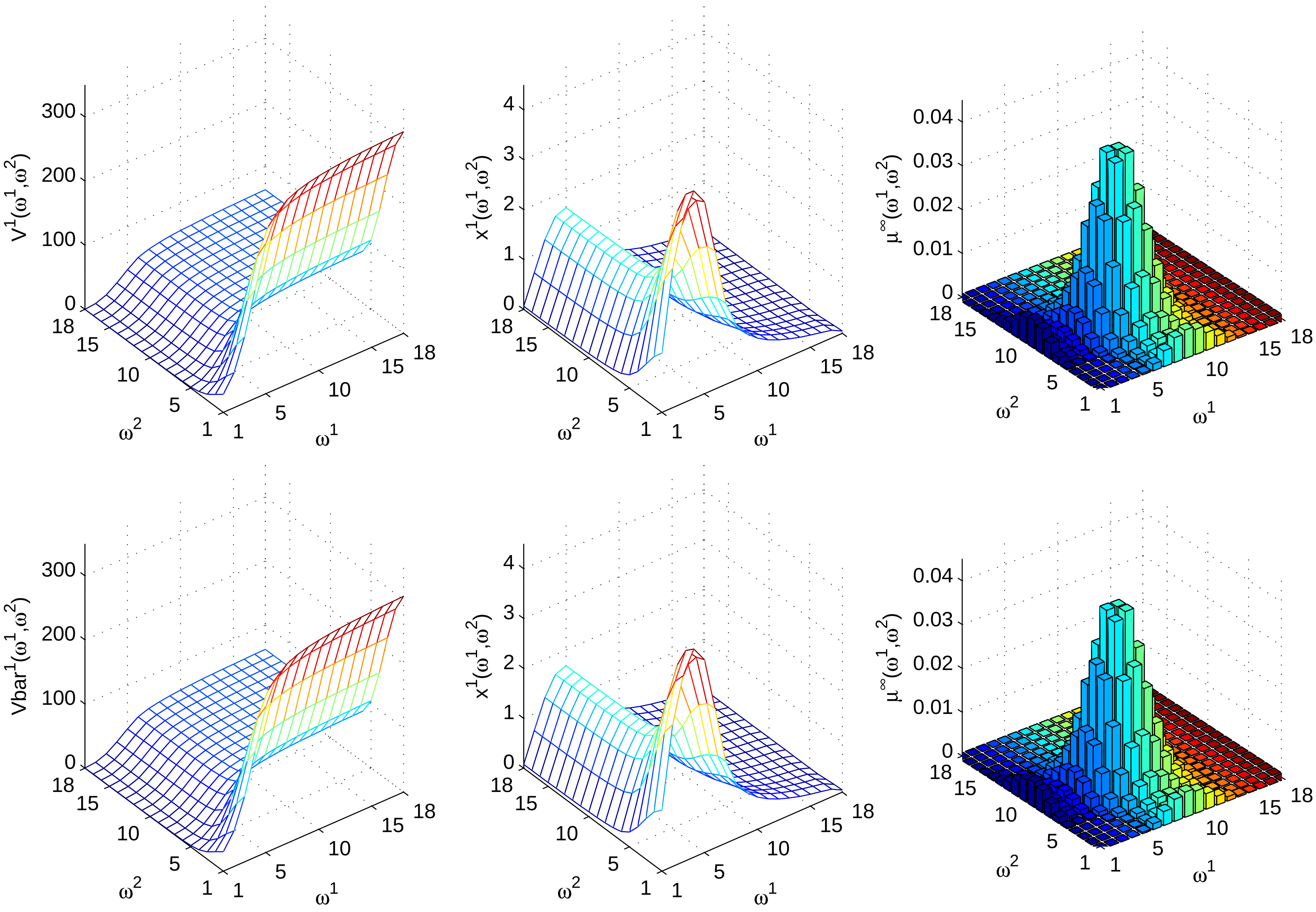

Figure 1: Equilibrium value and policy functions (left and middle panels) and limiting distribution (right panels). Simultaneous (upper panels) and sequential state-to-state transitions (lower panels). Quality ladder model with $N = 2$ firms, $M = 18$ quality levels per firm, and a discount factor of $\beta = 0.925$.

compute the limiting (or ergodic) distribution over states, $\mu^{\infty}(\cdot)$.[3] The transient distribution captures short-run dynamics and the limiting distribution captures long-run (or steady-state) dynamics.

Figure 1 also depicts the limiting distribution. As can be seen, the differences between sequential and simultaneous state-to-state transitions are very small. Table 2 summarizes the dynamics of the industry. We list the most likely industry structure (modal state) and its probability at various points in time. Note that, because we shorten the length of a period in the model with sequential transitions by a factor of 2, the transient distribution in period t in the model with simultaneous transitions is comparable to the transient distribution in period $2t$ in the model with sequential transitions. As can be seen, in the short run the industry evolves either in a symmetric or an asymmetric fashion. However, even if a firm is able to gain the upper hand over its rival in the short run, in the long run the most likely industry structure is symmetric and the limiting distribution leaves little probability mass on asymmetric industry structures (see again Figure 1). In addition, we report a firm's expected profit from product market competition and its expected investment in quality improvements along with their standard deviations in Table 2. Again the differences between sequential and simultaneous state-to-state transitions are very small. It is an open issue whether the same is true in other examples besides the quality ladder model.

4.3 Discussion of Assumptions and Range of Applications

Not all economic problems are well-suited to be formulated with sequential state-to-state transitions. In the remainder of this section, we argue that the structure of per-period payoffs and law of motion in Assumptions 1 and 2, respectively, may be questionable in some applications of EP's framework. We show that relaxing either one of these restrictions substantially increases the computational burden: Absent Assumption 1 the dimension of the state vector is expanded; in addition, the curse of dimensionality is resurrected absent Assumption 2.

4.3.1 Per-Period Payoffs

According to Assumption 1, if player i has the move, then its per-period payoff depends on its own action but not those of its rivals as well as on the state, whereas its per-period payoff depends solely on the state if it does not have the move. This structure of per-period payoffs entails a hidden assumption in most applications of EP's framework. The key fact to remember is that per-period payoffs are usually derived from the Nash equilibrium of a product market game. In the quality ladder model, for example, in each period firm i

[3] Let P be the $M^2 \times M^2$ transition probability matrix. The transient distribution in period t is given by $\mu^t = \mu^0 P^t$, where μ^0 is the $1 \times M^2$ initial distribution and P^t the t^{th} matrix power of P. The Markov process turns out to be irreducible. That is, all its states belong to a single closed communicating class and the $1 \times M^2$ limiting distribution μ^{∞} solves the system of linear equations $\mu^{\infty} = \mu^{\infty} P$.

t	most likely industry structure	prob.	profit mean	profit std. dev.	investment mean	investment std. dev.
5	(2,1), (1,2)	0.1177	1.20	3.02	2.20	1.12
10	(3,3)	0.0510	4.57	7.12	2.32	1.19
25	(7,1), (1,7)	0.0301	8.60	9.32	1.46	0.93
50	(7,7)	0.0307	7.78	8.18	0.97	0.61
100	(8,7), (7,8)	0.0360	6.74	6.56	0.90	0.53
∞	(8,7), (7,8)	0.0404	6.14	5.27	0.91	0.51

t	most likely industry structure	prob.	profit mean	profit std. dev.	investment mean	investment std. dev.
10	(2,1), (1,2)	0.1160	1.43	3.68	2.25	1.15
20	(3,3)	0.0488	4.86	7.56	2.35	1.21
50	(7,1), (1,7)	0.0300	8.75	9.46	1.46	0.94
100	(7,7)	0.0305	7.78	8.17	0.98	0.61
200	(8,7), (7,8)	0.0363	6.66	6.39	0.91	0.53
∞	(8,7), (7,8)	0.0406	6.08	5.10	0.91	0.50

Table 2: Most likely industry structure and its probability, expected profit and investment and their standard deviations. Simultaneous (upper panel) and sequential state-to-state transitions (lower panel). Quality ladder model with $N = 2$ firms, $M = 18$ quality levels per firm, and a discount factor of $\beta = 0.925$.

chooses the price p^i of product i after deciding on its investment x^i in quality improvements (see Section 4 of DJ for details). Allowing firm i's per-period payoff to depend solely on the quality of its and its rivals' products as encoded in $\omega = \left(\omega^1, \ldots, \omega^N\right)$ and on its investment x^i if it has the move thus presumes that firm i is able to reoptimize its price in response to each change in ω. Hence, recasting the quality ladder model as a game with sequential state-to-state transitions and per-period payoffs satisfying Assumption 1 entails that a firm gets to choose its price once in every period but its investment on average once in every N periods. Depending on the details of the institutional setting one may be able to justify this assumption by arguing that prices are less rigid than investments. If this is not possible, then one has to replace Assumption 1 with the assumption that a firm can choose both its price and its investment only when it has the move. However, as we show below, this alternative assumption causes the dimension of the state vector to expand and thus adds to the computational burden.

Assumption 1 is also difficult to maintain in applications of EP's framework where there is no distinction between actions (such as price) that affect per-period payoffs but not state-to-state transitions and actions (such as investment) that affect both. For example, in a learning-by-doing model such as Benkard (2004), a firm's output decision determines not only how much it receives in the product market but also how far it moves down the learning curve. In cases like this, per-period payoffs depend on the actions of all players, irrespective of whether they have the move or not.

To explore the consequences of giving up Assumption 1 let us assume that a player remains committed to the chosen action until it once again has the move. Since the chosen action affects the per-period payoff of player i for the entire duration, we write the per-period payoff to player i as $\pi^{i,j}(x_t, \omega_t)$. In the learning-by-doing model, x_t^i is the output of firm i. In the quality ladder model, we have $x_t^i = \left(x_t^{i,1}, x_t^{i,2}\right)$, where $x_t^{i,1}$ denotes investment and $x_t^{i,2}$ denotes price, and the per-period payoff to player i can be simplified to

$$\begin{cases} \pi^{i,i}(x_t^{i,1}, x_t^{1,2}, \ldots, x_t^{N,2}, \omega_t), & j = i, \\ \pi^{i,j}(x_t^{1,2}, \ldots, x_t^{N,2}, \omega_t), & j \neq i. \end{cases}$$

That is, if player i has the move ($j = i$), then its per-period payoff depends on its investment and all prices and it depends on all prices if player i does not have the move ($j \neq i$).

The key fact is that if the per-period payoff to player i is $\pi^{i,j}(x_t, \omega_t)$, then, from the point of view of the player with the move, the actions of the other players are part of the state of the game, similar to Maskin & Tirole (1988*a*, 1988*b*). To see this more clearly, let $V^{i,j}(\omega, x^{-j})$ denote the expected net present value of future cash flows to player i if player j has the move, the state is ω, and players other than j are locked into actions x^{-j}. If player i has the move ($j = i$), then the Bellman equation (3) becomes

$$V^{i,i}(\omega, x^{-i}) = \max_{x^i} \pi^{i,i}(x^i, x^{-i}, \omega) + \beta \mathrm{E}_{k',(\omega')^i} \left\{ V^{i,k'}\left(\left(\omega'\right)^i, \omega^{-i}, x^i, x^{-i,k'} \right) | \omega^i, x^i \right\}, \quad (12)$$

where $x^{-i,k'}$ is obtained from x by deleting x^i and $x^{k'}$. Note that while the locked-in actions in x^{-i} do not enter into the state-to-state transitions, they enter into the per-period payoff of player i and are therefore part of the state of the game. Similarly, if player i does not have the move ($j \neq i$), then the Bellman equation (5) becomes

$$V^{i,j}(\omega, x^{-j}) = \pi^{i,j}(X^j(\omega, x^{-j}), x^{-j}, \omega) + \beta \mathrm{E}_{k',(\omega')^j} \left\{ V^{i,k'}\left(\left(\omega'\right)^j, \omega^{-j}, X^j(\omega, x^{-j}), x^{-j,k'} \right) | \omega^j, X^j(\omega, x^{-j}) \right\}. \quad (13)$$

As can be seen by comparing equations (12) and (13) with equations (3) and (5), respectively, the dimension of the state vector is expanded. In this setting, discrete actions pose a problem insofar as the number of states that have to be examined to compute an equilibrium grows rapidly in the number of state variables, thereby increasing both time and memory requirements. Continuous actions, by contrast, require an entirely different set of techniques to compute an equilibrium since players' values and policies can no longer be represented as vectors and instead have to be represented as functions. Judd (1998) gives an introduction to these so-called projection techniques and Rui & Miranda (1996, 2001), Doraszelski (2003), and Hall, Royer & Van Audenrode (2003) apply them to dynamic games with a continuum of states.

4.3.2 Law of Motion

Turning to the law of motion, according to Assumption 2 a player's state can change only if it has the move. This structure may be inappropriate for some applications of EP's framework. Consider again the learning-by-doing model. Given that the firm remains committed to its output decision on average for N periods in the model with sequential state-to-state transitions, it seems strange to assume that its position on the learning curve as represented by its state can change only in the period it has the move. A more natural assumption as that the firm's output decision affects its per-period payoff as well as the transitions in its state for the entire duration until it once again has the move, contrary to Assumption 2.

To explore the consequences of giving up Assumption 2 let us assume that a player remains committed to the chosen action until it once again has the move. Moreover, the state-to-state transitions in a period are controlled by the currently chosen action of the player with the move and the previously chosen actions of its rivals. In this case the law of motion is

$$\Pr\left(\omega'|\omega_t, x_t\right) = \prod_{i=1}^{N} \Pr^i\left(\left(\omega'\right)^i |\omega_t^i, x_t^i\right)$$

and the curse of dimensionality in computing the expectation over successor states is resurrected. In addition, because the locked-in actions enter into the state-to-state transitions, players' actions are part of the state of the game.

5 Concluding Remarks

There are many ways to formulate dynamic stochastic games, and some of them are computationally more tractable than others. DJ propose continuous-time models in order to avoid the curse of dimensionality in the expectation over successor states. In this note we study discrete-time games with sequential state-to-state transitions. Our assumption is that each period one player is picked at random to choose an action and that the player's state then changes in response.

There are a number of issues that require more thought. First, it may be possible to alleviate the iteration penalty in the model with sequential state-to-state transitions, thus making it more competitive to the continuous-time model in DJ.

Second, the issue how players' behavior in equilibrium and the dynamics of the game implied by that behavior change as the model is recast with sequential state-to-state transitions merits further attention. While the differences are minor in the quality ladder model, this is just a particular example.

Third, there clearly are alternative assumptions that avoid the simultaneous state-to-state transitions in EP, PM1, and PM2 and thus also the curse of dimensionality. For example, one may assume that each period all players choose actions but that then one player is picked at random and its state is allowed to change in response to its previously chosen action. An important task for future research is therefore to describe more generally the class of discrete-time models that do not suffer from the curse of dimensionality.

References

Basar, T. & Olsder, J. (1999), *Dynamic noncooperative game theory*, 2nd edn, Society for Industrial and Applied Mathematics, Philadelphia.

Benkard, L. (2004), 'A dynamic analysis of the market for wide-bodied commercial aircraft', *Review of Economic Studies* **71**(3), 581–611.

Doraszelski, U. (2003), 'An R&D race with knowledge accumulation', *Rand Journal of Economics* **34**(1), 19–41.

Doraszelski, U. & Judd, K. (2004), Avoiding the curse of dimensionality in dynamic stochastic games, Technical working paper no. 304, NBER, Cambridge.

Ericson, R. & Pakes, A. (1995), 'Markov-perfect industry dynamics: A framework for empirical work', *Review of Economic Studies* **62**, 53–82.

Fershtman, C. & Pakes, A. (2000), 'A dynamic oligopoly with collusion and price wars', *Rand Journal of Economics* **31**, 294–326.

Filar, J. & Vrieze, K. (1997), *Competitive Markov decision processes*, Springer, New York.

Fudenberg, D., Gilbert, R., Stiglitz, J. & Tirole, J. (1983), 'Preemption, leapfrogging and competition in patent races', *European Economic Review* **22**, 3–31.

Gowrisankaran, G. (1999), 'A dynamic model of endogenous horizontal mergers', *Rand Journal of Economics* **30**(1), 56–83.

Hall, R., Royer, J. & Van Audenrode, M. (2003), Potential competition and the prices of network goods: Desktop software, Working paper, Hoover Institution, Stanford.

Harris, C. & Vickers, J. (1987), 'Racing with uncertainty', *Review of Economic Studies* **54**(1), 1–21.

Judd, K. (1998), *Numerical methods in economics*, MIT Press, Cambridge.

Langohr, P. (2003), Competitive convergence and divergence: Capability and position dynamics, Working paper, Northwestern University, Evanston.

Maskin, E. & Tirole, J. (1988*a*), 'A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs', *Econometrica* **56**(3), 549–569.

Maskin, E. & Tirole, J. (1988*b*), 'A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles', *Econometrica* **56**(3), 571–599.

Pakes, A. & McGuire, P. (1994), 'Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model', *Rand Journal of Economics* **25**(4), 555–589.

Pakes, A. & McGuire, P. (2001), 'Stochastic algorithms, symmetric Markov perfect equilibrium, and the "curse" of dimensionality', *Econometrica* **69**(5), 1261–1281.

Rui, X. & Miranda, M. (1996), 'Solving nonlinear dynamic games via orthogonal collocation: An application to international commodity markets', *Annals of Operations Research* **68**, 89–108.

Vedenov, D. & Miranda, M. (2001), 'Numerial solution of dynamic oligopoly games with capital investment', *Economic Theory* **18**, 237–261.