Dynamic Games: Numerical Methods and Applications
Problem Set

Prof. Ulrich Doraszelski, University of Pennsylvania

If you have little or no experience with dynamic stochastic games, try to solve exercises 1a and 1b. If you have more experience, try exercises 2a and 2b (or go directly to exercise 2b). Please work in groups of 2-3 people.

Exercise 1a: Quality Ladder Monopoly

The exercise below is inspired by the quality ladder model of Pakes & McGuire (1994).

The monopolist produces a product of quality ω. Quality is assumed to be discrete, i.e., $\omega \in \{1, \ldots, L\}$, and evolves over time in response to investment and depreciation. The model is cast in discrete time and has an infinite horizon to avoid end effects. I first describe the static product market model and then turn to investment dynamics.

Product market. There is a continuum of consumers with heterogenous tastes. Each consumer purchases at most one unit of one product. The utility consumer m derives from purchasing the product is $g(\omega) - p + \epsilon_m$, where

$$g(\omega) = \begin{cases} 3\omega - 4 & \text{if} \quad \omega \leq 5, \\ 12 + \ln\left(2 - \exp\left(16 - 3\omega\right)\right) & \text{if} \quad \omega > 5 \end{cases}$$

maps the quality of the product into the consumer's valuation for it and ϵ_m represents taste differences among consumers. There is a no-purchase alternative with utility normalized to zero. I assume that the idiosyncratic shock ϵ_m is independently and identically logistically distributed across consumers; therefore, the monopolist's demand is

$$D(p; \omega) = M \frac{\exp\left(g(\omega) - p\right)}{1 + \exp\left(g(\omega) - p\right)},$$

where $M > 0$ is the size of the market (the measure of consumers).

The monopolist chooses the price p of its product in each period to maximize profits, thereby solving

$$\max_{p\geq 0} D(p;\omega)\,(p-c)\,,$$

where $c \geq 0$ is the marginal cost of production. The FOC for an interior solution is

$$0 = 1 - \frac{1}{1+\exp(g(\omega)-p)}\,(p-c)\,.$$

Let $p(\omega)$ denote the solution to the monopolist's problem. The monopolist's profit is given by

$$\pi(\omega) = D(p(\omega);\omega)(p(\omega)-c).$$

Investment dynamics. The monopolist's state ω represents the quality of its product in the present period. The quality of its product in the subsequent period, ω', is determined by its investment $x \geq 0$ in quality improvements and by depreciation. The outcomes of the investment and depreciation processes are assumed to be stochastic. If the investment is successful, then the quality increases by one level. Expenditures in investment enhance the probability of success; in particular, the probability of success is $\frac{\alpha x}{1+\alpha x}$, where $\alpha > 0$ is a measure of the effectiveness of investment. If the firm is hit by a depreciation shock, then the quality decreases by one level; this is assumed to happen with probability $\delta \in [0,1]$.

Combining the investment and depreciation processes, the quality of the monopolist's product changes according to the transition probability

$$\Pr(\omega'|\omega,x) = \begin{cases} \frac{(1-\delta)\alpha x}{1+\alpha x} & \text{if} \quad \omega' = \omega+1, \\ \frac{1-\delta+\delta\alpha x}{1+\alpha x} & \text{if} \quad \omega' = \omega, \\ \frac{\delta}{1+\alpha x} & \text{if} \quad \omega' = \omega-1 \end{cases}$$

if $\omega \in \{2,\ldots,L-1\}$. Since the monopolist cannot move further down (up) from the lowest (highest) product quality, I set

$$\begin{aligned} \Pr(\omega'|1,x) &= \begin{cases} \frac{(1-\delta)\alpha x}{1+\alpha x} & \text{if} \quad \omega' = 2, \\ \frac{1+\delta\alpha x}{1+\alpha x} & \text{if} \quad \omega' = 1, \end{cases} \\ \Pr(\omega'|L,x) &= \begin{cases} \frac{1-\delta+\alpha x}{1+\alpha x} & \text{if} \quad \omega' = L, \\ \frac{\delta}{1+\alpha x} & \text{if} \quad \omega' = L-1. \end{cases} \end{aligned}$$

Let $V(\omega)$ denote the expected net present value of future cash flows to the monopolist if the current state is ω. The Bellman equation is

$$V(\omega) = \max_{x\geq 0} \pi(\omega) - x + \beta \sum_{\omega'=1}^{L} V(\omega')\Pr(\omega'|\omega,x),$$

where $\beta \in [0,1)$ is the discount factor. The Bellman equation adds the current cash flow, $\pi(\omega) - x$, to the appropriately discounted expected future cash flow, $\sum_{\omega'=1}^{L} V(\omega') \Pr(\omega'|\omega, x)$, where the expectation is taken over the successor states ω'.

The monopolist's investment strategy is given by

$$x(\omega) = \arg\max_{x \geq 0} \pi(\omega) - x + \beta \sum_{\omega'=1}^{L} V(\omega') \Pr(\omega'|\omega, x).$$

The FOC for an interior solution is

$$-1 + \beta \sum_{\omega'=1}^{L} V(\omega') \frac{\partial \Pr(\omega'|\omega, x)}{\partial x} = 0.$$

Consider $\omega \in \{2, \ldots, L-1\}$. Given the assumed functional forms, solving the FOC for x yields

$$\frac{-1 + \sqrt{\beta\alpha\left((1-\delta)(V(\omega+1) - V(\omega)) + \delta(V(\omega) - V(\omega-1))\right)}}{\alpha}.$$

The second-order condition (SOC) reduces to

$$-\left((1-\delta)(V(\omega) - V(\omega)) + \delta(V(\omega) - V(\omega-1))\right) < 0.$$

Hence, the SOC is satisfied whenever a solution to the FOC exists. Moreover, the objective function equals $\pi(\omega) + \beta\{(1-\delta)V(\omega) + \delta V(\omega-1)\}$ at $x = 0$ and approaches $-\infty$ as x approaches ∞. Hence, the objective function is decreasing when a solution to the FOC fails to exist. Thus,

$$x(\omega) = \frac{-1 + \sqrt{\max\left\{1, \beta\alpha\left((1-\delta)(V(\omega+1) - V(\omega)) + \delta(V(\omega) - V(\omega-1))\right)\right\}}}{\alpha}.$$

If $\omega = 1$, the monopolist's investment strategy can be derived using similar arguments; it is

$$x(1) = \frac{-1 + \sqrt{\max\left\{1, \beta\alpha(1-\delta)\left(V(2) - V(1)\right)\right\}}}{\alpha}.$$

Finally, if $\omega = L$, then

$$x(L) = \frac{-1 + \sqrt{\max\left\{1, \beta\alpha\delta\left(V(L) - V(L-1)\right)\right\}}}{\alpha}.$$

Parameterization. The number of quality levels is $L = 18$, the size of the market is $M = 5$, the marginal cost of production is $c = 5$, the effectiveness of investment is $\alpha = 3$, the depreciation probability is $\delta = 0.7$, and the discount factor is $\beta = 0.925$, which corresponds to a yearly interest rate of 8.1%.

Exercise. First compute the per-period profit function $\pi(\omega)$ by numerically solving the FOC. Then solve the monopolist's dynamic programming problem using value function iteration (or another method such as policy function iteration).

Exercise 1b: Quality Ladder Duopoly

This exercise introduces competition into the quality ladder model from Exercise 1a. Since many of the ingredients of the model remain the same, I focus on pointing out the differences between the dynamic game and the single-agent dynamic programming problem from a previous exercise.

There are two firms. Firm n produces a product of quality $\omega_n \in \{1, \ldots, L\}$. The state space is thus $\Omega \in \{1, \ldots, L\}^2$. I refer to $\omega = (\omega_1, \omega_2)$ as the state of the industry and to ω_n as the state of firm n.

Product market. The demand for firm n's product is

$$D_n(p_1, p_2; \omega) = M \frac{\exp\left(g(\omega_n) - p_n\right)}{1 + \sum_{k=1}^{2} \exp\left(g(\omega_k) - p_k\right)}.$$

Firm n chooses the price of its product in each period to maximize profits, thereby solving

$$\max_{p_n \geq 0} D_n(p_1, p_2; \omega)(p_n - c).$$

The FOC for an interior solution is

$$0 = 1 - \frac{1 + \exp\left(g(\omega_{-n}) - p_{-n}\right)}{1 + \exp\left(g(\omega_n) - p_n\right) + \exp\left(g(\omega_{-n}) - p_{-n}\right)}(p_n - c), \quad n \neq -n.$$

Let $(p_1(\omega), p_2(\omega))$ denote the Nash equilibrium of the product market game. Firm n's profit is given by

$$\pi_n(\omega) = D_n(p_1(\omega), p_2(\omega); \omega)(p_n(\omega) - c).$$

Investment dynamics. The Bellman equation for firm n is

$$V_n(\omega) = \max_{x_n \geq 0} \pi_n(\omega) - x_n + \beta \sum_{\omega'_n = 1}^{L} W_n(\omega'_n; \omega_{-n}, x_{-n}(\omega)) \Pr(\omega'_n | \omega_n, x_n),$$

where

$$W_n(\omega_n'; \omega_{-n}, x_{-n}(\omega)) = \sum_{\omega_{-n}'=1}^{L} V_n(\omega') \Pr(\omega_{-n}'|\omega_{-n}, x_{-n}(\omega))$$

is the expectation (with respect to its rival's successor state) of the value function of firm n conditional on firm n's successor state being ω_n' and $x_{-n}(\omega)$ is the rival's investment strategy.

The investment strategy of firm n is given by

$$x_n(\omega) = \arg\max_{x_n \geq 0} \pi_n(\omega) - x_n + \beta \sum_{\omega_n'=1}^{L} W_n(\omega_n') \Pr(\omega_n'|\omega_n, x_n),$$

where $W_n(\omega_n')$ is shorthand for $W_n(\omega_n'; \omega_{-n}, x_{-n}(\omega))$. Specifically, if $\omega_n \in \{2, \ldots, L-1\}$, then

$$x_n(\omega) = \frac{-1 + \sqrt{\max\{1, \beta\alpha\left((1-\delta)(W_n(\omega_n+1) - W_n(\omega_n)) + \delta(W_n(\omega_n) - W_n(\omega_n-1))\right)\}}}{\alpha}.$$

If $\omega_n = 1$, then

$$x_n(\omega) = \frac{-1 + \sqrt{\max\{1, \beta\alpha(1-\delta)\left(W_n(2) - W_n(1)\right)\}}}{\alpha}.$$

Finally, if $\omega_n = L$, then

$$x_n(\omega) = \frac{-1 + \sqrt{\max\{1, \beta\alpha\delta\left(W_n(L) - W_n(L-1)\right)\}}}{\alpha}.$$

Equilibrium. Note that profits from product market competition are symmetric, i.e., $\pi_1(\omega_1, \omega_2) = \pi_2(\omega_2, \omega_1)$. The remaining primitives are also symmetric. I therefore restrict attention to symmetric Markov perfect equilibria (MPE). Such a MPE is characterized by a value function $V(\omega)$ and a policy function $x(\omega)$ such that, if $V(\omega)$ is firm 1's value function, then firm 2's value function is given by $V_2(\omega_1, \omega_2) = V(\omega_2, \omega_1)$. Similarly, if $x(\omega)$ is firm 1's policy function, then firm 2's policy function is given by $x_2(\omega_1, \omega_2) = x(\omega_2, \omega_1)$. Existence of a symmetric MPE in pure strategies follows from the arguments in Doraszelski & Satterthwaite (2010) provided that investment is bounded above.

Algorithm. To compute the MPE, Pakes & McGuire (1994) suggest an algorithm that essentially adapts value function iteration to dynamic games. The algorithm proceeds as follows:

1. Make initial guesses for the value and policy functions (or, more precisely, $L \times L$

matrices), $\mathbf{V}^0$ and $\mathbf{x}^0$, choose a stopping criterion $\epsilon > 0$, and initialize the iteration counter to $l = 1$.

2. For all states $\omega \in \Omega$ compute

$$x^{l+1}(\omega) = \arg\max_{x_1 \geq 0} \pi_1(\omega) - x_1 + \beta \sum_{\omega_1'=1}^{L} W^l(\omega_1') \Pr(\omega_1'|\omega_1, x_1)$$

and

$$V^{l+1}(\omega) = \pi_1(\omega) - x^{l+1}(\omega) + \beta \sum_{\omega_1'=1}^{L} W^l(\omega_1') \Pr(\omega_1'|\omega_1, x^{l+1}(\omega)),$$

where

$$W^l(\omega_1') = \sum_{\omega_2'=1}^{L} V^l(\omega') \Pr(\omega_2'|\omega_2, x^l(\omega_2, \omega_1)).$$

3. If

$$\max_{\omega \in \Omega} \left| \frac{V^{l+1}(\omega) - V^l(\omega)}{1 + |V^{l+1}(\omega)|} \right| < \epsilon \quad \wedge \quad \max_{\omega \in \Omega} \left| \frac{x^{l+1}(\omega) - x^l(\omega)}{1 + |x^{l+1}(\omega)|} \right| < \epsilon$$

then stop; else increment the iteration counter l by one and go to step 2.

Unlike value function iteration for single-agent dynamic programming problems, there is no guarantee that the above algorithm converges. If it fails to converge, a trick that often works is to go through an additional dampening step before returning to step 2. This dampening step assigns

$$\begin{aligned} \mathbf{V}^{l+1} &\leftarrow \lambda \mathbf{V}^{l+1} + (1-\lambda)\mathbf{V}^l, \\ \mathbf{x}^{l+1} &\leftarrow \lambda \mathbf{x}^{l+1} + (1-\lambda)\mathbf{x}^l \end{aligned}$$

for some $\lambda \in (0, 1)$.

Exercise. Compute and plot the MPE. Experiment with different starting values. Do you find multiple MPE?

Exercise 1a: Learning-by-Doing Monopoly

The exercise below is inspired by the learning-by-doing model of Besanko, Doraszelski, Kryukov & Satterthwaite (2010).

The monopolist is characterized by its production experience $\omega \in \{1, \ldots, L\}$. The marginal cost of production, $c(\omega)$, depends on this stock of know-how. In particular, the monopolist faces a learning curve given by

$$c(\omega) = \begin{cases} \kappa\omega^{\eta} & \text{if} \quad 1 \leq \omega < l, \\ \kappa l^{\eta} & \text{if} \quad l \leq \omega \leq L, \end{cases}$$

where $\eta = \frac{\ln \rho}{\ln 2}$ for a learning curve with a slope of ρ percent, κ is the marginal cost with minimal know-how (normalized to be $\omega = 1$), and $l < L$ represents the stock of know-how at which the firm reaches the bottom of its learning curve.

The model is cast in discrete time and has an infinite horizon to avoid end effects. I first describe the product market and then turn to pricing dynamics. Unlike the quality ladder model, the learning-by-doing model cannot be broken down in a static and a dynamic part.

Product market. The industry draws its customers from a large pool of potential buyers. In each period, one buyer enters the market and makes, at most, on purchase. The utility consumer m derives from purchasing the product is $v - p + \epsilon_m$, where v represents the quality of the product and ϵ_m represents the buyer's idiosyncratic preference for the product. There is a no-purchase alternative with utility normalized to zero. I assume that the idiosyncratic shock ϵ_m is independently and identically logistically distributed across consumers. Moreover, the monopolist does not observe the buyer's idiosyncratic preference. The probability that the monopolist makes the sale therefore is

$$\Pr(q = 1) = D(p) = \frac{\exp(v - p)}{1 + \exp(v - p)}.$$

Pricing dynamics. The monopolist's state ω represents its know-how in the present period. Its know-how in the subsequent period, ω', depends on whether or not it makes a sale and on whether or not its stock of know-how depreciates.

The probability that the stock of know-how depreciates is $\Pr(f = 1) = \Delta(\omega) = 1 - (1 - \delta)^{\omega}$, where $\delta \in [0, 1]$. This specification is conceptually similar to the deterministic "capital-stock" models of depreciation employed in the empirical work on organizational forgetting (Argote, Beckman & Epple 1990, Argote & Epple 1990, Benkard 2004), where the depreciation of know-how increases as the firm's stock of know-how increases.

The law of motion for the monopolist's stock of know-how is

$$\omega' = \omega + q - f,$$

where $q \in \{0, 1\}$ indicates whether the monopolist makes a sale and $f \in \{0, 1\}$ represents organizational forgetting. The monopolist's stock of know-how therefore changes according the transition function

$$\Pr(\omega'|\omega,q) = \begin{cases} 1-\Delta(\omega) & \text{if} \quad \omega' = \omega + q, \\ \Delta(\omega) & \text{if} \quad \omega' = \omega + q - 1, \end{cases}$$

At the upper and lower boundaries of the state space, I take the transition function to be $\Pr(L|L, q = 1) = 1$ and $\Pr(1|1, q = 0) = 1$, respectively.

Let $V(\omega)$ denote the expected net present value of future cash flows to the monopolist if the current state is ω. The Bellman equation is

$$V(\omega) = \max_p D(p)(p - c(\omega)) + \beta \left[D(p)W_1(\omega) + (1 - D(p))W_0(\omega)\right],$$

where $\beta \in [0, 1)$ is the discount factor. The Bellman equation adds the current cash flow to the appropriately discounted expected future cash flow, where

$$\begin{aligned} W_1(\omega) &= \sum_{\omega'=1}^{L} V(\omega') \Pr(\omega'|\omega, q = 1), \\ W_0(\omega) &= \sum_{\omega'=1}^{L} V(\omega') \Pr(\omega'|\omega, q = 0) \end{aligned}$$

are the expectation of the value function conditional on winning and losing the sale, respectively. Note that the monopolist is allowed to price below marginal cost or even set a negative price.

The monopolist's pricing strategy is given by

$$p(\omega) = \arg\max_p D(p)(p - c(\omega)) + \beta \left[D(p)W_1(\omega) + (1 - D(p))W_0(\omega)\right].$$

Let $h(p) = D(p)(p - c(\omega)) + \beta \left[D(p)W_1(\omega) + (1 - D(p))W_0(\omega)\right]$ denote the maximand on the RHS of the Bellman equation. Differentiating with respect to p and making use of the assumed functional forms yields

$$\frac{\partial h}{\partial p} = D(p) \left[1 - (p - c(\omega)) - \beta W_1(\omega) + h(p)\right].$$

Differentiating this again and combining terms gives

$$\frac{\partial^2 h}{\partial p^2} = -\frac{\partial h}{\partial p}(1 - 2D(p)) - D(p).$$

Thus, $\frac{\partial h}{\partial p} = 0 \Rightarrow \frac{\partial^2 h}{\partial p^2} = -D(p) < 0$, i.e., the objective function is strictly quasi-concave and the price choice $p(\omega)$ is therefore unique. It is found by numerically solving $\frac{\partial h}{\partial p} = 0$ or, equivalently,

$$0 = 1 - (1 - D(p))[(p - c(\omega)) + \beta(W_1(\omega) - W_0(\omega))].$$

Parameterization. The number of know-how levels is $L = 30$, the slope of the learning curve is $\rho = 0.85$, the marginal cost of production with minimal know-how is $\kappa = 10$, the learning curve flattens out at $l = 15$ units of know-how, the quality of the good is $v = 10$, the depreciation probability is $\delta = 0.03$, and the discount factor is $\beta = \frac{1}{1.05}$, which corresponds to a yearly interest rate of 5%.

Exercise. Solve the monopolist's dynamic programming problem using value function iteration (or another method such as policy function iteration).

Exercise 2b: Learning-by-Doing Duopoly

This exercise introduces competition into the learning-by-doing model from Exercise 2a. Since many of the ingredients of the model remain the same, I focus on pointing out the differences between the dynamic game and the single-agent dynamic programming problem from a previous exercise.

There are two firms. Firm n is characterized by its productoin experience $\omega_n \in \{1, \ldots, L\}$. The state space is thus $\Omega \in \{1, \ldots, L\}^2$. I refer to $\omega = (\omega_1, \omega_2)$ as the state of the industry and to ω_n as the state of firm n.

Product market. The probability that firm n makes the sale is

$$D_n(p_1, p_2) = \frac{\exp(v - p_n)}{1 + \sum_{k=1}^{2} \exp(v - p_k)}.$$

Pricing dynamics. The Bellman equation for firm n is

$$V_n(\omega) = \max_{p_n} D_n(p_n, p_{-n}(\omega))(p_n - c(\omega_n)) + \beta \sum_{k=0}^{2} D_k(p_n, p_{-n}(\omega)) W_{nk}(\omega),$$

where $p_{-n}(\omega)$ is the rival's pricing strategy, and

$$\begin{aligned}
W_{n0}(\omega) &= \sum_{\omega_1'=1}^{L}\sum_{\omega_2'=1}^{L} V_n(\omega')\Pr(\omega_1'|\omega_1, q_1=0)\Pr(\omega_2'|\omega_2, q_2=0),\\
W_{n1}(\omega) &= \sum_{\omega_1'=1}^{L}\sum_{\omega_2'=1}^{L} V_n(\omega')\Pr(\omega_1'|\omega_1, q_1=1)\Pr(\omega_2'|\omega_2, q_2=0),\\
W_{n2}(\omega) &= \sum_{\omega_1'=1}^{L}\sum_{\omega_2'=1}^{L} V_n(\omega')\Pr(\omega_1'|\omega_1, q_1=0)\Pr(\omega_2'|\omega_2, q_2=1)
\end{aligned}$$

is the expectation of the value function of firm n conditional on the buyer purchasing good $k \in \{0,1,2\}$ (good 0 is the outside good).

The pricing strategy of firm n is given by

$$p_n(\omega) = \arg\max_{p_n} D_n(p_n, p_{-n}(\omega))(p_n - c(\omega_n)) + \beta \sum_{k=0}^{2} D_k(p_n, p_{-n}(\omega))W_{nk}(\omega).$$

Let $h_n(p_n) = D_n(p_n, p_{-n}(\omega))(p_n - c(\omega_n)) + \beta \sum_{k=0}^{2} D_k(p_n, p_{-n}(\omega))W_{nk}(\omega)$ denote the maximand on the RHS of the Bellman equation. $h_n(p_n)$ is strictly quasi-concave and the price choice $p_n(\omega)$ is therefore unique. It is found by numerically solving $\frac{\partial h_n}{\partial p_n} = 0$ or, equivalently,

$$\begin{aligned}
0 = 1 - (1 - D_n(p_n, p_{-n}(\omega)))\,(p_n - c(\omega_n)) - \beta W_{nn}(\omega)\\
+\beta \sum_{k=0}^{2} D_k(p_n, p_{-n}(\omega))W_{nk}(\omega).
\end{aligned}$$

Equilibrium. The primitives are symmetric. I therefore restrict attention to symmetric Markov perfect equilibria (MPE). Such a MPE is characterized by a value function $V(\omega)$ and a policy function $p(\omega)$ such that, if $V(\omega)$ is firm 1's value function, then firm 2's value function is given by $V_2(\omega_1, \omega_2) = V(\omega_2, \omega_1)$. Similarly, if $p(\omega)$ is firm 1's policy function, then firm 2's policy function is given by $p_2(\omega_1, \omega_2) = p(\omega_2, \omega_1)$. Existence of a symmetric MPE in pure strategies follows from the arguments in Doraszelski & Satterthwaite (2010) provided that prices are bounded.

Algorithm. To compute the MPE, Pakes & McGuire (1994) suggest an algorithm that essentially adapts value function iteration to dynamic games. The algorithm proceeds as follows:

1. Make initial guesses for the value and policy functions (or, more precisely, $L \times L$ matrices), $\mathbf{V}^0$ and $\mathbf{p}^0$, choose a stopping criterion $\epsilon > 0$, and initialize the iteration

counter to $l = 1$.

2. For all states $\omega \in \Omega$ compute

$$p^{l+1}(\omega) = \arg\max_{p_1} D_1(p_1, p_2^l(\omega_2, \omega_1))(p_1 - c(\omega_1)) + \beta \sum_{k=0}^{2} D_k(p_1, p_2^l(\omega_2, \omega_1)) W_k^l(\omega)).$$

and

$$V^{l+1}(\omega) = D_1(p_1^{l+1}(\omega), p_2^l(\omega_2, \omega_1))(p_1^{l+1}(\omega) - c(\omega_1)) + \beta \sum_{k=0}^{2} D_k(p_1^{l+1}(\omega), p_2^l(\omega_2, \omega_1)) W_n^l(\omega).$$

where

$$\begin{aligned}
W_0^l(\omega) &= \sum_{\omega_1'=1}^{L} \sum_{\omega_2'=1}^{L} V^l(\omega') \Pr(\omega_1'|\omega_1, q_1 = 0) \Pr(\omega_2'|\omega_2, q_2 = 0), \\
W_1^l(\omega) &= \sum_{\omega_1'=1}^{L} \sum_{\omega_2'=1}^{L} V^l(\omega') \Pr(\omega_1'|\omega_1, q_1 = 1) \Pr(\omega_2'|\omega_2, q_2 = 0), \\
W_2^l(\omega) &= \sum_{\omega_1'=1}^{L} \sum_{\omega_2'=1}^{L} V^l(\omega') \Pr(\omega_1'|\omega_1, q_1 = 0) \Pr(\omega_2'|\omega_2, q_2 = 1).
\end{aligned}$$

3. If

$$\max_{\omega \in \Omega} \left| \frac{V^{l+1}(\omega) - V^l(\omega)}{1 + |V^{l+1}(\omega)|} \right| < \epsilon \quad \wedge \quad \max_{\omega \in \Omega} \left| \frac{p^{l+1}(\omega) - p^l(\omega)}{1 + |p^{l+1}(\omega)|} \right| < \epsilon$$

then stop; else increment the iteration counter l by one and go to step 2.

Unlike value function iteration for single-agent dynamic programming problems, there is no guarantee that the above algorithm converges. If it fails to converge, a trick that often works is to go through an additional dampening step before returning to step 2. This dampening step assigns

$$\begin{aligned}
\mathbf{V}^{l+1} &\leftarrow \lambda \mathbf{V}^{l+1} + (1 - \lambda)\mathbf{V}^l, \\
\mathbf{p}^{l+1} &\leftarrow \lambda \mathbf{p}^{l+1} + (1 - \lambda)\mathbf{p}^l
\end{aligned}$$

for some $\lambda \in (0, 1)$.

Exercise. Compute and plot the MPE. Experiment with different starting values. Do you find multiple MPE?

References

Argote, L., Beckman, S. & Epple, D. (1990). The persistence and transfer of learning in an industrial setting, *Management Science* **36**(2): 140–154.

Argote, L. & Epple, D. (1990). Learning curves in manufacturing, *Science* **247**(4945): 920–924.

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

Besanko, D., Doraszelski, U., Kryukov, Y. & Satterthwaite, M. (2010). Learning-by-doing, organizational forgetting, and industry dynamics, *Econometrica* **78**(2): 453–508.

Doraszelski, U. & Satterthwaite, M. (2010). Computable Markov-perfect industry dynamics, *Rand Journal of Economics* **41**(2): 215–243.

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.