Learning-by-Doing, Organizational Forgetting, and Industry Dynamics
– Online Appendix –

David Besanko* Ulrich Doraszelski† Yaroslav Kryukov‡
Mark Satterthwaite§

February 8, 2007

1 Model: Organizational forgetting

Below we show that the expected stock of know-how in the absence of further learning is a decreasing convex function of time provided that $\Delta(e_n)$ is increasing in e_n.

Omitting firm subscripts to simplify the notation, let $\varphi(t) = E(e_t|e_0)$ be the expected stock of know-how in period t assuming that the initial stock of know-how is e_0 and that there is no further learning.

Proposition A1 *If $\Delta(e_t)$ is constant in e_t, then $\varphi(t)$ is a decreasing linear function of t. If $\Delta(e_t)$ is increasing in e_t, then $\varphi(t)$ is a decreasing convex function of t.*

Proof. In the absence of further learning, $q_t = 0$ and the evolution of the stock of know-how is governed by the law of motion

$$e_{t+1} = e_t - f_t.$$

Taking expectations (conditional on e_t) gives us

$$E(e_{t+1}|e_t) = e_t - E(f_t|e_t) = e_t - \Delta(e_t).$$

Since for any two random variables X and Y, $E_Y(E_X(X|Y)) = E_X(X)$, we can take expectations (conditional on e_0) on both sides of the above equation to obtain

$$E(e_{t+1}|e_0) = E(e_t|e_0) - E(\Delta(e_t)|e_0).$$

*Kellogg School of Management, Northwestern University, Evanston, IL 60208, d-besanko@kellogg.northwestern.edu.
†Department of Economics, Harvard University, Cambridge, MA 02138, doraszelski@harvard.edu.
‡Department of Economics, Northwestern University, Evanston, IL 60208, kryukov@northwestern.edu.
§Kellogg School of Management, Northwestern University, Evanston, IL 60208, m-satterthwaite@kellogg.northwestern.edu.

This implies

$$\varphi(1) = e_0 - \Delta(e_0),$$
$$\varphi(t+1) - \varphi(t) = -E(\Delta(e_t)|e_0), \quad t \geq 1.$$

Since $-E(\Delta(e_t)|e_0) < 0$, $\varphi(t)$ is a decreasing function of t. Let $\Delta\varphi(t) = \varphi(t) - \varphi(t-1)$ be its slope, so that

$$\Delta\varphi(t+1) - \Delta\varphi(t) = E(\Delta(e_{t-1})|e_0) - E(\Delta(e_t)|e_0).$$

If $\Delta(e_t)$ is constant in e_t, then $\Delta\varphi(t+1) - \Delta\varphi(t) = 0$, and $\varphi(t)$ is a linear function of t. If, by contrast, $\Delta(e_t)$ is increasing in e_t, then $\Delta\varphi(t+1) - \Delta\varphi(t) > 0$ because the distribution of e_{t-1} stochastically dominates the distribution of e_t in the absence of further learning. Thus, $\varphi(t)$ is a convex function of t. ■

2 Model: Parameterization

Below we show how to map the empirical estimates of rates of depreciation into in our specification.

Empirical work on organizational forgetting employs a capital-stock model. This model is defined by the deterministic law of motion

$$e_n' = (1 - \xi)e_n + y_n,$$

where ξ is the rate of depreciation and y_n is the flow of orders. If the flow of orders is equal to a constant y, the steady-state stock of know-how is

$$\frac{y}{\xi}.$$

Recall our stochastic law of motion:

$$e_n' = e_n + q_n - f_n.$$

Taking expectations yields

$$E(e_n'|e_n) = e_n + \gamma - \Delta(e_n),$$

where $\gamma = \Pr(q_n = 1)$ is the probability that the firm makes a sale and $\Delta(e_n) = \Pr(f_n = 1) = 1 - (1-\delta)^{e_n}$ is the probability that it loses a unit of know-how through organizational forgetting. The steady-state stock of know-how is

$$\frac{\ln(1-\gamma)}{\ln(1-\delta)}.$$

We now ask what is the value of the forgetting rate δ so that the two specifications generate the same steady-state stock of know-how? The answer is given by

$$\frac{y}{\xi} = \frac{\ln(1-\gamma)}{\ln(1-\delta)}$$

or, equivalently,

$$\delta = 1 - (1-\gamma)^{\frac{\xi}{y}}.$$

To illustrate, consider Benkard's (2000) empirical analysis of organizational forgetting in the production of wide-bodied airframes. There were 250 L-1011 aircraft produced over a 14 year period. Assuming a smooth flow of orders, this implies $y = 1.5$ units per month. Benkard (2000) estimates a rate of depreciation of $\xi = 4$ percent per month. This implies that the steady-state stock of know-how is equal to $\frac{1.5}{0.04} = 37.50$ units. Matching steady states implies that δ falls in the range between 0.0014 and 0.077 as γ ranges between 0.05 and 0.95, with $\delta = 0.018$ when $\gamma = 0.5$.

3 Computation: Pakes & McGuire (1994) algorithm

In this section we first relate the Pakes & McGuire (1994) algorithm to our homotopy algorithm. Then we discuss in more detail the extent and source of the convergence failure of the Pakes & McGuire (1994) algorithm.

3.1 Relationship of Jacobians

To relate the Pakes & McGuire (1994) algorithm to our homotopy algorithm, recall that $\mathbf{x} = (\mathbf{V}^*, \mathbf{p}^*)$ and consider a parametric path $(\mathbf{x}(s), \delta(s)) \in \mathbf{F}^{-1}$ in the equilibrium correspondence.

Proposition A2 *Let $(\mathbf{x}(s), \delta(s)) \in \mathbf{F}^{-1}$. We have*

$$\left.\frac{\partial \mathbf{G}(\mathbf{x}(s))}{\partial \mathbf{x}}\right|_{\delta=\delta(s)} = \frac{\partial \mathbf{F}(\mathbf{x}(s), \delta(s))}{\partial \mathbf{x}} + \mathbf{I},$$

where $\mathbf{I}$ denotes the $(2M^2 \times 2M^2)$ identity matrix.

Homotopy algorithm. Before proving the proposition, we provide some notation. Recall that the homotopy algorithm searches for a zero of $\mathbf{F}(\cdot)$, the collection of equations (6) in the main paper that defines an equilibrium. Further recall that the return function

$$h_n(\mathbf{e}, p_n, p_{-n}(\mathbf{e}), \mathbf{V}_n) = D_n(p_n, p_{-n}(\mathbf{e}))\,(p_n - c(e_n)) + \beta \sum_{k=1}^{2} D_k(p_n, p_{-n}(\mathbf{e}))\overline{V}_{nk}(\mathbf{e})$$

is defined as the maximand in the Bellman equation (1) in the main paper. It is convenient to reformulate $F^1_{\mathbf{e}}(\mathbf{x},\delta)$ and $F^2_{\mathbf{e}}(\mathbf{x},\delta)$ in equations (4) and (5) in the main paper as

$$F^1_{\mathbf{e}}(\mathbf{x},\delta) = -V^*(\mathbf{e}) + h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*),$$

$$F^2_{\mathbf{e}}(\mathbf{x},\delta) = \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} \Bigg/ \frac{1}{\sigma} D_1(p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}))$$
$$= q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*),$$

where

$$q_n(\mathbf{e}, p_n, p_{-n}(\mathbf{e}), \mathbf{V}_n)$$
$$= \sigma - (1 - D_n(p_n, p_{-n}(\mathbf{e})))\,(p_n - c(e_n)) - \beta \overline{V}_{nn}(\mathbf{e}) + \beta \sum_{k=1}^{2} D_k(p_n, p_{-n}(\mathbf{e})) \overline{V}_{nk}(\mathbf{e}), \quad \text{(A1)}$$

and we make explicit that $F^1_{\mathbf{e}}(\mathbf{x},\delta)$ and $F^2_{\mathbf{e}}(\mathbf{x},\delta)$ depend on δ.

Pakes & McGuire (1994) algorithm. Recall that the Pakes & McGuire (1994) algorithm searches for a fixed point $\mathbf{x} = (\mathbf{V}^*, \mathbf{p}^*)$ of $\mathbf{G}(\cdot)$, the collection of equations (15) in the main paper that maps old guesses for the value and policy functions of firm 1 into new guesses. Again it is convenient to reformulate $G^2_{\mathbf{e}}(\mathbf{x})$ and $G^1_{\mathbf{e}}(\mathbf{x})$ in equations (13) and (14) in the main paper as

$$G^2_{\mathbf{e}}(\mathbf{x}) = \arg\max_{p_1} h_1(\mathbf{e}, p_1, p^*(\mathbf{e}^{[2]}), \mathbf{V}^*) = \left\{ p_1 \,\middle|\, \frac{\partial h_1(\mathbf{e}, p_1, p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} = 0 \right\}$$
$$= \left\{ p_1 \,\middle|\, q_1(\mathbf{e}, p_1, p^*(\mathbf{e}^{[2]}), \mathbf{V}^*) = 0 \right\},$$
$$G^1_{\mathbf{e}}(\mathbf{x}) = \max_{p_1} h_1(\mathbf{e}, p_1, p^*(\mathbf{e}^{[2]}), \mathbf{V}^*) = h_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*),$$

where, recall, $G^2_{\mathbf{e}}(\mathbf{x})$ is uniquely determined because $h_1(\cdot)$ is strictly quasi-concave in p_1.

Proof of Proposition A2. Let $(\mathbf{x},\delta) \in \mathbf{F}^{-1}$, where we suppress the dependence of $\mathbf{x}$ and δ on s to simplify the notation. Then

$$F^1_{\mathbf{e}}(\mathbf{x},\delta) = 0, \quad \text{(A2)}$$
$$F^2_{\mathbf{e}}(\mathbf{x},\delta) = 0 \quad \text{(A3)}$$

because the equilibrium is a zero of $\mathbf{F}(\cdot)$ and also

$$V^*(\mathbf{e}) = G^1_{\mathbf{e}}(\mathbf{x}), \tag{A4}$$

$$p^*(\mathbf{e}) = G^2_{\mathbf{e}}(\mathbf{x}) \tag{A5}$$

because the equilibrium is a fixed point of $\mathbf{G}(\cdot)$.

Letting x_i denote the ith element of $\mathbf{x}$, we have to show that

$$\begin{aligned} \frac{\partial G^1_{\mathbf{e}}(\mathbf{x})}{\partial x_i} &= \frac{\partial F^1_{\mathbf{e}}(\mathbf{x},\delta)}{\partial x_i} + 1(x_i = V^*(\mathbf{e})), \quad i = 1,\dots,2M^2, \\ \frac{\partial G^2_{\mathbf{e}}(\mathbf{x})}{\partial x_i} &= \frac{\partial F^2_{\mathbf{e}}(\mathbf{x},\delta)}{\partial x_i} + 1(x_i = p^*(\mathbf{e})), \quad i = 1,\dots,2M^2 \end{aligned}$$

for all states $\mathbf{e} \in \{1,\dots,M\}^2$.

Case (i): Consider first $F^1_{\mathbf{e}}(\mathbf{x},\delta)$ and $G^1_{\mathbf{e}}(\mathbf{x})$ for an arbitrary state $\mathbf{e} \in \{1,\dots,M\}^2$. In what follows we repeatedly use the fact that equation (A3) implies

$$\frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} = 0. \tag{A6}$$

Firm's price: If $e_1 \neq e_2$, then we have

$$\frac{\partial F^1_{\mathbf{e}}(\mathbf{x},\delta)}{\partial p^*(\mathbf{e})} = \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} = 0$$

because of equation (A6) and

$$\frac{\partial G^1_{\mathbf{e}}(\mathbf{x})}{\partial p^*(\mathbf{e})} = \frac{\partial h_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} \frac{\partial G^2_{\mathbf{e}}(\mathbf{x})}{\partial p^*(\mathbf{e})} = 0$$

because of equations (A5) and (A6).

Both prices: If $e_1 = e_2$, then $p^*(\mathbf{e}) = p^*(\mathbf{e}^{[2]})$ and we have

$$\begin{aligned} \frac{\partial F^1_{\mathbf{e}}(\mathbf{x},\delta)}{\partial p^*(\mathbf{e})} &= \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} + \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_2} \\ &= \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_2} \end{aligned}$$

because of equation (A6) and

$$\begin{aligned} \frac{\partial G^1_{\mathbf{e}}(\mathbf{x})}{\partial p^*(\mathbf{e})} &= \frac{\partial h_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} \frac{\partial G^2_{\mathbf{e}}(\mathbf{x})}{\partial p^*(\mathbf{e})} + \frac{\partial h_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_2} \\ &= \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_2} \end{aligned}$$

because of equations (A5) and (A6).

Other: If $x_i \neq p^*(\mathbf{e})$, then we have

$$\frac{\partial F^1_{\mathbf{e}}(\mathbf{x}, \delta)}{\partial x_i} = -1(x_i = V^*(\mathbf{e})) + \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial x_i},$$

where $-1(x_i = V^*(\mathbf{e}))$ is the derivative of $-V^*(\mathbf{e})$ with respect to x_i (with $1(\cdot)$ being the indicator function), and

$$\begin{aligned}\frac{\partial G^1_{\mathbf{e}}(\mathbf{x})}{\partial x_i} &= \frac{\partial h_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1}\frac{\partial G^2_{\mathbf{e}}(\mathbf{x})}{\partial x_i} + \frac{\partial h_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial x_i} \\ &= \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial x_i}\end{aligned}$$

because of equations (A5) and (A6).

Case (ii): Next consider $F^2_{\mathbf{e}}(\mathbf{x}, \delta)$ and $G^2_{\mathbf{e}}(\mathbf{x})$ for an arbitrary state $\mathbf{e} \in \{1, \ldots, M\}^2$. In what follows we repeatedly use the fact that equation (A3) implies

$$\frac{\partial q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), V^*)}{\partial p_1} = -1 + \frac{\partial h_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), V^*)}{\partial p_1} = -1. \tag{A7}$$

Firm's price: If $e_1 \neq e_2$, then we have

$$\frac{\partial F^2_{\mathbf{e}}(\mathbf{x}, \delta)}{\partial p^*(\mathbf{e})} = \frac{\partial q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} = -1$$

because of equation (A7) and

$$\frac{\partial G^2_{\mathbf{e}}(\mathbf{x})}{\partial p^*(\mathbf{e})} = 0$$

because the construction of $G^2_{\mathbf{e}}(\mathbf{x})$ does not depend on $p^*(\mathbf{e})$.

Both prices: If $e_1 = e_2$, then $p^*(\mathbf{e}) = p^*(\mathbf{e}^{[2]})$ and we have

$$\begin{aligned}\frac{\partial F^2_{\mathbf{e}}(\mathbf{x}, \delta)}{\partial p^*(\mathbf{e})} &= \frac{\partial q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_1} + \frac{\partial q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_2} \\ &= -1 + \frac{\partial q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_2}\end{aligned}$$

because of equation (A7) and

$$\begin{aligned}\frac{\partial G^2_{\mathbf{e}}(\mathbf{x})}{\partial p^*(\mathbf{e})} &= -\frac{\partial q_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)/\partial p_2}{\partial q_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)/\partial p_1} \\ &= \frac{\partial q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial p_2}\end{aligned}$$

because the construction of $G^2_{\mathbf{e}}(\mathbf{x})$ does not depend on $p^*(\mathbf{e})$, the implicit function theorem, and equations (A5) and (A7).

Other: If $x_i \neq p^*(\mathbf{e})$, then we have

$$\frac{\partial F^2_{\mathbf{e}}(\mathbf{x}, \delta)}{\partial x_i} = \frac{\partial q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial x_i}$$

and

$$\begin{aligned} \frac{\partial G^2_{\mathbf{e}}(\mathbf{x})}{\partial x_i} &= -\frac{\partial q_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)/\partial x_i}{\partial q_1(\mathbf{e}, G^2_{\mathbf{e}}(\mathbf{x}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)/\partial p_1} \\ &= \frac{\partial q_1(\mathbf{e}, p^*(\mathbf{e}), p^*(\mathbf{e}^{[2]}), \mathbf{V}^*)}{\partial x_i} \end{aligned}$$

because of the implicit function theorem and equations (A5) and (A7). ■

3.2 Extent of convergence failure

Next we illustrate the extent of the convergence failure of the Pakes & McGuire (1994) algorithm. Figure A1 summarizes Proposition 1 and Result 1 in the main paper by marking equilibria with $\rho\left(\frac{\partial \mathbf{G}(\mathbf{x}(s))}{\partial \mathbf{x}}\Big|_{\delta=\delta(s)}\right) \geq 1$ using a dotted line and equilibria with $\rho\left(\frac{\partial \mathbf{G}(\mathbf{x}(s))}{\partial \mathbf{x}}\Big|_{\delta=\delta(s)}\right) < 1$ using a solid line. The former cannot be computed by the Pakes & McGuire (1994) algorithm.

3.3 Source of convergence failure

Finally we explore the source of the convergence failure of the Pakes & McGuire (1994) algorithm. We show that, holding fixed the value of continued play, the best reply dynamics are contractive and therefore converge to a unique fixed point irrespective of the initial guess. In addition, we show that the value function iteration also is contractive holding fixed the policy function.

Best reply dynamics. Defining

$$\mathbf{G}^2(\mathbf{p}; \mathbf{V}) = \begin{pmatrix} G^2_{(1,1)}(\mathbf{V}, \mathbf{p}) \\ G^2_{(2,1)}(\mathbf{V}, \mathbf{p}) \\ \vdots \\ G^2_{(M,M)}(\mathbf{V}, \mathbf{p}) \end{pmatrix}$$

we write the Pakes & McGuire (1994) algorithm with the value function held fixed as

$$\mathbf{p}^{k+1} = \mathbf{G}^2(\mathbf{p}^k; \mathbf{V}), \quad k = 0, 1, 2, \ldots.$$

The following proposition establishes that $\mathbf{G}^2$ is a contraction. This implies that the best reply dynamics converge to a unique fixed point irrespective of the initial guess.

Proposition A3 *Holding fixed* $\mathbf{V} \in [\check{V}, \hat{V}]^{M^2}$, *where* $-\infty < \check{V} \leq \hat{V} < \infty$, $\mathbf{G}^2$ *is a contraction.*

Proof. Recall that $G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p})$ is the solution to the equation

$$q_1(\mathbf{e}, p_1, p(\mathbf{e}^{[2]}), \mathbf{V}) = 0, \tag{A8}$$

where $q_1(\cdot)$ is defined in equation (A1). To avoid having to deal with corner solutions, pick $-\infty < \check{p} \leq \hat{p} < \infty$ large enough so that $\mathbf{G}^2$ maps $[\check{p}, \hat{p}]^{M^2}$ into itself. Note that $[\check{p}, \hat{p}]^{M^2}$ is convex and that $\mathbf{G}^2$ is continuously differentiable. Moreover, since $G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p})$ is the solution to equation (A8), it is straightforward to show using the implicit function theorem that the entries of the Jacobian $\frac{\partial \mathbf{G}^2(\mathbf{p};\mathbf{V})}{\partial \mathbf{p}}$ are generated by

$$\frac{\partial G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p})}{\partial p(\mathbf{e}^{[2]})} = \frac{D_2(G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p}), p(\mathbf{e}^{[2]}))}{\sigma}$$
$$\times \left(D_1(G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p}), p(\mathbf{e}^{[2]})) \left(G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p}) - c(e_1)\right) - \beta \overline{V}_2(\mathbf{e}) + \beta \sum_{k=1}^{2} D_k(G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p}), p(\mathbf{e}^{[2]})) \overline{V}_k(\mathbf{e}) \right). \tag{A9}$$

It is helpful to re-write equation (A9): Since $G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p})$ is the solution to equation (A8), we have

$$G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p}) - c(e_1) = \frac{1}{1 - D_1(G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p}), p(\mathbf{e}^{[2]}))} \left(\sigma - \beta \overline{V}_1(\mathbf{e}) + \beta \sum_{k=1}^{2} D_k(G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p}), p(\mathbf{e}^{[2]})) \overline{V}_k(\mathbf{e}) \right).$$

Substituting into equation (A9) and simplifying yields

$$\frac{\partial G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p})}{\partial p(\mathbf{e}^{[2]})} = D_1(G^2_{\mathbf{e}}(\mathbf{V}, \mathbf{p}), p(\mathbf{e}^{[2]})) \in [D_1(\hat{p}, \check{p}), D_1(\check{p}, \hat{p})] \subseteq (0, 1).$$

The rest of the proof is a minor modification of the proof of Proposition 1.10 in Section 3.1 of Bertsekas & Tsitsiklis (1997). (In their notation set $m = 1$ and $g(t) = \mathbf{G}^2(t\mathbf{p}^\dagger + (1-t)\mathbf{p}; \mathbf{V})$ to show that $\left|\left|\mathbf{G}^2(\mathbf{p}^\dagger; \mathbf{V}) - \mathbf{G}^2(\mathbf{p}; \mathbf{V})\right|\right|_\infty = ||g(1) - g(0)||_\infty \leq \alpha \left|\left|\mathbf{p}^\dagger - \mathbf{p}\right|\right|_\infty$ with $\alpha = D_1(\check{p}, \hat{p}) < 1$.) ■

Value function iteration. Defining

$$\mathbf{G}^1(\mathbf{V};\mathbf{p}) = \begin{pmatrix} G^1_{(1,1)}(\mathbf{V},\mathbf{p}) \\ G^1_{(2,1)}(\mathbf{V},\mathbf{p}) \\ \vdots \\ G^1_{(M,M)}(\mathbf{V},\mathbf{p}) \end{pmatrix}$$

we write the Pakes & McGuire (1994) algorithm with the policy function held fixed as

$$\mathbf{V}^{k+1} = \mathbf{G}^1(\mathbf{V}^k;\mathbf{p}), \quad k = 0,1,2,\ldots.$$

The following proposition establishes that $\mathbf{G}^1$ is a contraction, so that the value function iteration converges.

Proposition A4 *Holding fixed* $\mathbf{p} \in [\check{p},\hat{p}]^{M^2}$*, where* $-\infty < \check{p} \leq \hat{p} < \infty$*,* $\mathbf{G}^1$ *is a contraction.*

Proof. Recall that

$$G^1_{\mathbf{e}}(\mathbf{V},\mathbf{p}) = \max_{p_1} h_1(\mathbf{e}, p_1, p(\mathbf{e}^{[2]}), \mathbf{V}).$$

Pick $-\infty < \check{V} \leq \hat{V} < \infty$ large enough so that $\mathbf{G}^1$ maps $[\check{V},\hat{V}]^{M^2}$ into itself.[1] The proof is completed by applying Blackwell's sufficient conditions (monotonicity and discounting, see e.g. p. 54 of Stokey & Lucas (1989)) to show that $\mathbf{G}^1$ is a contraction. ∎

4 Equilibrium correspondence

The value functions in Figure A2 correspond to the policy functions in Figure 4 in the main paper. The smooth value functions in the upper panels are typical for flat equilibria. While its value function is increasing in a firm's state, it is not decreasing by too much in its rival's state. Turning to the trenchy and extra-trenchy equilibria, the value functions in the lower panels are much less smooth. Both the leader and the follower experience a rise in value as the industry moves from a state on the diagonal of the state space with extremely intense price competition to an asymmetric state. In other words, the diagonal trench in the policy function is mirrored by a diagonal trench in the value function. Further, in an extra-trenchy equilibrium, the value of being a clear leader is very high while the value of being a distant follower is very low.

5 Robustness checks: Product differentiation

Figure A3 displays the limiting and maximum expected Herfindahl indices for the case of weaker product differentiation with $\sigma = 0.2$. Figures A4 and A5 do the same for the case

[1] For example, if $\check{V}$ and $\hat{V}$ solve $\check{V} = \min_{p_1\in[\check{p},\hat{p}],p_2\in[\check{p},\hat{p}]} D_1(p_1,p_2)(p_1 - c(M)) + \beta\check{V}$ and $\hat{V} = \max_{p_1\in[\check{p},\hat{p}],p_2\in[\check{p},\hat{p}]} D_1(p_1,p_2)(p_1 - c(1)) + \beta\hat{V}$, then $\check{V} \leq G^1_{\mathbf{e}}(\mathbf{V},\mathbf{p}) \leq \hat{V}$.

σ	ρ	0.95	0.85	0.75	0.65	0.55	0.35	0.15	0.05
0.2	$\bar{\delta}(\rho)$	0.65	0.88	0.91	0.96	0.98	0.99	0.99	1.00
1	$\bar{\delta}(\rho)$	0.57	0.60	0.62	0.71	0.78	0.81	0.88	0.90
2	$\bar{\delta}(\rho)$	0.56	0.57	0.60	0.61	0.63	0.69	0.73	0.76
10	$\bar{\delta}(\rho)$	0.55	0.55	0.56	0.56	0.57	0.57	0.58	0.59

Table A1: Product differentiation with $\sigma \in \{0.2, 1, 2, 10\}$. Critical value $\bar{\delta}(\rho)$ for investment stifling.

$v_0 - c_0$	ρ	0.95	0.85	0.75	0.65	0.55	0.35	0.15	0.05
$-\infty$	$\bar{\delta}(\rho)$	0.55	0.60	0.62	0.71	0.78	0.81	0.88	0.90
0	$\bar{\delta}(\rho)$	0.17	0.38	0.56	0.67	0.74	0.83	0.86	0.88
3	$\bar{\delta}(\rho)$	0.01	0.08	0.23	0.39	0.53	0.71	0.81	0.83
5	$\bar{\delta}(\rho)$	0.00	0.01	0.06	0.15	0.29	0.49	0.67	0.68
10	$\bar{\delta}(\rho)$	0	0	0.00	0.00	0.01	0.01	0.02	0.03

Table A2: Outside good with $v_0 - c_0 \in \{-\infty, 0, 3, 5, 10\}$. Critical value $\bar{\delta}(\rho)$ for investment stifling.

of stronger product differentiation with $\sigma \in \{2, 10\}$. These figures may be compared to our baseline parameterization with $\sigma = 1$ in Figure 3 in the main paper.

Table A1 documents the investment-stifling effect of organizational forgetting. If δ exceeds the critical value $\bar{\delta}(\rho)$ listed in Table A1, then firms cannot expect to make it down their learning curves. As can be seen, investment stifling sets in for ever lower forgetting rates as the degree of horizontal product differentiation becomes higher.

6 Robustness checks: Outside good

Figure A6 illustrates the extent of multiplicity for the case of an outside good with $v_0 - c_0 = 0$ It shows the number of equilibria for each combination of forgetting rate δ and progress ratio ρ. Darker shades indicate more equilibria. As can be seen, we have found up to nine equilibria for some values of δ and ρ. We no longer have sunspots for a progress ratio of $\rho = 1$. This figure may be compared to our baseline parameterization with $v_0 - c_0 = -\infty$ in Figure 2 in the main paper.

Figure A7 displays the limiting and maximum expected Herfindahl indices for the case an outside good with $v_0 - c_0 = 0$. Figures A8–A10 do the same for the case of a more attractive outside good with $v_0 - c_0 \in \{3, 5, 10\}$. These figures may be compared to our baseline parameterization with $v_0 - c_0 = -\infty$ in Figure 3 in the main paper.

Table A2 documents the investment-stifling effect of organizational forgetting. If δ exceeds the critical value $\bar{\delta}(\rho)$ listed in Table A2, then firms cannot expect to make it down their learning curves. As can be seen, investment stifling sets in for ever lower forgetting rates as the outside good becomes more attractive.

7 Robustness checks: Choke price

Figure A11 displays the limiting and maximum expected Herfindahl indices for logit demand (right panels) and linear demand (left panel) and various degrees of product differentiation. Note that the horizontal axis is the progress ratio ρ.

8 Robustness checks: Frequency of sales

Figure A12 exemplifies the policy functions of the typical equilibria. Figures A13 and A14 display the transient distribution in period 8 and 32 (subperiod 16 and 64), respectively, and Figure A15 displays the limiting distribution for the four typical cases. The parameter values are $\rho = 0.85$ and $\delta \in \{0, 0.02, 0.09\}$.

9 Robustness checks: Learning-by-doing

Figure A16 displays the limiting and maximum expected Herfindahl indices for the bottomless learning specification with $m = 30 = M$. This figure may be compared to our baseline parameterization with $m = 15 < M$ in Figure 3 in the main paper.

Figure A17 provides another example of a *plateau equilibrium*. It displays the policy function (upper left panel), the transient distribution in period 8 and 32 (upper right and lower left panels), and the limiting distribution (lower right panel). The parameter values are $\rho = 0.9$ and $\delta = 0.04$.

10 Robustness checks: Organizational forgetting

Figure A18 illustrates the extent of multiplicity for the constant forgetting specification with $\Delta(e_n) = \delta$.[2] It shows the number of equilibria for each combination of forgetting rate δ and progress ratio ρ. Darker shades indicate more equilibria. As can be seen, we have found up to eleven equilibria for some values of δ and ρ. Multiplicity is especially pervasive for forgetting rates δ between 0.4 and 0.5. Note that the horizontal axis is on a linear scale.

Figure A19 displays the limiting and maximum expected Herfindahl indices for the constant forgetting specification. Note that the horizontal axis is on a linear scale.

11 Robustness checks: Entry and exit

Below we describe the N-firm version of our model with entry and exit.

[2] Figure A18 looks somewhat rough because we use a grid of 20 rather than 100 values of $\rho \in (0, 1]$.

Order of moves. In each period the sequence of events is as follows:

1. Each of the N^* incumbent firms learns its own salvage value and makes an exit decision. Each of the $N - N^*$ potential entrants learns its own set-up cost and makes an entry decision. Entry and exit decisions are made simultaneously. In this process, the industry transits from state $\mathbf{e}$ to state $\mathbf{e}'$. Specifically, incumbent firm n transits from state $e_n \neq 0$ to state $e'_n = 0$ upon exiting and potential entrant n transits from state $e_n = 0$ to state $e'_n = e^0 \neq 0$ upon entering the industry.

2. Price competition takes place among active firms, where firm n is active if and only if $e'_n \neq 0$. Learning-by-doing and organizational forgetting occur. In this process, the industry transits from state $\mathbf{e}'$ to state $\mathbf{e}''$.

Before making their entry and exit decisions, all firms observe state $\mathbf{e}$, and all firms observe state $\mathbf{e}'$ prior to making their pricing decisions.

Entry and exit. Before price competition takes place, incumbent firms can choose to exit the industry and potential entrants can choose to enter it. If an incumbent firm exits the industry, it receives a salvage value and perishes. We assume that at the beginning of each period, incumbent firm n draws a salvage value X_n from a uniform distribution $G_X(\cdot)$ with support $[\overline{X} - a, \overline{X} + a]$, where $a > 0$ is a parameter. Salvage values are independently and identically distributed across firms and periods, and firm n's realization is observed only by itself but not by its rivals. Let $\tau_n(\mathbf{e}, X_n) = 1$ denote the decision of incumbent firm n to remain in the industry in state $\mathbf{e}$ when it has drawn salvage value X_n, while $\tau_n(\mathbf{e}, X_n) = 0$ denotes the decision to exit.

Simultaneous with the exit decisions of incumbent firms, potential entrants make entry decisions. If a potential entrant decides not to enter, it receives nothing and perishes; if it enters, it incurs a set-up cost. At the beginning of each period, potential entrant n draws a set-up cost S_n from a uniform distribution $G_S(\cdot)$ with support $[\overline{S} - b, \overline{S} + b]$, where $b > 0$ is a parameter. Set-up costs are independently and identically distributed across firms and periods, and its realization is private to a firm. Let $\tau_n(\mathbf{e}, S_n) = 1$ denote the decision of potential entrant n to enter the industry in state $\mathbf{e}$ when it has drawn set-up cost S_n, while $\tau_n(\mathbf{e}, S_n) = 0$ denotes the decision to stay out.

Combining the firms' entry and exit decisions, let $\lambda_n(\mathbf{e})$ denote the probability that firm n operates in the industry in state $\mathbf{e}$. If $e_n \neq 0$ so that firm n is an incumbent, then $\lambda_n(\mathbf{e}) = \int \tau_n(\mathbf{e}, X_n) dG_X(X_n)$. If $e_n = 0$ so that firm n is an entrant, then $\lambda_n(\mathbf{e}) = \int \tau_n(\mathbf{e}, S_n) dG_S(S_n)$.

Bellman equation. To develop the Bellman equation, we first consider firms' pricing decisions. We then consider the exit decisions of incumbent firms and the entry decisions of potential entrants. Throughout we use $V_n(\mathbf{e})$ to denote the expected net present value of future cash flows to firm n in state $\mathbf{e}$ *before* entry and exit decisions have been made. In addition, we use $U_n(\mathbf{e}')$ to denote the expected net present value of future cash flows to active firm n in state $\mathbf{e}'$ *after* entry and exit decisions have been made.

Pricing decisions. Consider an industry that, via a process of entry and exit, has transitioned from state $\mathbf{e}$ to state $\mathbf{e}'$. The expected net present value of future cash flows to active firm n is given by

$$U_n(\mathbf{e}') = \max_{p_n} D_n(p_n, \mathbf{p}_{-n}(\mathbf{e}'))(p_n - c(e'_n)) + \beta \sum_{k=0}^{N} D_k(p_n, \mathbf{p}_{-n}(\mathbf{e}'))\overline{V}_{nk}(\mathbf{e}'), \tag{A10}$$

where $p_{-n}(\mathbf{e}')$ denotes the prices charged by the other firms in state $\mathbf{e}'$ and $\overline{V}_{nk}(\mathbf{e})$ is the expectation of firm n's value function conditional on the buyer purchasing good $k \in \{0, 1, \ldots, N\}$ as given by

$$\begin{aligned}
\overline{V}_{n0}(\mathbf{e}') &= \sum_{e''_1=e'_1-1}^{e'_1} \cdots \sum_{e''_N=e'_N-1}^{e'_N} V_n(\mathbf{e}'') \prod_{i=1}^{N} \Pr(e''_i|e'_i, 0), \\
\overline{V}_{nk}(\mathbf{e}') &= \sum_{e''_1=e_1-1}^{e'_1} \cdots \sum_{e''_{k-1}=e'_{k-1}-1}^{e'_{k-1}} \sum_{e''_k=e'_k}^{e'_k+1} \sum_{e''_k=e'_{k+1}-1}^{e'_{k+1}} \cdots \sum_{e''_N=e'_N-1}^{e'_N} V_n(\mathbf{e}'') \\
&\quad \prod_{j \neq k} \Pr(e''_j|e'_j, 0) \Pr(e''_k|e'_k, 1), \quad k \in \{1, \ldots, N\}.
\end{aligned}$$

Note that we include an outside good (good 0) in the specification to ensure a well-posed monopoly problem.

Let $h_n(\mathbf{e}', p_n, \mathbf{p}_{-n}(\mathbf{e}'), \mathbf{V}_n)$ denote the maximand in equation (A10). Using the same argument as in Section 2 in the main paper, if the FOC $\frac{\partial h_n(\cdot)}{\partial p_n} = 0$ is satisfied, then $\frac{\partial^2 h_n(\cdot)}{\partial p_n^2} = -\frac{1}{\sigma} D_n(p_n, p_{-n}(\mathbf{e}')) < 0$. The return function $h_n(\cdot)$ is therefore strictly quasi-concave in p_n, so that the pricing decision $p_n(\mathbf{e}')$ is uniquely determined by the solution to the FOC (given $\mathbf{p}_{-n}(\mathbf{e}')$). If firm n is inactive, we assign $p_n(\mathbf{e}') = \infty$.

Exit decisions. To develop the Bellman equation determining $V_n(\mathbf{e})$, consider the exit decision $\tau_n(\mathbf{e}, X_n)$ of incumbent firm n who has drawn salvage value X_n. It remains in the industry in state $\mathbf{e}$ if its realized salvage value is less than or equal to the expected value of continuing forward to the price-setting stage:

$$\tau_n(\mathbf{e}, X_n) = \begin{cases} 1 & \text{if} \quad X_n \leq \widehat{X}_n(\mathbf{e}), \\ 0 & \text{if} \quad X_n \geq \widehat{X}_n(\mathbf{e}), \end{cases}$$

where

$$\widehat{X}_n(\mathbf{e}) = E[U_n(\mathbf{e}')|\mathbf{e}, e'_n = e_n, \lambda_{-n}(\mathbf{e})]$$

is the expected value to incumbent firm n of continuing forward to the price-setting stage as an active firm with its current stock of know-how, i.e., $e'_n = e_n$, taking into account the operating probabilities $\lambda_{-n}(\mathbf{e})$ of the other firms. $\widehat{X}_n(\mathbf{e})$ is computed as

$$\sum_{e'_1 \in \mathcal{E}'_1} \cdots \sum_{e'_{n-1} \in \mathcal{E}'_{n-1}} \sum_{e'_{n+1} \in \mathcal{E}'_{n+1}} \cdots \sum_{e'_N \in \mathcal{E}'_N} U_n(e'_1, \ldots e'_{n-1}, e_n, e'_{n+1}, \ldots, e'_N)$$
$$\prod_{k \neq n, e'_k \neq 0} \lambda_k(\mathbf{e}) \prod_{k \neq n, e'_k = 0} (1 - \lambda_k(\mathbf{e})),$$

where

$$\mathcal{E}'_n = \begin{cases} \{0, e_n\} & \text{if} \quad e_n \neq 0, \\ \{0, e^0\} & \text{if} \quad e_n = 0. \end{cases}$$

The expected net present value of future cash flows $V_n(\mathbf{e}, X_n)$ to incumbent firm n who has drawn salvage value X_n is given by

$$V_n(\mathbf{e}, X_n) = \max\left\{\widehat{X}_n(\mathbf{e}), X_n\right\}.$$

Integrating over all possible salvage values yields the value function $V_n(\mathbf{e}) = \int V_n(\mathbf{e}, X_n) dG_X(X_n)$ for incumbent firm n in state $\mathbf{e}$:

$$V_n(\mathbf{e}) = \begin{cases} \overline{X} & \text{if} \quad \widehat{X}_n(\mathbf{e}) < \overline{X} - a, \\ \frac{1}{4a}\left[\widehat{X}_n(\mathbf{e})^2 - 2\widehat{X}_n(\mathbf{e})(\overline{X} - a) + (\overline{X} + a)^2\right] & \text{if} \quad \widehat{X}_n(\mathbf{e}) \in [\overline{X} - a, \overline{X} + a], \\ \widehat{X}_n(\mathbf{e}) & \text{if} \quad \widehat{X}_n(\mathbf{e}) > \overline{X} + a, \end{cases} \tag{A11}$$

where, recall, $\overline{X}$ is the expected salvage value.

Since salvage values are private, from the point of view of the other firms, the probability that incumbent firm n remains in the industry is

$$\lambda_n(\mathbf{e}) = G_X(\widehat{X}_n(\mathbf{e})) = \begin{cases} 0 & \text{if} \quad \widehat{X}_n(\mathbf{e}) < \overline{X} - a, \\ \frac{1}{2} + \frac{\widehat{X}_n(\mathbf{e}) - \overline{X}}{2a} & \text{if} \quad \widehat{X}_n(\mathbf{e}) \in [\overline{X} - a, \overline{X} + a], \\ 1 & \text{if} \quad \widehat{X}_n(\mathbf{e}) > \overline{X} + a. \end{cases} \tag{A12}$$

Entry decisions. Consider the entry decision $\tau_n(\mathbf{e}, S_n)$ of potential entrant n who has drawn set-up cost S_n. It enters the industry in state $\mathbf{e}$ if its realized set-up cost is less than or equal to the expected value of continuing forward to the price-setting stage:

$$\tau_n(\mathbf{e}, S_n) = \begin{cases} 1 & \text{if} \quad S_n \leq \widehat{S}_n(\mathbf{e}'), \\ 0 & \text{if} \quad S_n \geq \widehat{S}_n(\mathbf{e}'), \end{cases}$$

where

$$\widehat{S}_n(\mathbf{e}) = E[U_n(\mathbf{e}')|\mathbf{e}, e'_n = e^0, \lambda_{-n}(\mathbf{e})]$$

is the expected value of continuing forward to the price-setting stage as an active firm with the initial stock of know-how, i.e., $e'_n = e^0$, taking into account the operating probabilities $\lambda_{-n}(\mathbf{e})$ of the other firms. $\widehat{S}_n(\mathbf{e})$ is computed analogously to $\widehat{X}_n(\mathbf{e})$.

The expected net present value of future cash flows $V_n(\mathbf{e}, S_n)$ to potential entrant n who has drawn set-up cost S_n is given by

$$V_n(\mathbf{e}, S_n) = \max\left\{\widehat{S}_n(\mathbf{e}) - S_n, 0\right\}.$$

Integrating over all possible set-up costs yields the value function $V_n(\mathbf{e}) = \int V_n(\mathbf{e}, S_n) dG_S(S_n)$ for potential entrant n in state $\mathbf{e}$:

$$V_n(\mathbf{e}) = \begin{cases} 0 & \text{if} \quad \widehat{S}_n(\mathbf{e}) < \overline{S} - b, \\ \frac{1}{4b}\left[\widehat{S}_n(\mathbf{e})^2 - 2\widehat{S}_n(\mathbf{e})(\overline{S} - b) + (\overline{S} - b)^2\right] & \text{if} \quad \widehat{S}_n(\mathbf{e}) \in [\overline{S} - b, \overline{S} + b], \\ \widehat{S}_n(\mathbf{e}) - \overline{S} & \text{if} \quad \widehat{S}_n(\mathbf{e}) > \overline{S} + b. \end{cases} \tag{A13}$$

Finally, from the point of view of the other firms, the probability that potential entrant n enters the industry is

$$\lambda_n(\mathbf{e}) = G_S(\widehat{S}_n(\mathbf{e})) = \begin{cases} 0 & \text{if} \quad \widehat{S}_n(\mathbf{e}) < \overline{S} - b, \\ \frac{1}{2} + \frac{\widehat{S}_n(\mathbf{e}) - \overline{S}}{2b} & \text{if} \quad \widehat{S}_n(\mathbf{e}) \in [\overline{S} - b, \overline{S} + b], \\ 1 & \text{if} \quad \widehat{S}_n(\mathbf{e}) > \overline{S} + b. \end{cases} \tag{A14}$$

Equilibrium. We restrict ourselves to symmetric and anonymous Markov perfect equilibria. Symmetry allows us to focus on the problem of firm 1 and anonymity (also called exchangeability) says that firm 1 does not care about the identity of its rivals, only about the distribution of their states (see, e.g., Doraszelski & Satterthwaite (2007) for a formal definition). It therefore suffices to determine the value and policy functions of firm 1, and we define $V^*(\mathbf{e}) = V_1(\mathbf{e})$, $p^*(\mathbf{e}) = p_1(\mathbf{e})$, and $\lambda^*(\mathbf{e}) = \lambda_1(\mathbf{e})$ for each state $\mathbf{e}$.. The corresponding value and policy functions for firm n in state $\mathbf{e}$ are recovered as $V_n(\mathbf{e}) = V^*(\mathbf{e}^{[n]})$, $p_n(\mathbf{e}) = p^*(\mathbf{e}^{[n]})$, and $\lambda_n(\mathbf{e}) = \lambda^*(\mathbf{e}^{[n]})$, where $\mathbf{e}^{[n]}$ is constructed from $\mathbf{e}$ by interchanging the stocks of know-how of firms 1 and n.

Parameterization. Since the analysis of entry and exit requires a well-posed monopoly problem, we include an outside good. Specifically, we set $v = 10$ and $v_0 - c_0 = 0$. We further set $\overline{X} = 1.5$ and $a = 1.5$ to ensure that salvage values are uniformly distributed with support $[0, 3]$ and $\overline{S} = 4.5$ and $b = 1.5$ to ensure that set-up costs are uniformly distributed with support $[3, 6]$.

Results. Figure A20 displays the limiting and maximum expected Herfindahl indices for the general model with entry and exit. Because our parameterization includes an outside good with $v_0 - c_0 = 0$, this figure may be compared to Figure A7.

References

Benkard, L. (2000), 'Learning and forgetting: The dynamics of aircraft production', *American Economic Review* **90**(4), 1034–1054.

Bertsekas, D. & Tsitsiklis, J. (1997), *Parallel and distributed computation: Numerical methods*, Athena Scientific, Belmont.

Doraszelski, U. & Satterthwaite, M. (2007), Computable Markov-perfect industry dynamics: Existence, purification, and multiplicity, Working paper, Harvard University, Cambridge.

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.

Stokey, N. & Lucas, R. with Prescott, E. (1989), *Recursive methods in economic dynamics*, Harvard University Press, Cambridge.

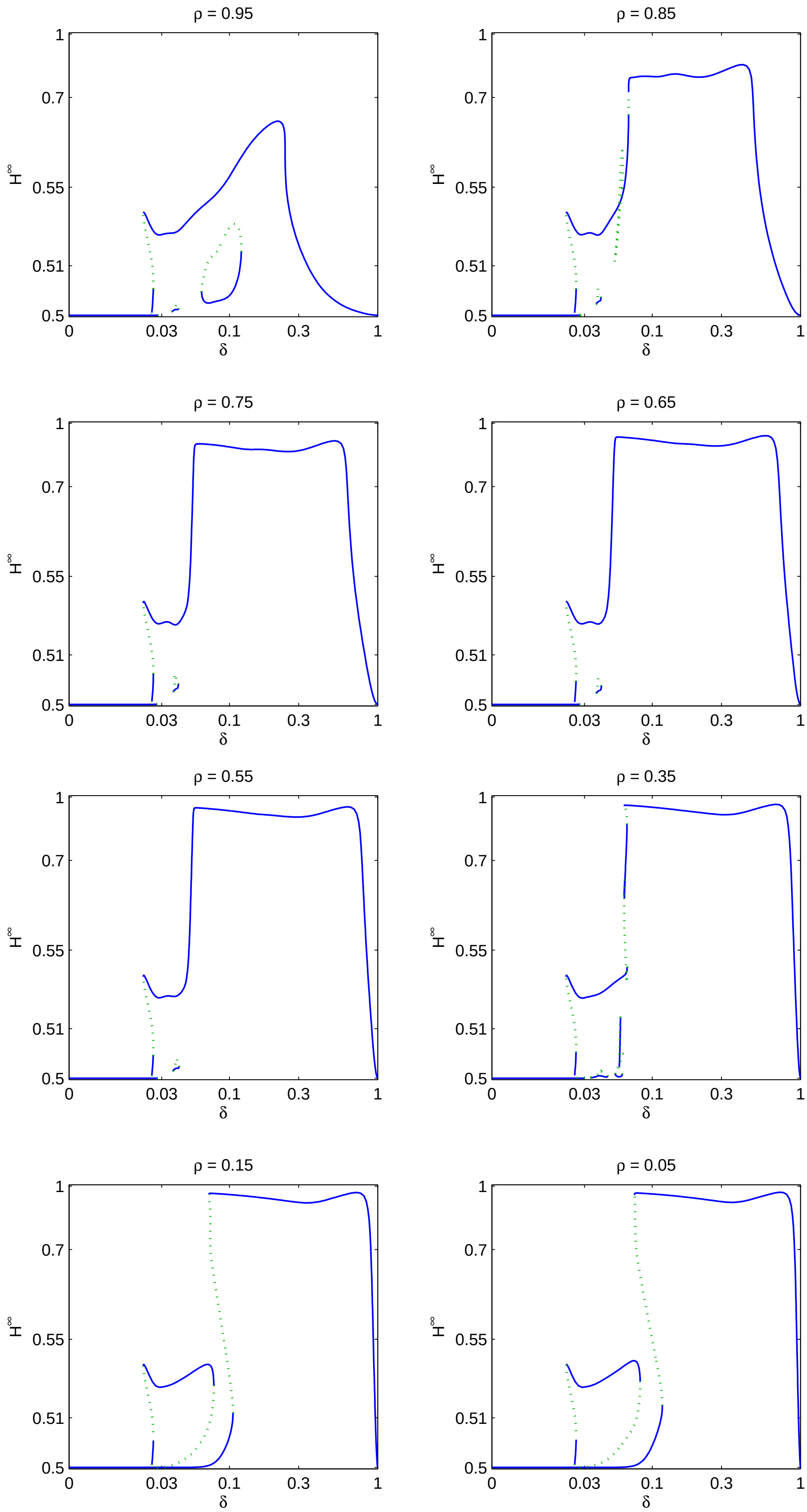

Figure A1: Limiting expected Herfindahl index H^{∞}. Equilibria with $\rho\left(\left.\frac{\partial \mathbf{G}(\mathbf{x}(s))}{\partial \mathbf{x}}\right|_{\delta=\delta(s)}\right) <$ 1 (solid line) and equilibria with $\rho\left(\left.\frac{\partial \mathbf{G}(\mathbf{x}(s))}{\partial \mathbf{x}}\right|_{\delta=\delta(s)}\right) \geq 1$ (dotted line).

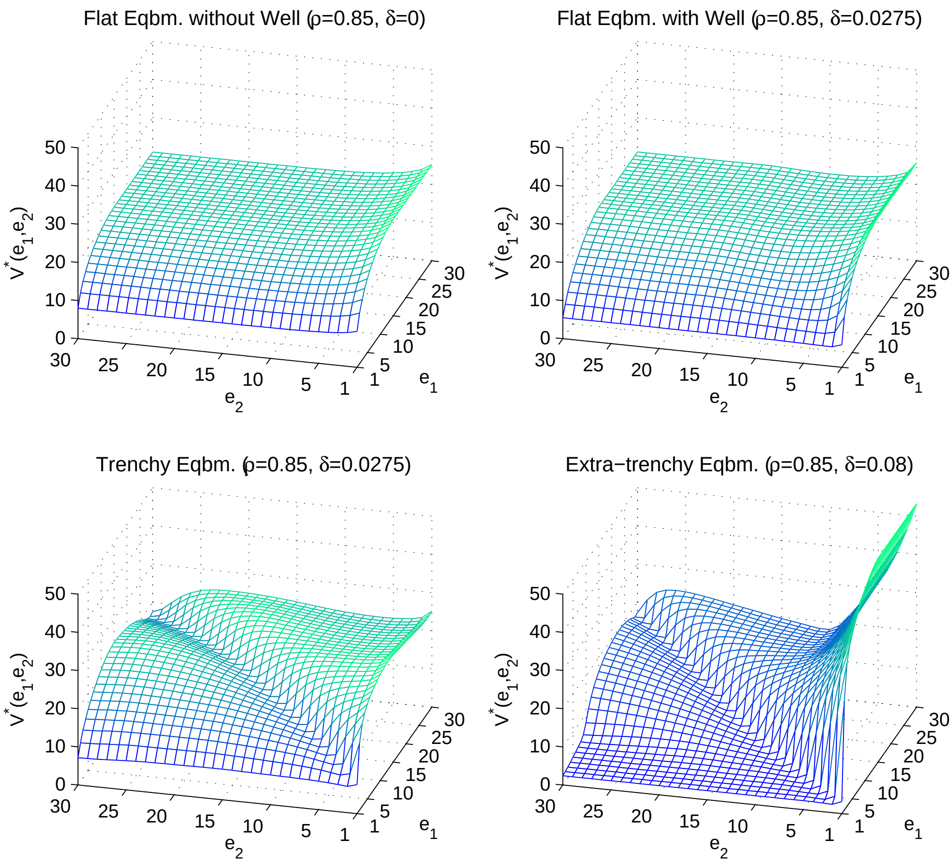

Figure A2: Value function $V^*(e_1, e_2)$.

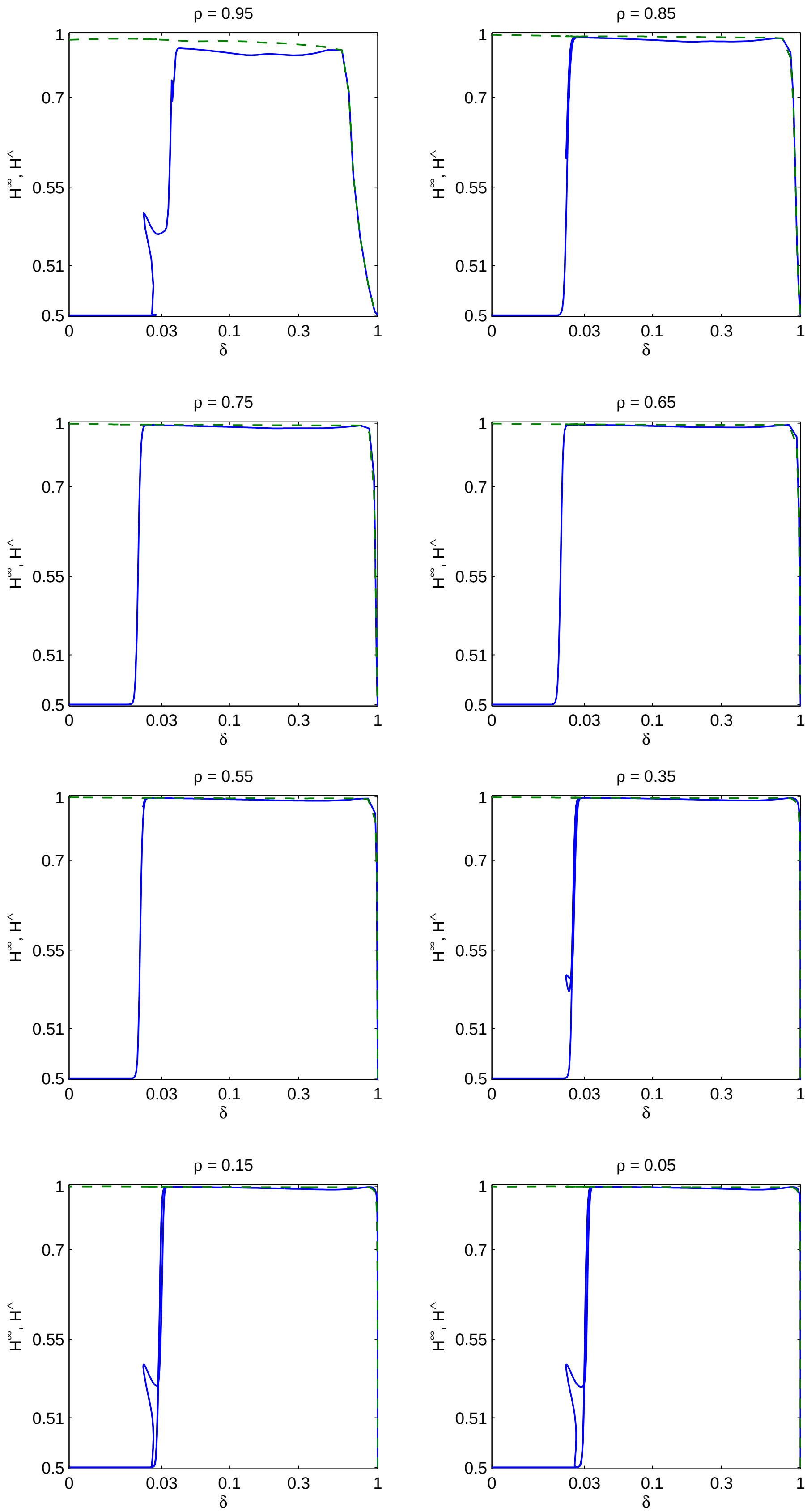

Figure A3: Product differentiation with $\sigma = 0.2$. Limiting expected Herfindahl index H^{∞} (solid line) and maximum expected Herfindahl index $H^{\wedge}$ (dashed line).

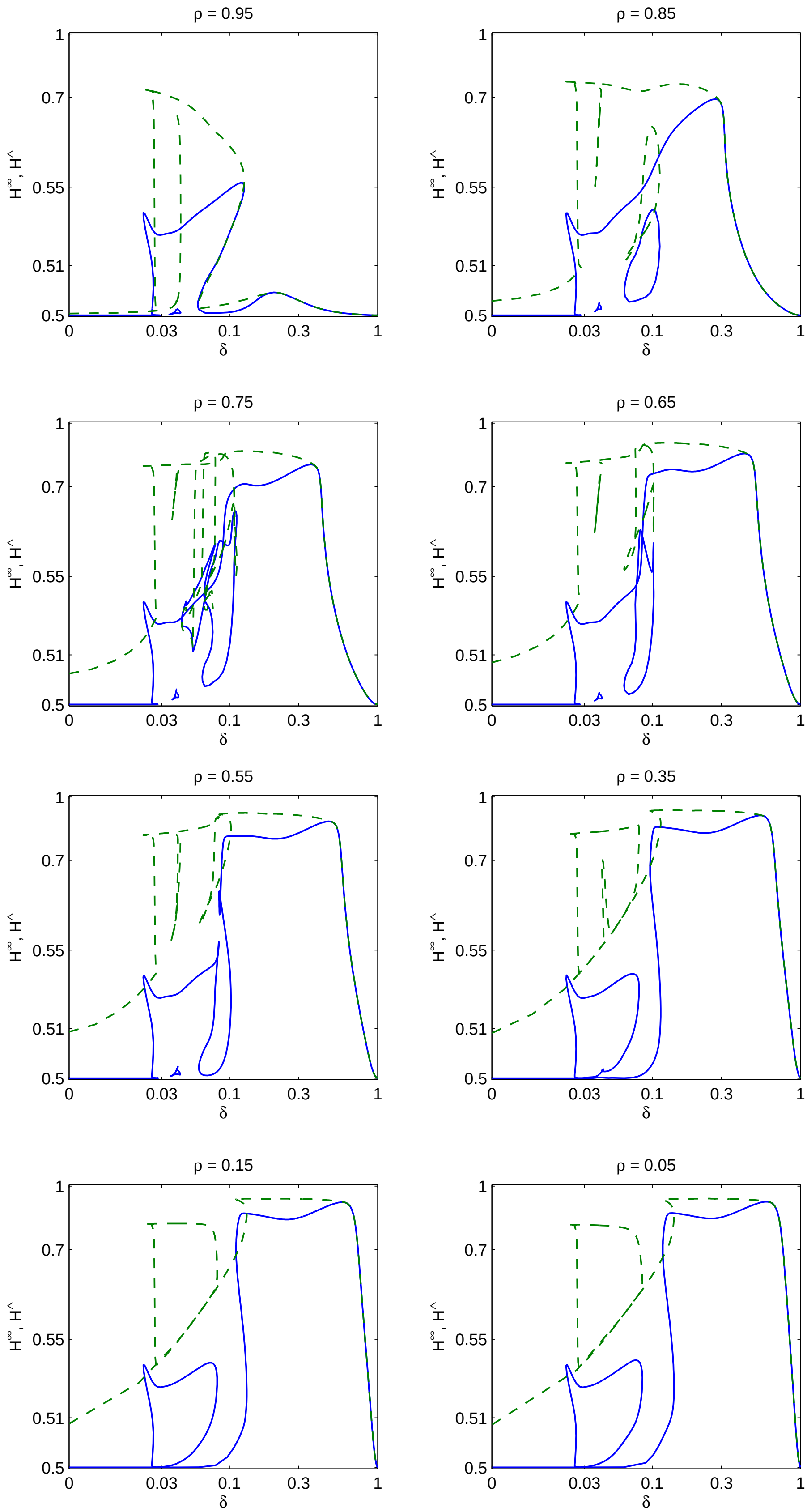

Figure A4: Product differentiation with $\sigma = 2$. Limiting expected Herfindahl index H^{∞} (solid line) and maximum expected Herfindahl index $H^{\wedge}$ (dashed line).

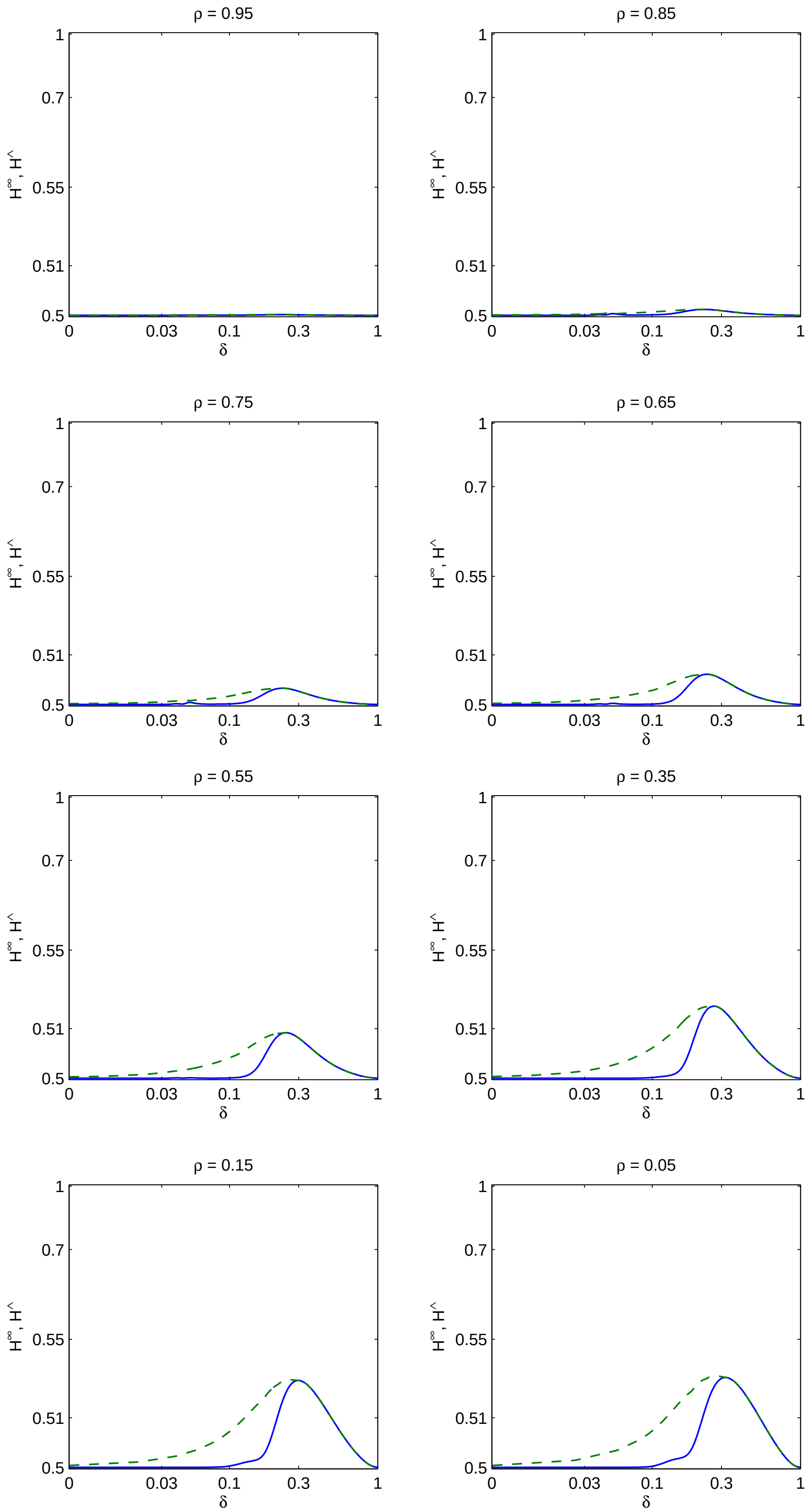

Figure A5: Product differentiation with $\sigma = 10$. Limiting expected Herfindahl index H^∞ (solid line) and maximum expected Herfindahl index $H^\wedge$ (dashed line).

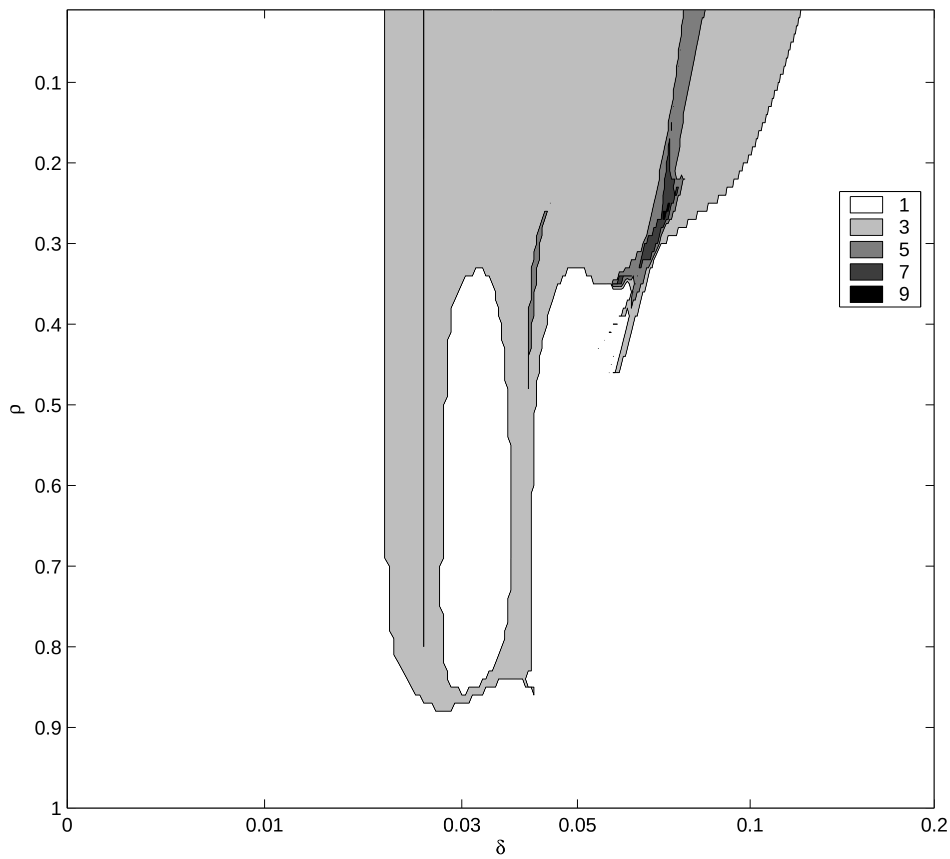

Figure A6: Outside good with $v_0 - c_0 = 0$. Number of equilibria.

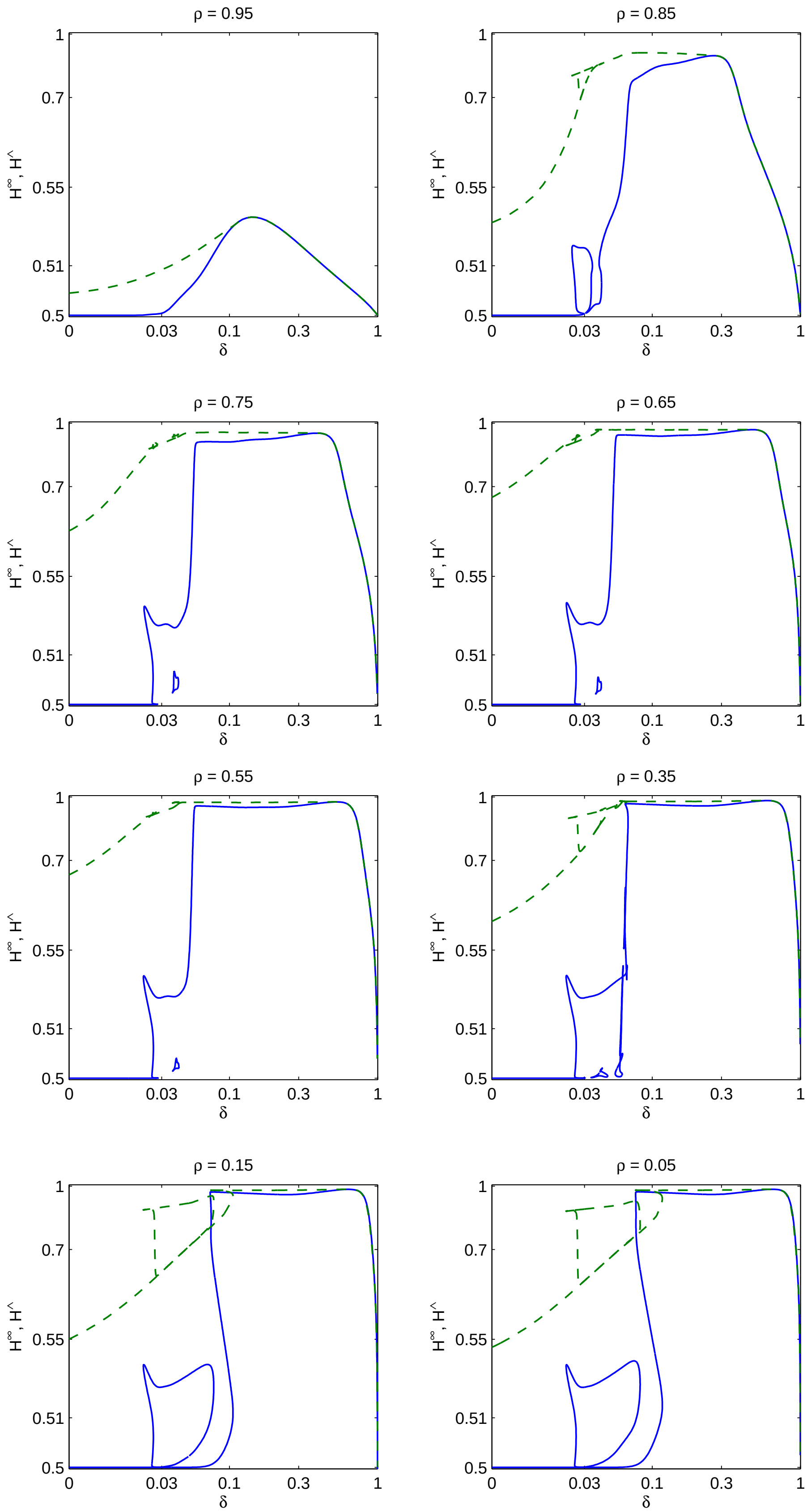

Figure A7: Outside good with $v_0 - c_0 = 0$. Limiting expected Herfindahl index H^∞ (solid line) and maximum expected Herfindahl index $H^\wedge$ (dashed line).

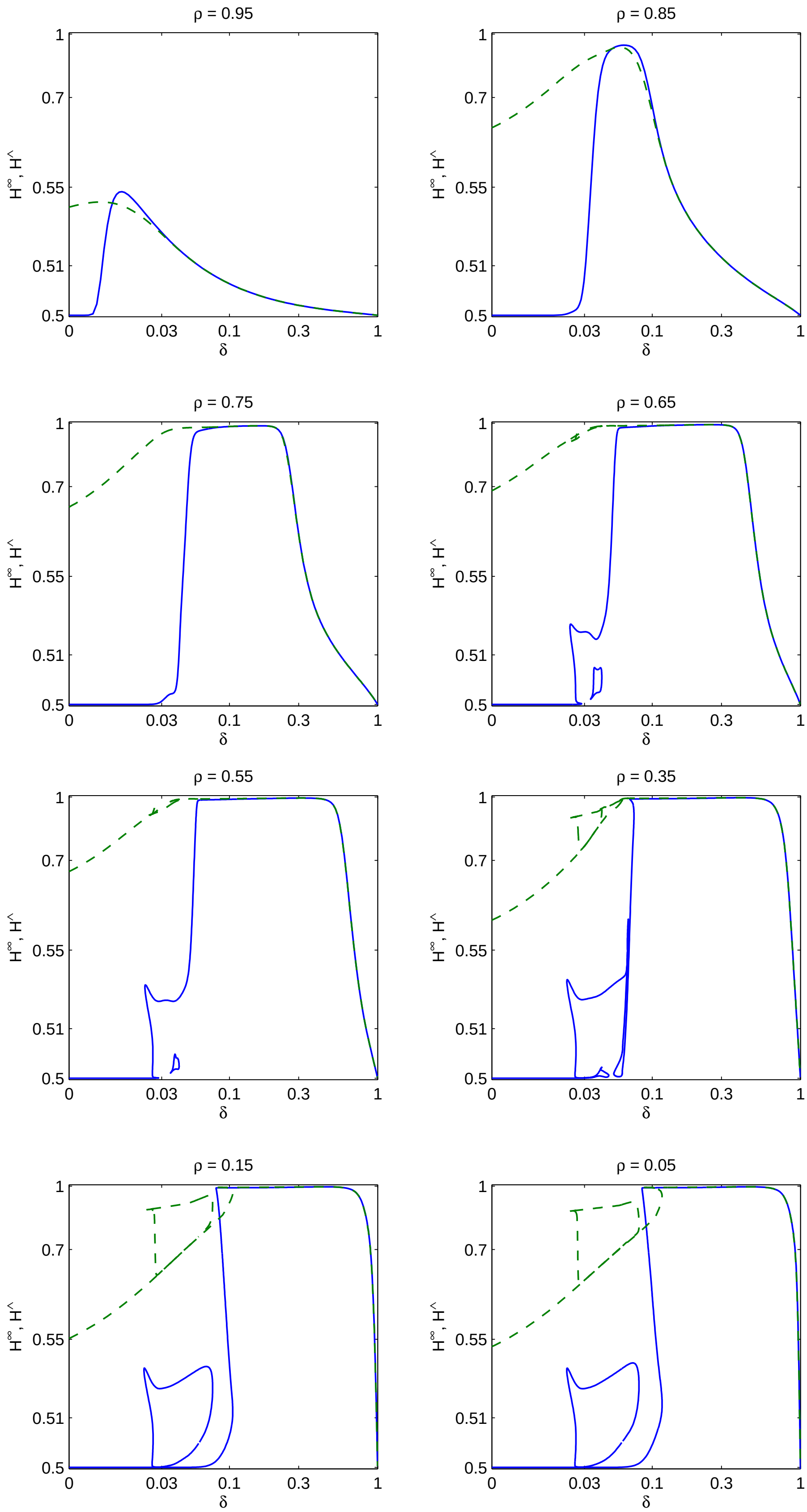

Figure A8: Outside good with $v_0 - c_0 = 3$. Limiting expected Herfindahl index H^∞ (solid line) and maximum expected Herfindahl index $H^\wedge$ (dashed line).

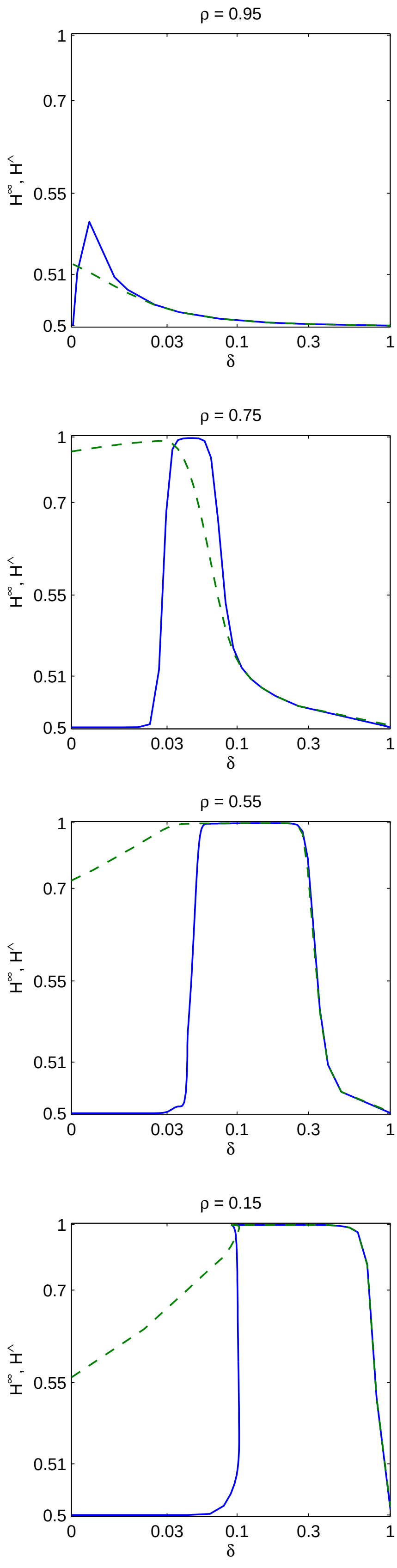

Figure A9: Outside good with $v_0 - c_0 = 5$. Limiting expected Herfindahl index H^∞ (solid line) and maximum expected Herfindahl index $H^\wedge$ (dashed line).

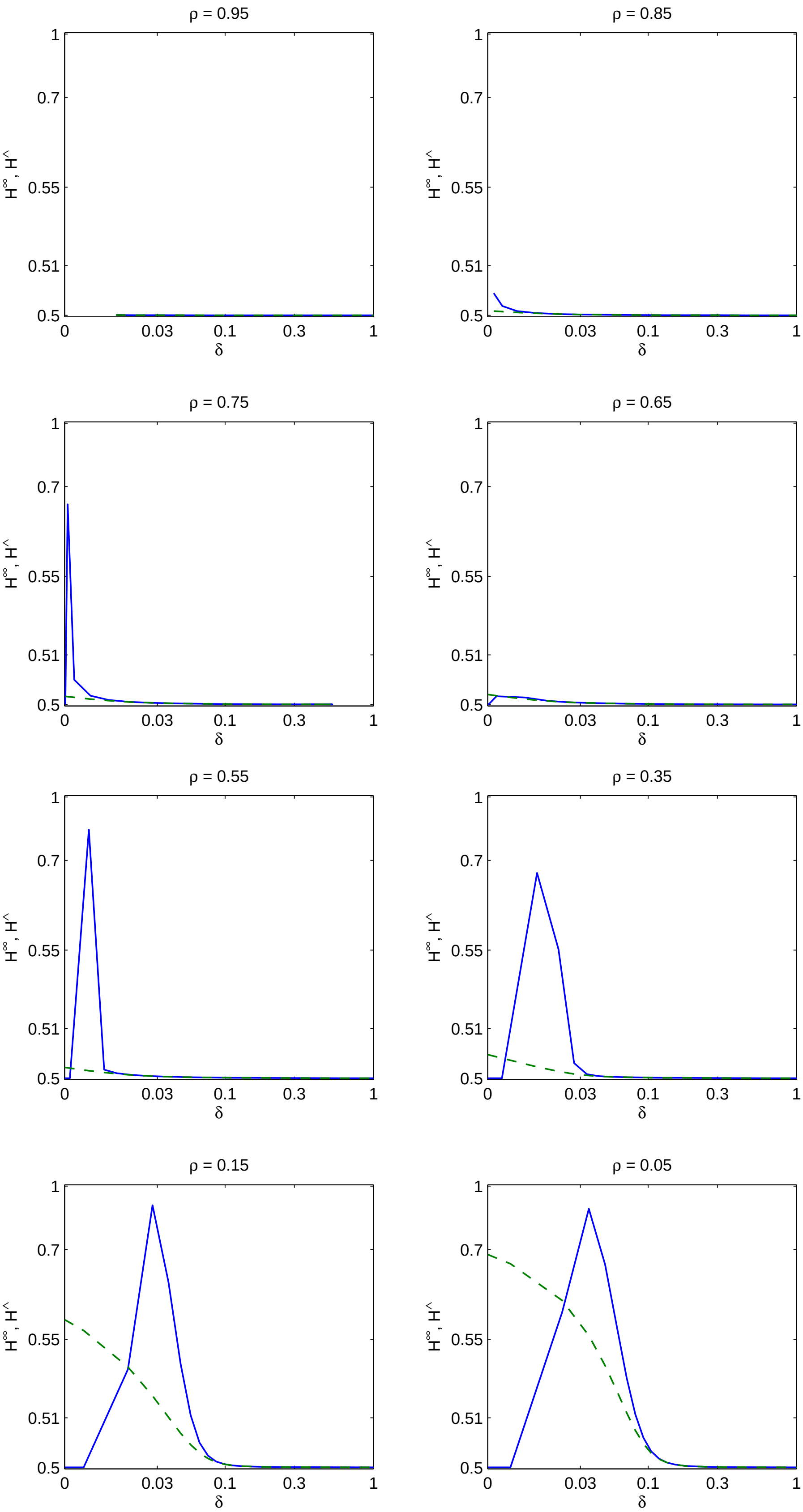

Figure A10: Outside good with $v_0 - c_0 = 10$. Limiting expected Herfindahl index H^∞ (solid line) and maximum expected Herfindahl index $H^\wedge$ (dashed line).

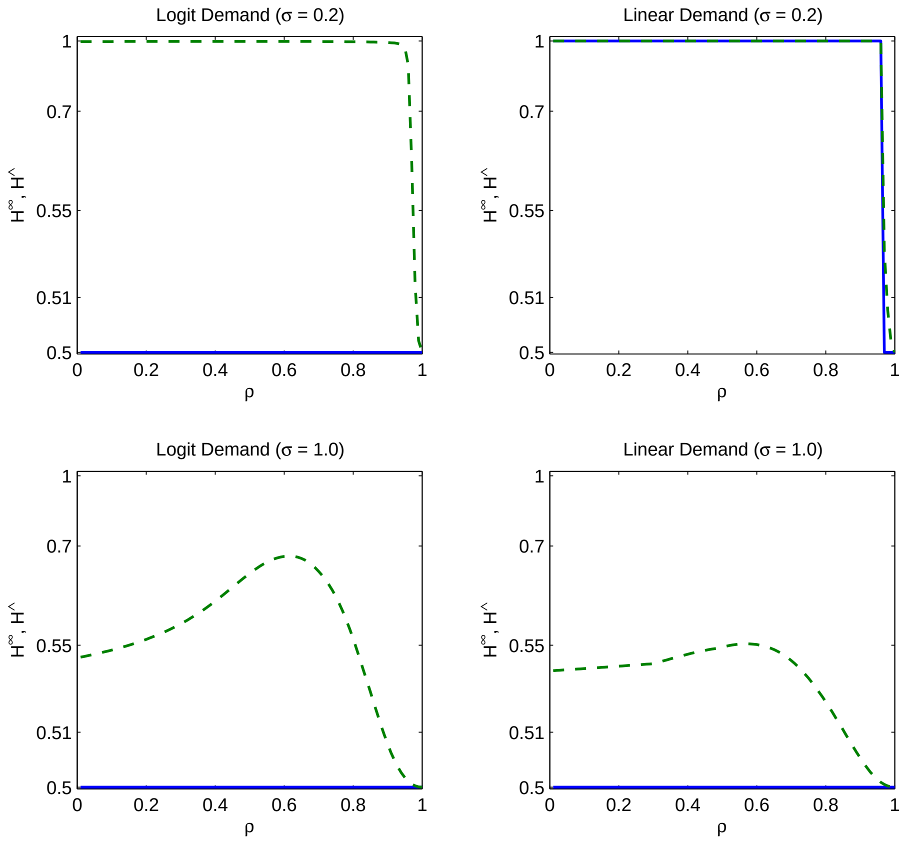

Figure A11: Choke price. Limiting expected Herfindahl index H^{∞} (solid line) and maximum expected Herfindahl index $H^{\wedge}$ (dashed line).

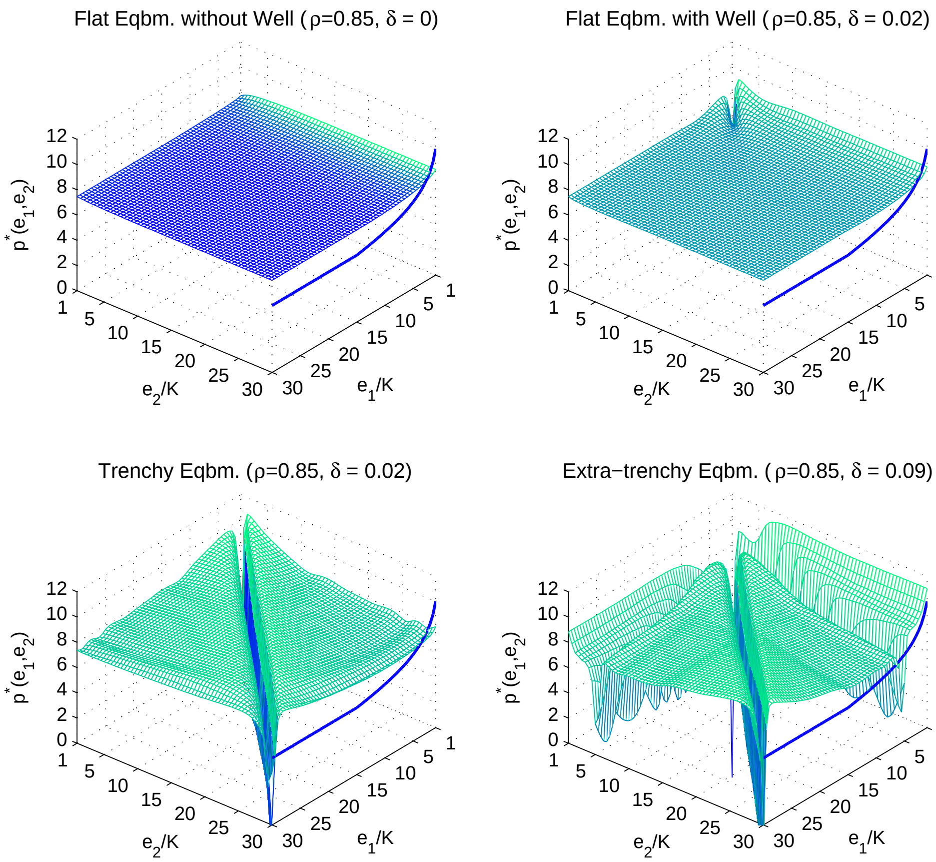

Figure A12: Frequency of sales with $K = 2$. Policy function $p^*(e_1, e_2)$. Marginal cost $c(e_1)$ (solid line in $e_2 = 30$-plane).

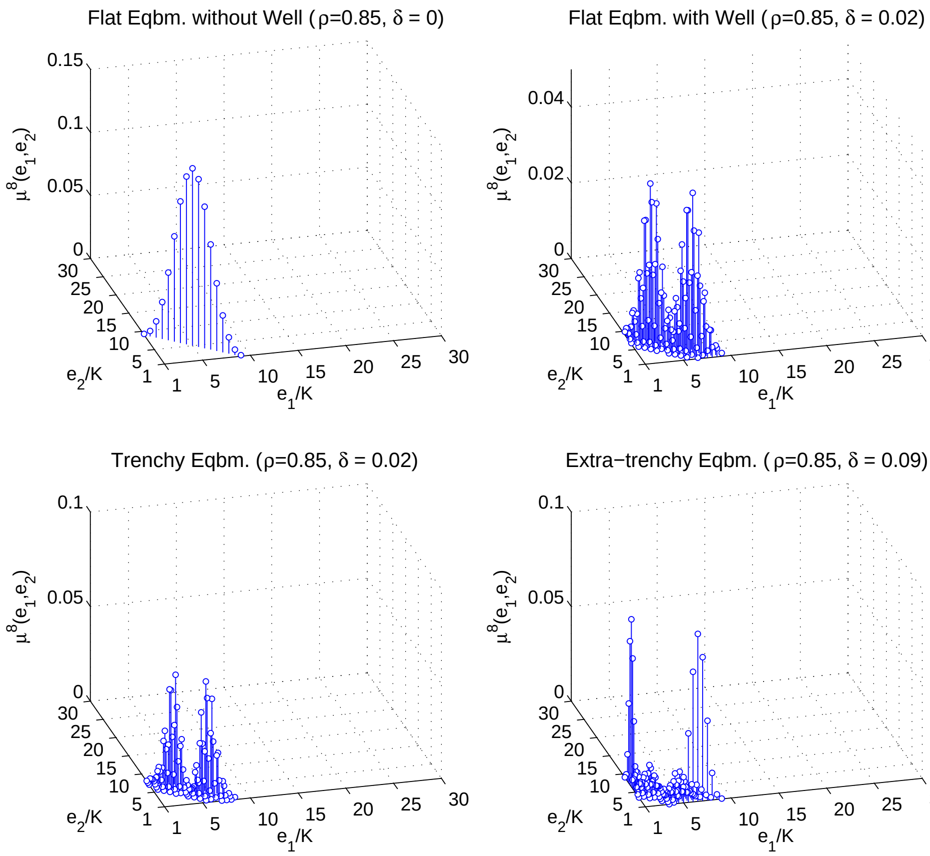

Figure A13: Frequency of sales with $K = 2$. Transient distribution over states in period 8 (subperiod 16) given initial state $(1, 1)$.

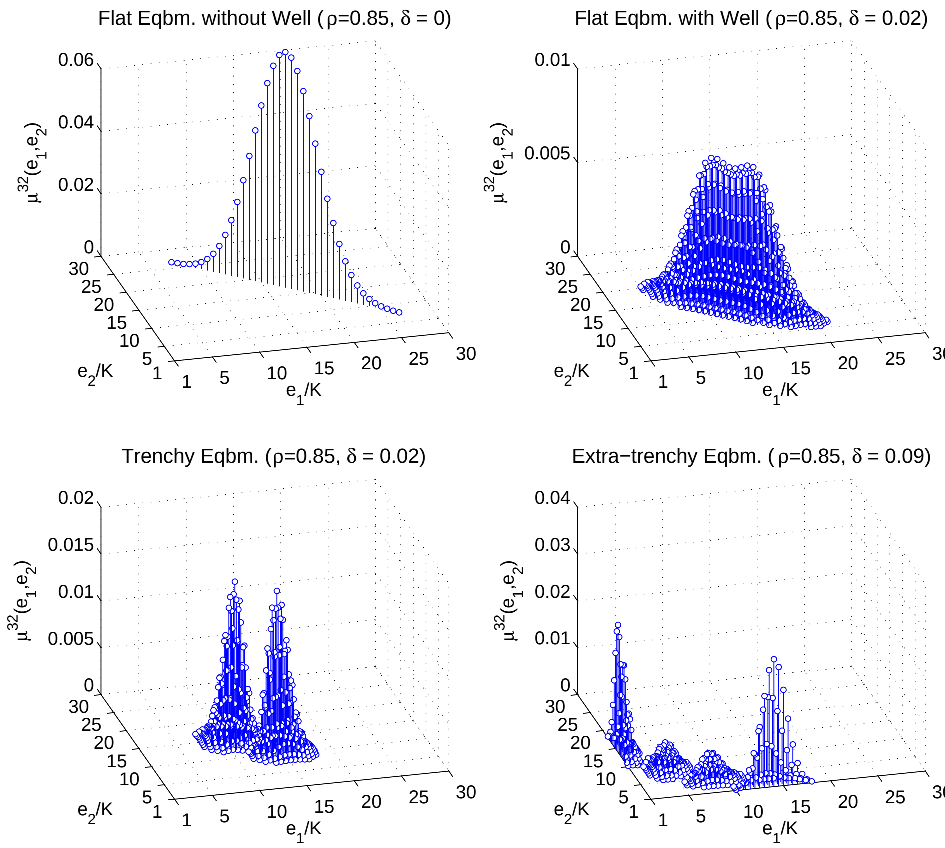

Figure A14: Frequency of sales with $K = 2$. Transient distribution over states in period 32 (subperiod 64) given initial state $(1, 1)$.

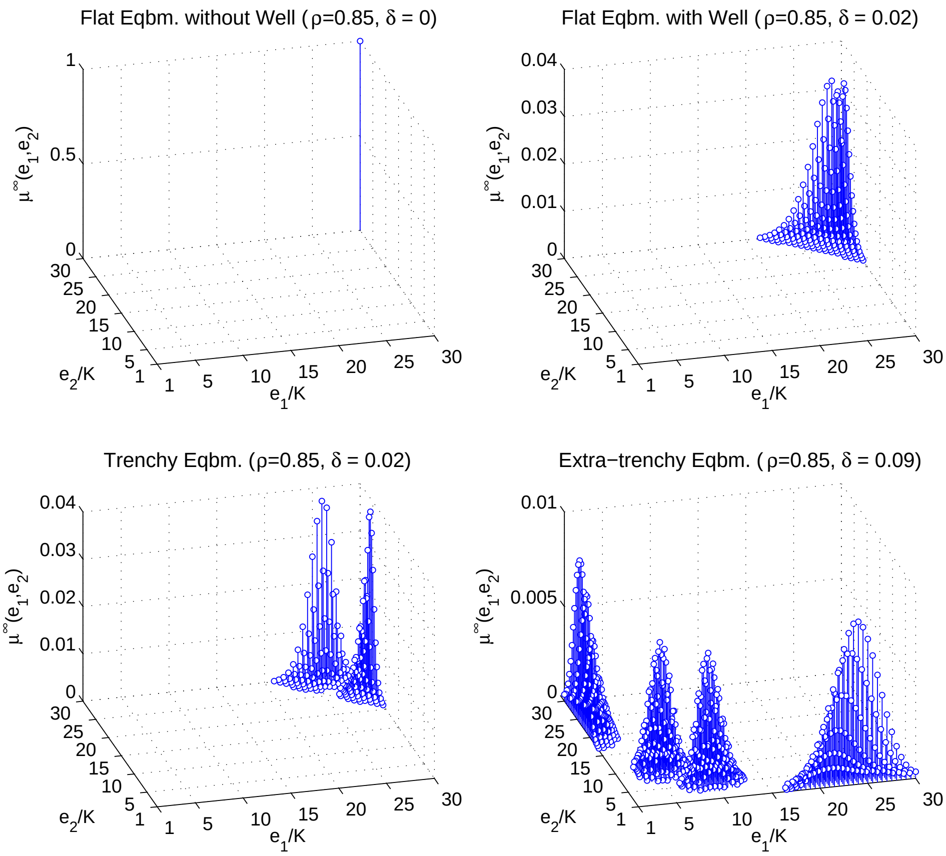

Figure A15: Frequency of sales with $K = 2$. Limiting distribution over states.

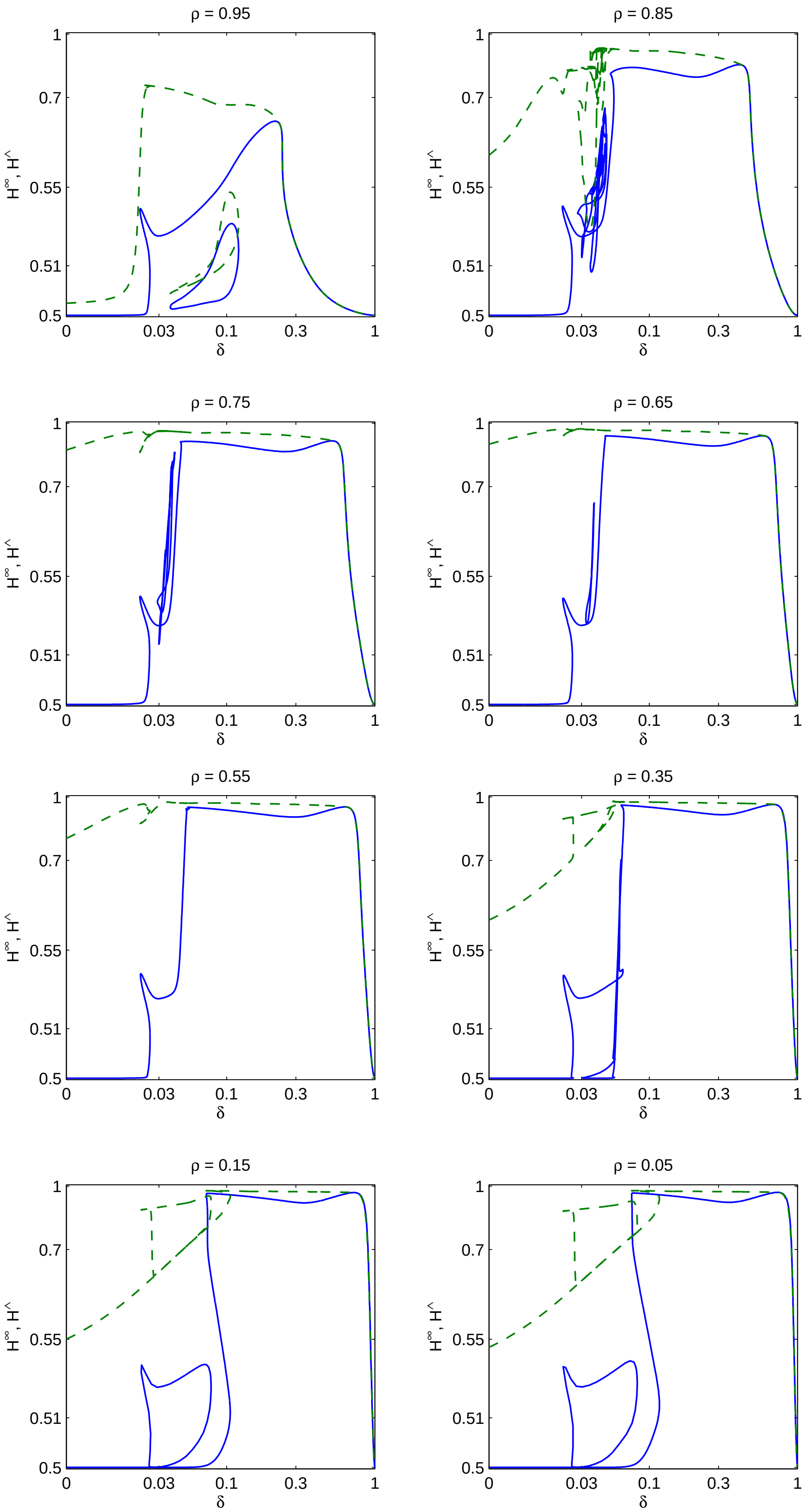

Figure A16: Bottomless learning. Limiting expected Herfindahl index H^{∞} (solid line) and maximum expected Herfindahl index $H^{\wedge}$ (dashed line).

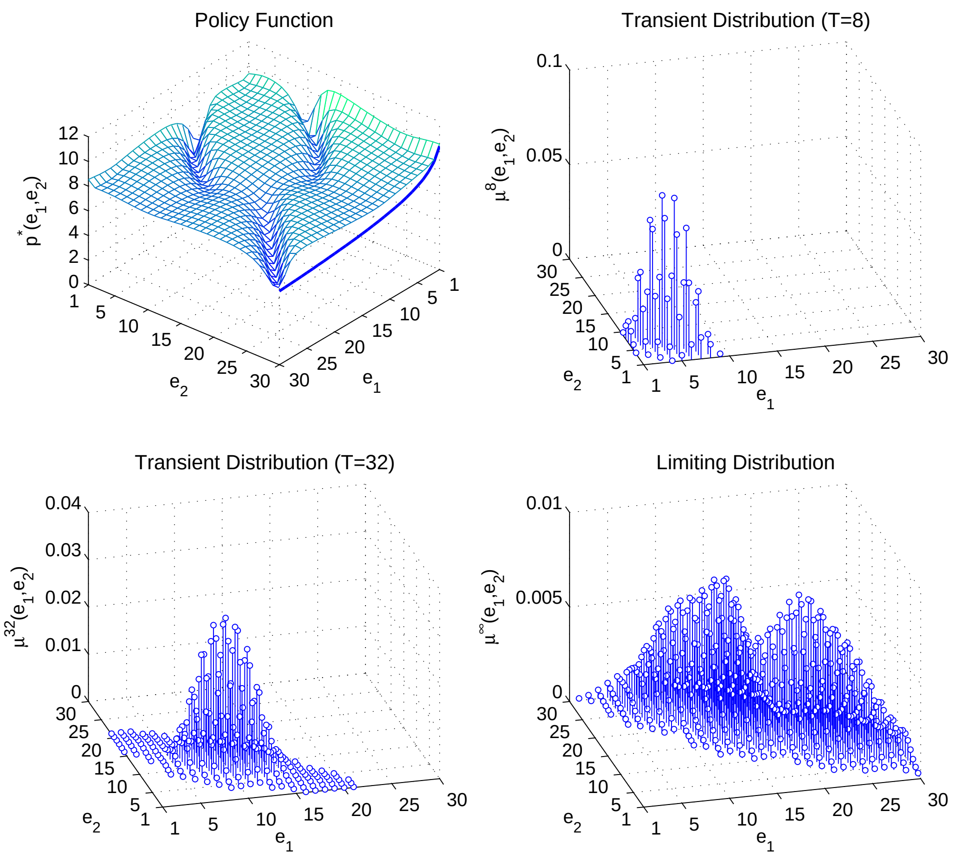

Figure A17: Bottomless learning. Policy function $p^*(e_1, e_2)$. Marginal cost $c(e_1)$ (solid line in $e_2 = 30$-plane) (upper left panel). Transient distribution over states in period 8 and 32 given initial state $(1, 1)$ (upper right and lower left panels). Limiting distribution over states (lower right panel). Plateau equilibrium ($\rho = 0.9$, $\delta = 0.04$).

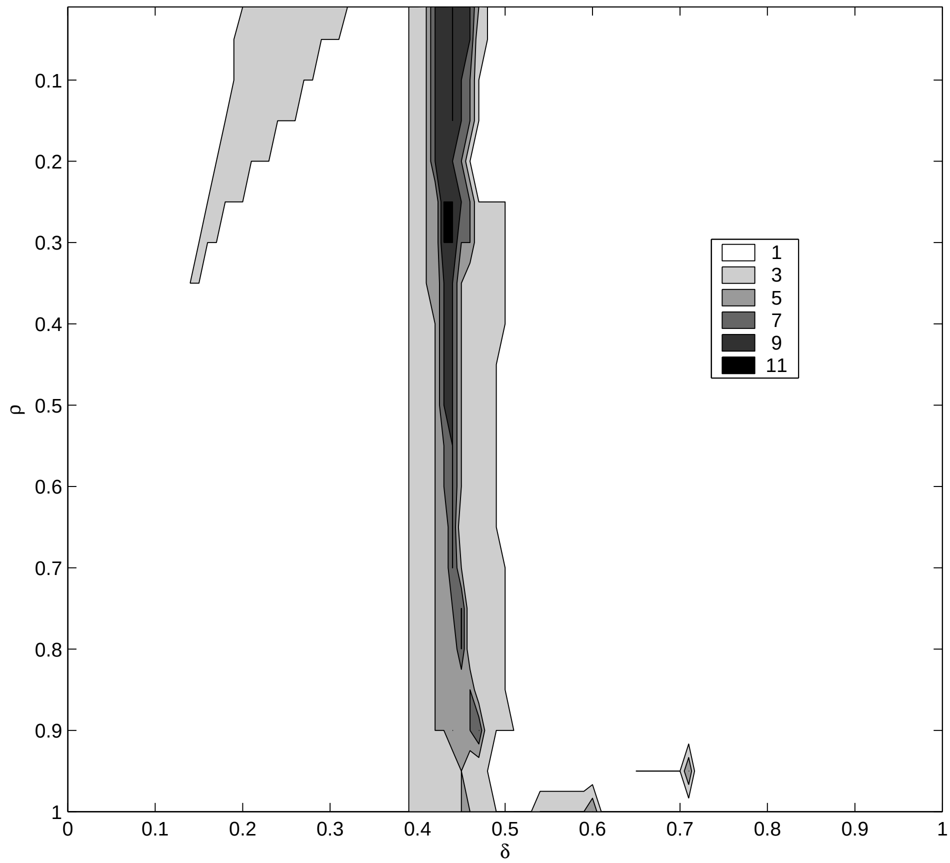

Figure A18: Constant forgetting. Number of equilibria.

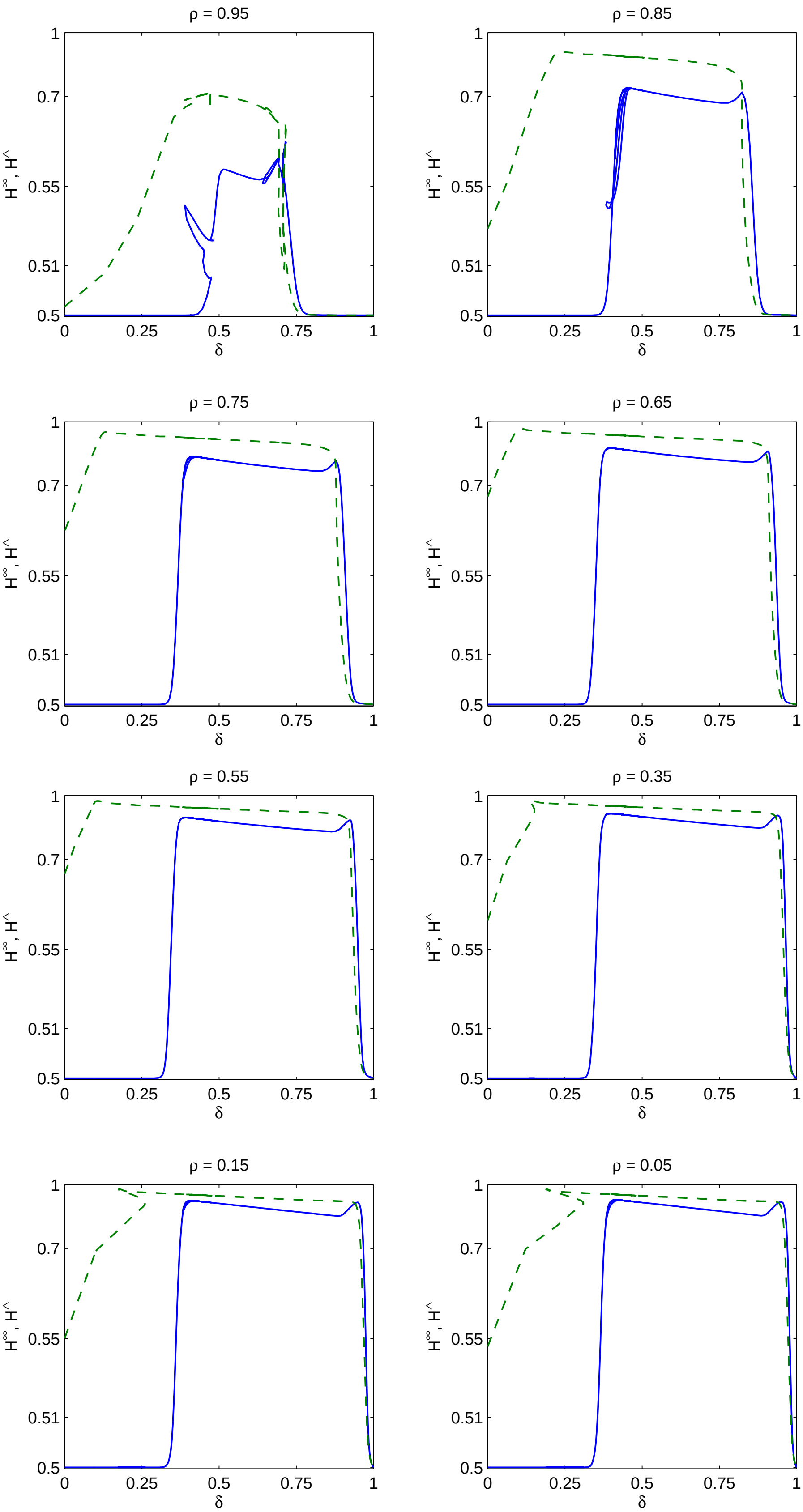

Figure A19: Constant forgetting. Limiting expected Herfindahl index H^∞ (solid line) and maximum expected Herfindahl index $H^\wedge$ (dashed line).

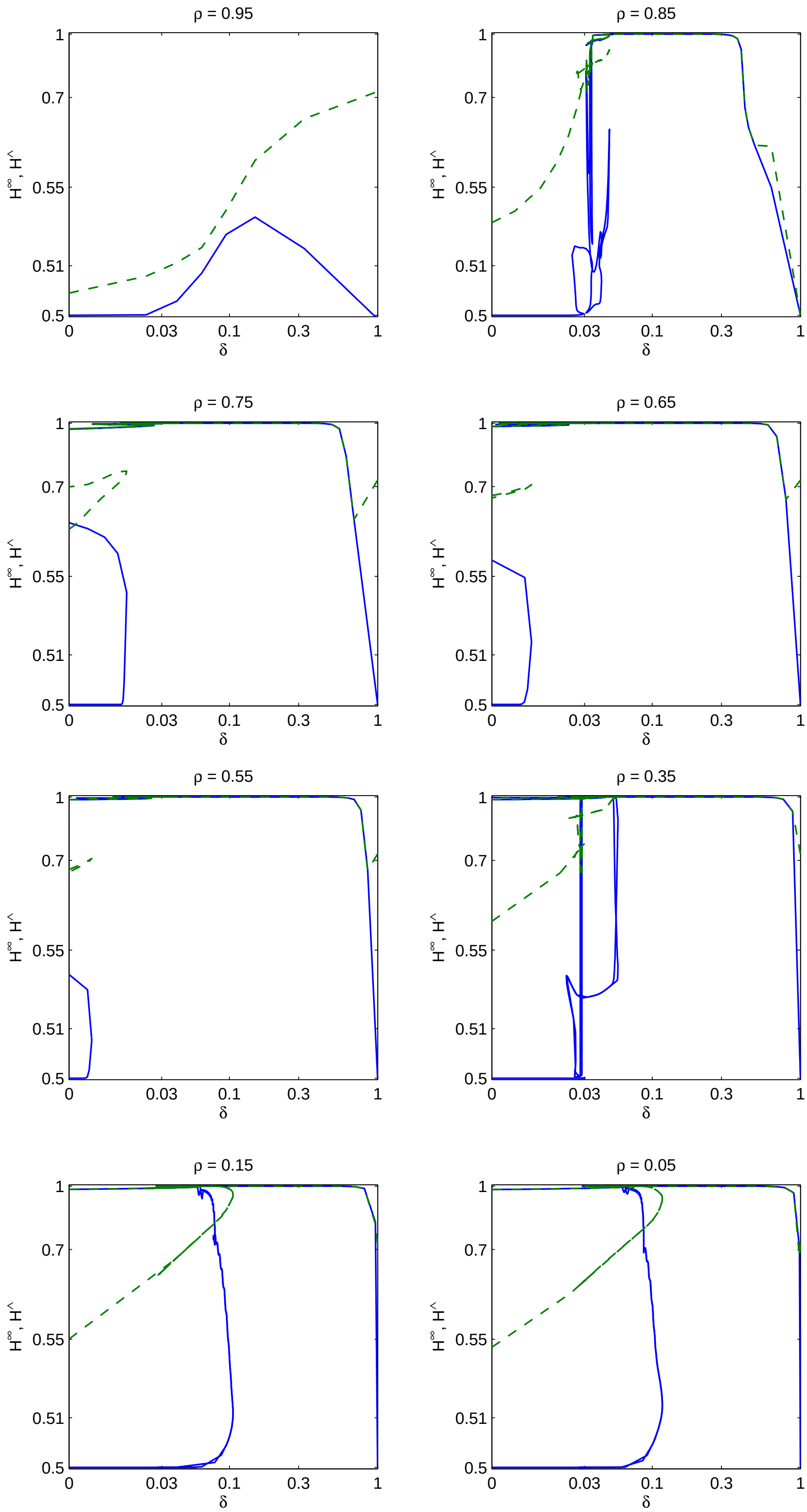

Figure A20: Entry and exit. Limiting expected Herfindahl index H^∞ (solid line) and maximum expected Herfindahl index $H^\wedge$ (dashed line).