Learning-by-Doing, Organizational Forgetting, and Industry Dynamics*

David Besanko[†] Ulrich Doraszelski[‡] Yaroslav Kryukov[§]
Mark Satterthwaite[¶]

May 9, 2005
— PRELIMINARY AND INCOMPLETE —

Abstract

Learning-by-doing and organizational forgetting have been shown to be important in a variety of industrial settings. This paper provides a general model of dynamic competition that accounts for these economic fundamentals and shows how they shape industry structure and dynamics. Previously obtained results regarding the dominance properties of firms' pricing behavior no longer hold in this more general setting. We show that organizational forgetting does not simply negate learning-by-doing. Rather, learning-by-doing and organizational forgetting are distinct economic forces. In particular, a model with both learning-by-doing and organizational forgetting can give rise to aggressive pricing behavior, market dominance, and multiple equilibria, whereas a model with learning-by-doing alone cannot.

1 Introduction

{INTRODUCTIO

Empirical studies provide ample evidence that the marginal cost of production decreases with cumulative experience in a variety of industrial settings (see, e.g., Alchian 1963, Zimmerman 1982, Lieberman 1984, Irwin & Klenow 1994, Thompson 2001, Thornton &

*We are indebted to Luis Cabral, Stefano Demichelis, Michaela Draganska, Ken Judd, Ariel Pakes, Michael Ryall, Karl Schmedders, Chris Shannon, Scott Stern, Michael Whinston, and Huseyin Yildirim for comments and suggestions. We have also bene¯tted from the comments of participants at the 2004 Duke-Northwestern-Texas IO Theory Conference and the 2004 Harvard Business School Strategy Conference.

†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.

Thompson 2001). This fall in marginal cost is known as learning-by-doing. More recent empirical studies also suggest that organizations can forget the know-how gained through learning-by-doing due to labor turnover, periods of inactivity, and failure to institutionalize tacit knowledge (Argote, Beckman & Epple 1990, Thompson 2003, Darr, Argote & Epple 1995, Benkard 2000). Organizational forgetting has been largely ignored by the existing literature. This is especially troubling because Benkard (2004) shows that organizational forgetting is needed to explain the dynamics in the market for wide-bodied airframes in the 1970s and 1980s. Moreover, the existing literature has mainly focused on firms' pricing behavior and has fallen short of directly examining the impact of learning-by-doing on industry dynamics.

This paper provides a general model of dynamic competition with learning-by-doing and organizational forgetting. Using the Markov-perfect equilibrium framework of Ericson & Pakes (1995) we show how these economic fundamentals shape the structure and dynamics of an industry. Closest in spirit to our model is the model with learning-by-doing alone that Cabral & Riordan (1994) presented in their seminal paper. Their analysis is centered around "two concepts of self-reinforcing market dominance" (p. 1115), increasing dominance (ID) and increasing increasing dominance (IID). If firms' pricing behavior satisfies ID, then the leader charges a lower price than the follower, and the gap between the leader's price and the follower's price widens with the length of the lead if it satisfies IID. We go beyond analyzing these dominance properties of firms' pricing behavior and directly examine the industry dynamics implied by that behavior using tools from stochastic process theory. Since the Cabral & Riordan (1994) model is a special case of ours, we are able to show that ID and IID may not be sufficient for economically meaningful market dominance. The reason is that, while the leader charges a *slightly* lower price than the follower and this gap widens *a bit* over time, with even a modest degree of horizontal product differentiation, the firms still split sales more or less equally and thus move down the learning curve in tandem. Hence, even if ID and IID hold, they may not tell us as much as we like about the structure and dynamics of an industry.

We generalize Cabral & Riordan's (1994) and other existing models of learning-by-doing through the addition of organizational forgetting.[1] This seemingly small change has surprisingly large effects. Dynamic competition with learning-by-doing and organizational forgetting is akin to racing down an upward moving escalator. As long as a firm makes sales sufficiently frequently so that the gain in know-how from learning-by-doing outstrips the loss in know-how from organizational forgetting, it moves down its learning curve and its marginal cost decreases. However, if sales slow down or come to a halt, then the firm slides

[1] Prior to the in¯nite-horizon price-setting model of Cabral & Riordan (1994), the literature has studied learning-by-doing using ¯nite-horizon quantity-setting models (e.g., Spence 1981, Fudenberg & Tirole 1983, Ghemawat & Spence 1985, Ross 1986, Cabral & Riordan 1997).

back up its learning curve and its marginal cost increases. This cannot happen in a model with learning-by-doing alone. Due to this qualitative difference, many previously obtained results no longer hold in a model with both learning-by-doing and organizational forgetting. In particular, by going from Cabral & Riordan's (1994) setting without organizational forgetting to a setting with organizational forgetting, we are able to show that the dominance properties that are at the center of their analysis break down in the presence of even a small degree of organizational forgetting.

Besides demonstrating that the dominance properties of firms' pricing behavior are neither very informative nor very robust, we carefully examine the role of organizational forgetting. It is often said that learning-by-doing promotes market dominance because it gives a more experienced firm the ability to profitably underprice its less experienced rival. But if learning-by-doing can be "undone" by organizational forgetting, this raises the question whether organizational forgetting is an antidote to market dominance for two reasons. First, to the extent that the leader has more to forget than the follower, organizational forgetting should work to equalize differences between firms. Second, because organizational forgetting makes improvements in competitive position from learning-by-doing more transitory, it should make firms more reluctant to invest in the acquisition of know-how through price cuts. We reach the opposite conclusion: organizational forgetting tends to make firms more instead of less aggressive. This aggressive pricing behavior, in turn, puts the industry on a path towards market dominance.

Specifically, we first show that organizational forgetting is a source of aggressive pricing behavior. The price that a firm sets reflects its twin goals of building and defending a competitive advantage over its rival. In the presence of organizational forgetting, bidirectional movements through the state space are possible, and this opens up new strategic possibilities for firms that work to enhance the advantage-building and advantage-defending motives relative to a model with learning-by-doing alone. By winning a sale, a firm makes itself less vulnerable to future losses from organizational forgetting, thus enhancing the advantage-building motive. It also makes its rival more vulnerable to future losses from organizational forgetting, thus enhancing the advantage-defending motive. Because these additional benefits are achieved by winning a sale, organizational forgetting creates strong incentives to cut prices.

Next we show that organizational forgetting is a source of—and not an antidote to—market dominance. If organizational forgetting is sufficiently weak, then asymmetries may arise but they cannot persist. If organizational forgetting is sufficiently strong, then asymmetries cannot arise in the first place because organizational forgetting stifles investment in learning-by-doing altogether. By contrast, for intermediate degrees of organizational forgetting, asymmetries arise and persist. Even extreme asymmetries akin to near-monopoly are

possible. This is because organizational forgetting in effect commits the leader to aggressively defending its position against imminent and distant threats. This more than offsets the increased vulnerability to organizational forgetting as the stock of know-how grows and makes the leadership position more secure than it would have been in the absence of organizational forgetting.

In addition, we show that organizational forgetting is a source of multiple equilibria. If the inflow of know-how into the industry due to learning-by-doing is substantially smaller than the outflow of know-how due to organizational forgetting, then it is virtually impossible that both firms reach the bottom of their learning curves. Conversely, if the inflow is substantially greater than the outflow, then it is virtually inevitable that they do. In both cases, the primitives of the model tie down the equilibrium. This is no longer the case if the inflow roughly balances the outflow, and the stage is set for multiple equilibria. If firms believe that they cannot profitably coexist at the bottom of their learning curves and that instead one firm comes to dominate the market, then both firms cut their prices in the hope of acquiring a competitive advantage early on and maintaining it throughout. This aggressive pricing behavior, in turn, leads to market dominance. However, if firms believe that they can profitably coexist, then neither firm cuts its price, thereby ensuring that the anticipated symmetric industry structure actually emerges. Consequently, in addition to the degree of organizational forgetting, the equilibrium by itself is an important determinant of pricing behavior and industry dynamics.

In our model multiple equilibria do not arise because of the specification of the primitives. In fact, we are able to show that multiple equilibria arise from firms' expectations regarding the value of continued play. In this sense multiplicity is rooted in the dynamics of the model. Our finding of multiplicity is important for two reasons. First, to our knowledge, all applications of Ericson & Pakes's (1995) framework have found a single equilibrium. It indeed is often held that "nonuniqueness does not seem to be a problem" in this setting (Pakes & McGuire 1994, p. 570). It is therefore striking that we obtain up to nine equilibria (for a given set of parameter values). Second, being able to pinpoint the driving force behind multiple equilibria is a first step towards tackling the multiplicity problem that plagues the estimation of dynamic stochastic games and inhibits the use of counterfactuals in policy analysis (see Pakes, Ostrovsky & Berry (2004) for a discussion of the issue).

In sum, learning-by-doing and organizational forgetting are distinct economic forces. Organizational forgetting, in particular, does not simply negate learning-by-doing. The unique role played by organizational forgetting comes about because it makes bidirectional movements through the state space possible. As a consequence, a model with both learning-by-doing and organizational forgetting can give rise to aggressive pricing behavior, market dominance, and multiple equilibria, whereas a model with learning-by-doing alone cannot.

We also make two methodological contributions. First, we point to a hitherto unknown weakness of the Pakes & McGuire (1994) algorithm for computing equilibria, the major tool in the literature following Ericson & Pakes (1995). Specifically, we prove that our dynamic stochastic game has equilibria that cannot be computed by it. Roughly speaking, in the presence of multiple equilibria, "inbetween" two equilibria that can be computed using the Pakes & McGuire (1994) algorithm, there is one equilibrium that cannot. This severely limits the ability of the Pakes & McGuire (1994) algorithm to provide a reasonably complete picture of the set of solutions to the model.

Second, we propose a homotopy or path-following algorithm. Our algorithm traces out the equilibrium correspondence by varying the degree of organizational forgetting and allows us to compute equilibria that cannot be computed using the Pakes & McGuire (1994) algorithm. We find that the equilibrium correspondence contains a unique path that starts at the equilibrium of the model with learning-by-doing alone. In addition, the equilibrium correspondence may contain (one or more) loops. To our knowledge, our paper is the first to describe in detail the structure of the set of equilibria of a dynamic stochastic game in the tradition of Ericson & Pakes (1995).

The organization of the remainder of the paper is as follows. Sections 2 and 3 describe the model specification and our computational strategy. Section 4 provides an overview of the equilibrium correspondence. Section 5 analyzes pricing behavior in equilibrium, and Section 6 characterizes the industry dynamics implied by it. In Section 7 we turn to the dominance properties of firms' pricing behavior. Section 8 describes how organizational forgetting can lead to multiple equilibria. Section 9 undertakes a number of robustness checks. Section 10 summarizes and concludes.

2 Model {MODEL}

For expositional simplicity, we focus on the basic model of an industry with two firms and neither entry nor exit. The general model is outlined in Section 9.

2.1 Firms and states

We consider a discrete-time, infinite-horizon dynamic stochastic game played by two firms. Firm $n \in \{1, 2\}$ is described by its state $e_n \in \{1, \ldots, M\}$. A firm's state indicates its cumulative experience or stock of know-how. By making a sale, a firm can add to its stock of know-how. Following Cabral & Riordan (1994), we take a period to be just long enough for at most one firm to make a sale. In contrast to Cabral & Riordan (1994), however, we incorporate organizational forgetting in our model as suggested by the empirical studies of Argote et al. (1990), Darr et al. (1995), Benkard (2000), and Thompson (2003). Accordingly,

the evolution of firm n's stock of know-how is governed by the law of motion

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

where e_n' and e_n is firm n's stock of know-how in the subsequent and current period, respectively, the random variable $q_n \in \{0,1\}$ indicates whether firm n makes a sale, and the random variable $f_n \in \{0,1\}$ represents organizational forgetting. If $q_n = 1$, the firm gains a unit of know-how through learning-by-doing, while it loses a unit of know-how through organizational forgetting if $f_n = 1$.

At any point in time, the industry is characterized by a vector of firms' states $\mathbf{e} = (e_1, e_2) \in \{1, \ldots, M\}^2$. We refer to $\mathbf{e}$ as the state of the industry. We use $\mathbf{e}^{[2]}$ to denote the vector (e_2, e_1) found by interchanging the stocks of know-how of firms 1 and 2.

2.2 Learning-by-doing

The marginal cost of production $c(e_n)$ of firm n depends on its stock of know-how e_n through a learning curve given by

$$c(e_n) = \begin{cases} \kappa e_n^{\eta} & \text{if} \quad 1 \leq e_n < m, \\ \kappa m^{\eta} & \text{if} \quad m \leq e_n \leq M, \end{cases}$$

where $\eta = \log_2 \rho$ for a progress ratio of $\rho \in (0, 1]$. Thus, marginal cost decreases by $1 - \rho$ percent as the stock of know-how doubles. The marginal cost of production at the top of the learning curve, $c(1)$, is $\kappa > 0$ and, in line with Cabral & Riordan (1994), m represents the stock of know-how at which a firm reaches the bottom of its learning curve.

2.3 Organizational forgetting

We let $\delta(e_n) = \Pr(f_n = 1)$ denote the probability that firm n loses a unit of know-how through organizational forgetting. We assume that this probability increases in the stock of know-how. This has several advantages. First, experimental evidence in the management literature suggests that forgetting is an increasing function of the amount learned (Bailey 1989). Second, it can be shown that the expected decay of know-how in the absence of further learning is a convex function of time as long as $\delta(e_n)$ is increasing in e_n.[2] This phenomenon, known in the psychology literature as Jost's second law, is consistent with experimental evidence on forgetting by individuals (Wixted & Ebbesen 1991). Third, in the capital-stock model employed in empirical work on organizational forgetting the amount of depreciation is assumed to be proportional to the stock of know-how. Hence, to counteract

[2] See the Online Appendix for a proof.

depreciation the addition of know-how through learning-by-doing must increase with the stock of know-how. Our specification has a similar feature but, unlike the capital-stock model, is consistent with a discrete state space.[3]

Specifically, we employ the functional form

$$\delta(e_n) = 1 - (1-\delta)^{e_n},$$

where we refer to $\delta \in [0,1]$ as the forgetting rate.[4] If $\delta > 0$, then $\delta(e_n)$ is increasing and concave in e_n; $\delta = 0$ corresponds to the absence of organizational forgetting, the special case analyzed by Cabral & Riordan (1994). Other functional forms are clearly possible, and we explore some of them in Section 9.

2.4 Demand

The industry draws its customers from a large pool of potential buyers. In each period, one buyer enters the market and makes, at most, one purchase, either from one of the two firms in the industry (inside goods 1 or 2) or chooses an alternative made from a substitute technology (outside good 0).[5] The net utility that a buyer obtains by purchasing good n is $v_n - p_n + \varepsilon_n$, where p_n is the price of good n and gross utility consists of a deterministic component v_n and a stochastic component ε_n that captures the buyer's idiosyncratic preference for good n. We assume that the deterministic component of gross utility is the same for the inside goods, $v_1 = v_2 = v$. Further, we assume that the outside good is supplied under conditions of perfect competition with price equal to marginal cost, $p_0 = c_0$.

A buyer's idiosyncratic preferences $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$ are unobservable to firms and are assumed to be independently and identically type 1 extreme value distributed with location parameter 0 and scale parameter $\sigma > 0$. The scale parameter governs the degree of horizontal product differentiation. As $\sigma \to 0$, goods become homogeneous.

The buyer purchases the good that gives it the highest net utility. Given our distributional assumptions the probability that firm n makes a sale is given by the logit specification

$$D_n(\mathsf{p}) = \frac{\exp(\frac{v-p_n}{\sigma})}{\exp(\frac{v_0-c_0}{\sigma}) + \sum_{k=1}^{2} \exp(\frac{v-p_k}{\sigma})},$$

[3] See Benkard (2004) for an alternative approximation to the capital-stock model.

[4] One way to motivate this functional form is to imagine that the stock of know-how is dispersed among a firm's workforce. In particular, assume that e_n is the number of skilled workers and that organizational forgetting is the result of labor turnover. Then, given a turnover rate of δ, $\delta(e_n)$ is the probability that at least one of the e_n skilled workers leaves the firm.

[5] Since there is a different buyer in each period, buyers are non-strategic. Lewis & Yildirim (2002) consider a model with a single buyer who optimally designs a multi-period procurement auction in order to influence the dynamics of the industry.

where $\mathsf{p} = (p_1, p_2)$ is a vector of prices. This specification has three economically meaningful parameters: v, σ, and $v_0 - c_0$. As $v_0 - c_0 \to -\infty$, $D_0(\mathsf{p}) = 1 - \sum_{n=1}^{2} D_n(\mathsf{p}) \to 0$ and we revert to the Cabral & Riordan (1994) setting in which the buyer always purchases from one of the two firms in the industry. We include an outside good because our analysis of entry and exit in Section 9 requires a well-posed monopoly problem.

2.5 State-to-state transitions

Conditional on firm n making a sale in the current period (an event denoted by w), its stock of know-how changes according to the transition function

$$\Pr(e_n'|e_n, w) = \begin{cases} 1 - \delta(e_n) & \text{if} \quad e_n' = e_n + 1, \\ \delta(e_n) & \text{if} \quad e_n' = e_n. \end{cases}$$

Conditional on firm n not making a sale (an event denoted by l), its stock of know-how changes according to the transition function

$$\Pr(e_n'|e_n, l) = \begin{cases} 1 - \delta(e_n) & \text{if} \quad e_n' = e_n, \\ \delta(e_n) & \text{if} \quad e_n' = e_n - 1. \end{cases}$$

At the upper and lower boundaries of the state space, we take the transition function to be $\Pr(M|M, w) = 1$ and $\Pr(1|1, l) = 1$, respectively.

2.6 Bellman equation

Consider an industry that is in state e. Letting $\beta \in (0, 1)$ denote the discount factor, the expected net present value of future cash flows to firm n, $V_n(\mathsf{e})$, is implicitly defined by the Bellman equation

$$V_n(\mathsf{e}) = \max_{p_n} D_n(p_n, p_{-n}(\mathsf{e}))(p_n - c(e_n)) + \beta \sum_{k=0}^{2} D_k(p_n, p_{-n}(\mathsf{e}))\overline{V}_{nk}(\mathsf{e}), \quad (1) \quad \{\text{BELLMAN}\}$$

where $p_{-n}(\mathsf{e})$ denotes the price charged by the other firm in state e and $\overline{V}_{nk}(\mathsf{e})$ is the expectation of firm n's value function conditional on the buyer purchasing good $k \in \{0, 1, 2\}$

in state e. Specifically,

$$\begin{aligned}
\overline{V}_{n0}(\mathsf{e}) &= \sum_{e_1'=e_1-1}^{e_1} \sum_{e_2'=e_2-1}^{e_2} V_n(\mathsf{e}') \Pr(e_1'|e_1,l) \Pr(e_2'|e_2,l), \\
\overline{V}_{n1}(\mathsf{e}) &= \sum_{e_1'=e_1}^{e_1+1} \sum_{e_2'=e_2-1}^{e_2} V_n(\mathsf{e}') \Pr(e_1'|e_1,w) \Pr(e_2'|e_2,l), \\
\overline{V}_{n2}(\mathsf{e}) &= \sum_{e_1'=e_1-1}^{e_1} \sum_{e_2'=e_2}^{e_2+1} V_n(\mathsf{e}') \Pr(e_1'|e_1,l) \Pr(e_2'|e_2,w).
\end{aligned}$$

Let $\Omega_n(\mathsf{e}, p_n)$ denote the maximand on the RHS of equation (1). Differentiating it with respect to p_n and using the fact that our demand specification implies

$$\frac{\partial D_k(\mathsf{p})}{\partial p_n} = \begin{cases} -\frac{1}{\sigma}(1 - D_n(\mathsf{p}))D_n(\mathsf{p}) & \text{if} \quad k = n, \\ \frac{1}{\sigma} D_k(\mathsf{p}) D_n(\mathsf{p}) & \text{if} \quad k \neq n, \end{cases}$$

we have

$$\frac{\partial \Omega_n(\mathsf{e}, p_n)}{\partial p_n} = \frac{1}{\sigma} D_n(p_n, p_{-n}(\mathsf{e})) \Big(\sigma - (p_n - c(e_n)) - \beta \overline{V}_{nn}(\mathsf{e}) + \Omega_n(\mathsf{e}, p_n) \Big). \qquad (2) \quad \{\texttt{FOC}\}$$

Differentiating $\frac{\partial \Omega_n(\mathsf{e}, p_n)}{\partial p_n}$ again and using equation (2) yields

$$\frac{\partial^2 \Omega_n(\mathsf{e}, p_n)}{\partial p_n^2} = \frac{1}{\sigma} \frac{\partial \Omega_n(\mathsf{e}, p_n)}{\partial p_n} \Big(2 D_n(p_n, p_{-n}(\mathsf{e})) - 1 \Big) - \frac{1}{\sigma} D_n(p_n, p_{-n}(\mathsf{e})).$$

Thus, the FOC $\frac{\partial \Omega_n(\mathsf{e}, p_n)}{\partial p_n} = 0$ implies $\frac{\partial^2 \Omega_n(\mathsf{e}, p_n)}{\partial p_n^2} = -\frac{1}{\sigma} D_n(p_n, p_{-n}(\mathsf{e})) < 0$, i.e., $\Omega_n(\mathsf{e}, p_n)$ is strictly quasi-concave in p_n and the pricing decision $p_n(\mathsf{e})$ therefore uniquely determined.

2.7 Equilibrium

Because our demand specification is symmetric in the sense that $D_2(\mathsf{p}) = D_1(\mathsf{p}^{[2]})$, we focus attention on symmetric Markov perfect equilibria. Existence of a symmetric equilibrium in pure strategies follows from the arguments in Doraszelski & Satterthwaite (2003). In a symmetric equilibrium, it suffices to determine the value and policy functions of firm 1, and we define $V^*(\mathsf{e}) = V_1(\mathsf{e})$ and $p^*(\mathsf{e}) = p_1(\mathsf{e})$ for each state e. We thus obtain the value and policy functions of firm 2 in state e as $V_2(\mathsf{e}) = V^*(\mathsf{e}^{[2]})$ and $p_2(\mathsf{e}) = p^*(\mathsf{e}^{[2]})$. Further, we let $\overline{V}_k^*(\mathsf{e}) = \overline{V}_{1k}(\mathsf{e})$ denote the conditional expectation of firm 1's value function and $D_k^*(\mathsf{e}) = D_k(p^*(\mathsf{e}), p^*(\mathsf{e}^{[2]}))$ the equilibrium probability that the buyer purchases good $k \in \{0, 1, 2\}$ in state e.

Given this notation, the Bellman equation and FOC for state $\mathbf{e}$ can be expressed as

$$V^*(\mathbf{e}) = D_1^*(\mathbf{e})\left(p^*(\mathbf{e}) - c(e_1)\right) + \beta \sum_{k=0}^{2} D_k^*(\mathbf{e})\overline{V}_k^*(\mathbf{e}), \tag{3}$$ {VALUE}

$$0 = \sigma - \left(1 - D_1^*(\mathbf{e})\right)\left(p^*(\mathbf{e}) - c(e_1)\right) - \beta\overline{V}_1^*(\mathbf{e}) + \beta \sum_{k=0}^{2} D_k^*(\mathbf{e})\overline{V}_k^*(\mathbf{e}). \tag{4}$$ {POLICY}

The system of equations given by the collection of equations (3) and (4) for each state $\mathbf{e} \in \{1, \ldots, M\}^2$ defines a symmetric equilibrium.

Note that if the firm is myopic, i.e., $\beta = 0$, then equation (4) reduces to the FOC of a static price-setting game. As a point of reference, consider the Nash equilibrium of this static game. The FOCs in state $\mathbf{e}$ can be re-written as

$$p_1^\dagger(\mathbf{e}) = c(e_1) + \frac{\sigma}{1 - D_1(p_1^\dagger(\mathbf{e}), p_2^\dagger(\mathbf{e}))}, \tag{5}$$ {STATICPOLICY

$$p_2^\dagger(\mathbf{e}) = c(e_2) + \frac{\sigma}{1 - D_2(p_1^\dagger(\mathbf{e}), p_2^\dagger(\mathbf{e}))}. \tag{6}$$ {STATICPOLICY

Existence and uniqueness of a static Nash equilibrium follows from Caplin & Nalebuff (1991). Moreover, $p_2^\dagger(\mathbf{e}) = p_1^\dagger(\mathbf{e}^{[2]})$, thereby allowing us to focus on firm 1 by defining $p^\dagger(\mathbf{e}) = p_1^\dagger(\mathbf{e})$ for each state $\mathbf{e}$. We also let $D_k^\dagger(\mathbf{e}) = D_k(p^\dagger(\mathbf{e}), p^\dagger(\mathbf{e}^{[2]}))$ denote the equilibrium probability that the buyer purchases good $k \in \{0, 1, 2\}$.

2.8 Parameterization

Our focus is on how learning-by-doing and organizational forgetting affect pricing behavior and the industry dynamics implied by that behavior. Accordingly, we explore the full range of values for the progress ratio ρ and the forgetting rate δ. To do so, we proceed as follows: First we specify a grid of 100 equidistant values of $\rho \in (0, 1]$. For each of them, we then trace the equilibrium as δ ranges from 0 to 1 (see Section 3 for details). This typically entails solving the model for a few thousand intermediate values of δ.

While our goal is to exhaust the parameter space (within the confines of our computational resources), we note that empirical estimates of progress ratios are in the range of 0.7 to 0.95 (Dutton & Thomas 1984). Using a capital-stock model of learning-by-doing and organizational forgetting, empirical studies estimate rates of depreciation ranging from 4 to 25 percent per month. In the context of our specification this suggests that the relevant values of δ fall below 0.1.[6]

For a large part of the paper, we fix a set of values for the remaining parameters. In Sec-

[6] See the Online Appendix for details.

tion 9 we then discuss how they influence the equilibrium and demonstrate the robustness of our conclusions regarding the critical parameters ρ and δ. In our baseline parameterization, we set $M = 30$ and $m = 15$. We assume that the marginal cost at the top of the learning curve κ is equal to 10. For a progress ratio of $\rho = 0.85$, these assumptions imply that the marginal cost of production declines from a maximum value of $c(1) = 10$ to a minimum value of $c(15) = \ldots = c(30) = 5.30$.

Turning to demand, we set $v = 10$, $\sigma = 1$, and $v_0 - c_0 = 0$. Because $v - c(1) = v_0 - c_0$, a firm at the top of the learning curve is on par with the outside good. The share of the outside good is quite small in general. Consider a progress ratio of $\rho = 0.85$. In the static Nash equilibrium, as the marginal cost of production declines, the share of the outside good declines from 0.63 in state $(1, 1)$ over 0.33 in state $(2, 2)$ and 0.15 in state $(4, 4)$ to 0.03 in state $(15, 15)$. Setting $\sigma = 1$ in our baseline parameterization gives rise to a moderate degree of horizontal product differentiation. To illustrate, in the static Nash equilibrium the own-price elasticity of demand ranges between -10.27 in state $(1, 15)$ and -2.52 in state $(15, 1)$. The cross-price elasticity of firm 1's demand with respect to firm 2's price is 0.80 in state $(15, 1)$ and 6.28 in state $(1, 15)$. In this sense the results reported below do not hinge on unrealistic parameterizations.

Finally, we set the discount factor to $\beta = \frac{1}{1.05}$. The discount factor can be thought of as $\beta = \frac{\zeta}{1+r}$, where $r > 0$ is the per-period discount rate and $\zeta \in (0, 1]$ is the exogenous probability that the industry survives from one period to the next. Consequently, our baseline parameterization corresponds to a variety of scenarios that differ in the length of a period. For example, it corresponds to a period length of one year, a yearly discount rate of 5 percent, and certain survival, but also to a period length of one month, a monthly discount rate of 1 percent (which translates into a yearly discount rate of 12.68 percent), and a monthly survival probability of 0.96. To put this scenario in perspective, technology companies such as IBM and Microsoft had costs of capital in the range of 11 to 15 percent per annum in the late 1990s. Further, an industry with a monthly survival probability of 0.96 has an expected lifetime of 26.25 months. Thus, this scenario is broadly consistent with a pace of innovative activity that is expected to make an industry's products obsolete within two to three years.

3 Computation

{COMPUTATION

In this section we first describe an algorithm for computing equilibria due to Pakes & McGuire (1994). While it is a cornerstone of the literature spawned by Ericson & Pakes (1995), we argue that it has a number of problems. Next we present our homotopy algorithm.

3.1 Pakes & McGuire (1994) algorithm

The Pakes & McGuire (1994) algorithm is intuitively appealing because it combines value function iteration as familiar from dynamic programming with best reply dynamics (akin to Cournot adjustment) as familiar from static games. The value of continued play is given by the conditional expectation of firm 1's value function. Holding $\overline{V}_0^*(\mathsf{e})$, $\overline{V}_1^*(\mathsf{e})$, and $\overline{V}_2^*(\mathsf{e})$ fixed, the strategic situation in state e is similar to a static game, and the Pakes & McGuire (1994) algorithm computes the best reply of firm 1 against $p^*(\mathsf{e}^{[2]})$ in this game. The best reply serves to update the value and policy functions of firm 1 in state e and the algorithm proceeds to the next state according to some pre-specified order. The algorithm continues to iterate until it has reached a fixed point.

More formally, let

$$\Omega_1(\mathsf{e}, p_1) = D_1(p_1, p^*(\mathsf{e}^{[2]}))\,(p_1 - c(e_1)) + \beta \sum_{k=0}^{2} D_k(p_1, p^*(\mathsf{e}^{[2]}))\overline{V}_k^*(\mathsf{e})$$

denote the maximand on the RHS of the Bellman equation. The best reply of firm 1 against $p^*(\mathsf{e}^{[2]})$ in state e is given by

$$BR(\mathsf{e}) = \arg\max_{p_1} \Omega_1(\mathsf{e}, p_1) \qquad (7) \quad \{\mathtt{BR}\}$$

and the value associated with it is

$$Val(\mathsf{e}) = \Omega_1(\mathsf{e}, BR(\mathsf{e})). \qquad (8) \quad \{\mathtt{VAL}\}$$

Let

$$\mathsf{x} = (V^*(1,1), \ldots, V^*(M,M), p^*(1,1), \ldots, p^*(M,M))$$

denote the $2M^2$ unknowns, $V^*(\mathsf{e})$ and $p^*(\mathsf{e})$ for each state $\mathsf{e} \in \{1, \ldots, M\}^2$

Given an initial guess x^0, the Pakes & McGuire (1994) algorithm executes the iteration

$$\mathsf{x}^{k+1} = G(\mathsf{x}^k), \quad k = 0, 1, 2, \ldots.$$

The algorithm aims to compute a fixed point $\mathsf{x} = G(\mathsf{x})$ by continuing to iterate until the changes in the value and policy functions of firm 1 are deemed small.[7]

The Pakes & McGuire (1994) algorithm has a number of problems. Most important, as we show in Section 4, our dynamic stochastic game has equilibria that cannot be computed by it. Roughly speaking, in the presence of multiple equilibria, "inbetween" two equilibria that can be computed using the Pakes & McGuire (1994) algorithm, there is one equilibrium that cannot. This severely limits the ability of the Pakes & McGuire (1994) algorithm to provide a reasonably complete picture of the set of solutions to the model.

The presence of multiple equilibria gives rise to another problem. Different initial guesses may or may not lead to different equilibria. Hence, to pick up more than one equilibrium, the Pakes & McGuire (1994) algorithm must be restarted from different initial guesses. This is a process of trial-and-error, and the Pakes & McGuire (1994) algorithm offers no systematic approach to computing multiple equilibria.

3.2 Homotopy algorithm

In marked contrast to the Pakes & McGuire (1994) algorithm, our approach is to treat the collection of equations (3) and (4) as a system of $2M^2$ nonlinear equations in $2M^2$ unknowns $\mathsf{x} = (V^*(1,1), \ldots, V^*(M,M), p^*(1,1), \ldots, p^*(M,M))$. Various methods such as Newton's method are available for solving a system of nonlinear equations (see e.g. Judd 1998). Unlike the Pakes & McGuire (1994) algorithm, some of these algorithms are guaranteed to converge provided that the system of nonlinear equations satisfies some conditions and the initial guess is good. Like the Pakes & McGuire (1994) algorithm, however, these algorithms are limited in their ability to compute multiple equilibria because a good initial guess must be supplied for each equilibrium.

The homotopy or path-following algorithm resolves this issue. In Section 4 we show that the equilibrium is unique if $\delta = 0$ as well as if $\delta = 1$ (under some conditions on the remaining parameters). The homotopy algorithm follows the path that connects the equilibrium at $\delta = 0$ with the equilibrium at $\delta = 1$.

The basic idea is as follows.[8] Write the system of $2M^2$ nonlinear equations that defines an equilibrium as $F(\mathsf{x}, \delta) = 0$ (holding fixed parameters other than δ). The object of interest

[7] The Pakes & McGuire (1994) algorithm is also called a pre-Gauss-Jacobi method in the literature on nonlinear equations (see Judd (1998) for details).

[8] See Zangwill & Garcia (1981) for an introduction to homotopy methods and Schmedders (1998, 1999) for an application to general equilibrium models with incomplete asset markets.

is the equilibrium correspondence $F^{-1} = \{(\mathsf{x}, \delta) | F(\mathsf{x}, \delta) = 0\}$. To trace it out, we define a parametric path to be a set of functions $(\mathsf{x}(s), \delta(s))$ such that $(\mathsf{x}(s), \delta(s)) \in F^{-1}$. Since as s varies $(\mathsf{x}(s), \delta(s))$ describes a path in F^{-1}, we have

$$F(\mathsf{x}(s), \delta(s)) = 0.$$

Differentiating both sides with respect to s yields

$$\sum_{i=1}^{2M^2} \frac{\partial F(\mathsf{x}(s), \delta(s))}{\partial x_i} x_i'(s) + \frac{\partial F(\mathsf{x}(s), \delta(s))}{\partial \delta} \delta'(s) = 0.$$

This system of $2M^2$ linear equations in $2M^2 + 1$ unknowns, $x_i'(s)$, $i = 1, \ldots, 2M^2$, and $\delta'(s)$, captures the conditions that are required to remain "on path." While there are many solutions, all of them describe the same path. One solution obeys the so-called basic differential equations

$$y_i'(s) = (-1)^{i+1} \det \left(\left(\frac{\partial F(\mathsf{y}(s))}{\partial \mathsf{y}} \right)_{-i} \right), \quad i = 1, \ldots, 2M^2 + 1, \tag{9}$$ {BDE}

where $\mathsf{y}(s) = (\mathsf{x}(s), \delta(s))$ and the notation $(\cdot)_{-i}$ is used to indicate that the ith column is removed from the $(2M^2 \times 2M^2 + 1)$ Jacobian $\frac{\partial F(\mathsf{y}(s))}{\partial \mathsf{y}}$.

The significance of the basic differential equations (9) is that they reduce the task of tracing out the equilibrium correspondence to solving a system of differential equations. This is easily done using a variety of methods (see e.g. Judd 1998). The unique equilibrium for $\delta = 0$ provides the initial condition. From there the homotopy algorithm follows the path until it reaches $\delta = 1$. Whenever $\delta'(s)$ switches sign from positive to negative, the path is "bending backward" and there are multiple equilibria. Conversely, whenever the sign of $\delta'(s)$ switches back from negative to positive, the path is "bending forward."

A major burden is computing the Jacobian; we use analytic derivatives to speed up the computations.[9] Nevertheless, the homotopy algorithm is computationally demanding and its application therefore restricted to relatively small systems of nonlinear equations. Our programs are based on Hompack (Watson, Billups & Morgan 1987) written in Fortran 77 and are available from the authors upon request.

4 Equilibrium correspondence {CORRESPONDE

In the absence of organizational forgetting, Cabral & Riordan (1994) show that the equilibrium is unique. The following proposition generalizes their result:

[9] See the Online Appendix for details.

{PROP1}

Proposition 1 *If organizational forgetting is either absent ($\delta = 0$) or certain ($\delta = 1$) and the outside good is sufficiently unattractive ($v_0 - c_0 \to -\infty$), then there is a unique equilibrium.*

Proof. See Appendix. ■

Two remarks are in order. First, note that Proposition 1 pertains to both symmetric and asymmetric equilibria. In what follows, we restrict attention to symmetric equilibria. Second, as $v_0 - c_0 \to -\infty$, we approach a setting in which the buyer always purchases from one of the two firms in the industry, similar to Cabral & Riordan (1994). Our computations indicate that Proposition 1 continues to hold even if the outside good has a nontrivial share.[10]

The cases of $\delta = 0$ and $\delta = 1$ are special in that they ensure that movements through the state space are unidirectional. Specifically, when $\delta = 0$, a firm can never move "backward" to a lower state and when $\delta = 1$, it can never move "forward" to a higher state. In contrast, when $\delta \in (0, 1)$, a firm can move in either direction. Our computations show that this has a substantive impact on the set of equilibria:

{RESULT1}

Result 1 *If organizational forgetting is neither absent ($\delta = 0$) nor certain ($\delta = 1$), then there may be multiple equilibria.*[11]

Figure 1 illustrates the extent of multiplicity. 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, there are up to nine equilibria. Multiplicity is especially pervasive for forgetting rates δ in the empirically relevant range below 0.1; indeed, we always obtained a unique equilibrium for sufficiently large forgetting rates (about $\delta \geq 0.15$; omitted in Figure 1), and similarly for sufficiently large progress ratios (about $\rho \geq 0.9$).

In dynamic stochastic games with finite actions, it has been shown that generically the number of Markov perfect equilibria is odd (Herings & Peeters 2004). While Herings & Peeters (2004) consider both symmetric and asymmetric equilibria, in a symmetric game like ours, asymmetric equilibria occur in pairs. Hence, their result immediately implies that generically the number of symmetric equilibria is odd. Figure 1 suggests that this carries over to our setting with continuous actions.

[10] Indeed, we have been unable to produce an example with multiple equilibria despite varying $v_0 - c_0$ between 0 and 15. If $v_0 - c_0 = 15$, then the outside good is so attractive that the inside goods have trivial shares.

[11] Because Proposition 1 presumes a su±ciently unattractive outside good, one might wonder whether it merely is the presence of the outside good that causes the multiple equilibria. Our computations indicate that there may be multiple equilibria even if $v_0 - c_0 = -\infty$, see Section 9.

We next take a closer look at the set of equilibria. Recall that our homotopy algorithm traces out the equilibrium correspondence $F^{-1} = \{(\mathsf{x}, \delta) | F(\mathsf{x}, \delta) = 0\}$ (holding fixed parameters other than δ). We have the following result:

{RESULT2}

Result 2 *The equilibrium correspondence F^{-1} contains a unique path that connects the equilibrium at $\delta = 0$ with the equilibrium at $\delta = 1$. In addition, F^{-1} may contain (one or more) loops that are connected to neither the equilibrium at $\delta = 0$ nor the equilibrium at $\delta = 1$.*

Figure 2 illustrates Result 2. An equilibrium is defined in terms of a value and a policy function and is thus an element of a high-dimensional space. To succinctly represent it, we use the policy function to construct the probability distribution over next period's state e' given this period's state e, i.e., the transition matrix that characterizes the Markov process of industry dynamics. We compute the limiting (or ergodic) distribution over states, $\mu^\infty(\cdot)$, implied by the equilibrium and from it the expected Herfindahl index

$$H^\infty = \sum_{\mathbf{e}} \frac{D_1^*(\mathsf{e})^2 + D_2^*(\mathsf{e})^2}{(1 - D_0^*(\mathsf{e}))^2} \mu^\infty(\mathsf{e})$$

as a summary measure of long-run industry concentration.[12] To visualize the equilibrium correspondence F^{-1}, we plot the limiting expected Herfindahl index H^∞ for forgetting rates $\delta \in [0, 1]$ and a variety of progress ratios $\rho \in \{0.95, 0.85, 0.75, 0.65, 0.55, 0.35, 0.15, 0.05\}$ in Figure 2. As can be seen, there are multiple equilibria whenever the path that connects the equilibrium at $\delta = 0$ with the equilibrium at $\delta = 1$ is bending backward. Moreover, the equilibrium correspondence contains one loop for $\rho \in \{0.75, 0.65, 0.55, 0.15, 0.05\}$ and two loops for $\rho = 0.35$, thus adding further equilibria.

A caveat is in order. To the best of our knowledge, no algorithm is guaranteed to find all equilibria,[13] and our homotopy algorithm is no exception. To follow a path or a loop, the homotopy algorithm requires an initial condition. The unique equilibrium for $\delta = 0$ provides such an initial condition, thus enabling us to find "all" equilibria along the path that connects the equilibrium at $\delta = 0$ with the equilibrium at $\delta = 1$. Unfortunately, there is no systematic approach for obtaining an initial condition for a loop that is connected to neither the equilibrium at $\delta = 0$ nor the equilibrium at $\delta = 1$. Moreover, we cannot rule out that there are more such loops. Unless we know at least one equilibrium on each loop, we are unable to pick up these additional equilibria. Neither can we be sure to pick up

[12] Since the Her¯ndahl index is based on ¯rms' conditional market shares, it is bounded below by 0.5 and above by 1.

[13] Unless the system of equations that de¯nes them happens to be polynomial. See Judd & Schmedders (2004) for some early e®orts along this line.

isolated equilibria or continua should they exist.[14] Nevertheless, the homotopy algorithm offers a substantial improvement over the trial-and-error approach to computing multiple equilibria that the Pakes & McGuire (1994) algorithm affords.

In addition, there are equilibria that cannot be computed using the Pakes & McGuire (1994) algorithm. Recall that the algorithm executes the iteration

$$\mathsf{x}^{k+1} = G(\mathsf{x}^k), \quad k = 0, 1, 2, \ldots.$$

Consider the $(2M^2 \times 2M^2)$ Jacobian $\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}}$ at a fixed point $\mathsf{x} = G(\mathsf{x})$ and let $\rho\left(\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}}\right)$ be its spectral radius. The fixed point is said to be locally stable if $\rho\left(\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}}\right) < 1$, i.e., if all eigenvalues are within the complex unit circle. In this case the Pakes & McGuire (1994) algorithm converges provided that the initial guess is sufficiently close to the fixed point. Conversely, the fixed point cannot be computed using the Pakes & McGuire (1994) algorithm if $\rho\left(\frac{\partial G(\mathbf{x})}{\partial \mathbf{x}}\right) \geq 1$.

To relate the Pakes & McGuire (1994) algorithm to our homotopy algorithm, consider a path or a loop $(\mathsf{x}(s), \delta(s))$ in the equilibrium correspondence F^{-1}. We show in the Online Appendix that

$$\frac{\partial G(\mathsf{x}(s))}{\partial \mathsf{x}} = \frac{\partial F(\mathsf{x}(s), \delta(s))}{\partial \mathsf{x}} + I, \tag{10}$$ {JAC}

where I denotes the $(2M^2 \times 2M^2)$ identity matrix. The following proposition points to a hitherto unknown weakness of the major tool in the literature following Ericson & Pakes (1995). It establishes that no equilibrium on the part of equilibrium correspondence where $\delta'(s) \leq 0$ can be computed using the Pakes & McGuire (1994) algorithm:

{PROP2}

Proposition 2 *If $\delta'(s) \leq 0$, then $\rho\left(\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}}\right) \geq 1$.*[15]

Proof. See Appendix. ■

Recall that whenever $\delta'(s)$ switches sign from positive to negative, the path that connects the equilibrium at $\delta = 0$ with the equilibrium at $\delta = 1$ is bending backward and there are multiple equilibria. Conversely, whenever the sign of $\delta'(s)$ switches back from negative to positive, the path is bending forward. Hence, holding fixed the forgetting rate, "inbetween" two equilibria with $\delta'(s) > 0$, there is one equilibrium with $\delta'(s) \leq 0$ that cannot be

[14] We have not been able to establish that our homotopy is regular, i.e., that $\frac{\partial F(\mathbf{y})}{\partial \mathbf{y}}$ has rank $2M^2$ whenever $\mathbf{y} \in F^{-1}$. Therefore, we cannot rule out isolated equilibria and continua.

[15] Local stability cannot be restored through dampening or extrapolation, i.e., executing the iteration $\mathbf{x}^{k+1} = \omega G(\mathbf{x}^k) + (1-\omega)\mathbf{x}^k$, where ω is a real scalar. To see this, note that the spectrum of the Jacobian becomes $\sigma \;\; \omega\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}} + (1-\omega)I \;\; = \omega\sigma \;\; \frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}} \;\; + 1 - \omega = \omega\sigma \;\; \frac{\partial F(\mathbf{x}(s))}{\partial \mathbf{x}} \;\; + 1$. As we show in the proof of Proposition 2, if $\delta'(s) \leq 0$, then $\frac{\partial F(\mathbf{x}(s))}{\partial \mathbf{x}}$ has at least one real nonnegative and one real negative eigenvalue. The real nonnegative eigenvalue ensures $\rho \;\; \omega\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}} + (1-\omega)I \;\; \geq 1$ for all $\omega \geq 0$, and the real negative eigenvalue ensures the same for all $\omega < 0$.

computed using the Pakes & McGuire (1994) algorithm. Similarly, a loop is necessarily composed of equilibria with $\delta'(s) > 0$ and equilibria with $\delta'(s) \leq 0$ that cannot be computed, although visualizing the sign of $\delta'(s)$ is not as easy as on the path.

The ability of the Pakes & McGuire (1994) algorithm to provide a reasonably complete picture of the set of solutions to the model is limited beyond the scope of Proposition 2. As our computations indicate, some equilibria on the part of the equilibrium correspondence where $\delta'(s) > 0$ also cannot be computed using the Pakes & McGuire (1994) algorithm:

{RESULT3}

Result 3 *If $\delta'(s) > 0$, then we may have $\rho\left(\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}}\right) \geq 1$.*

Figure 2 summarizes Proposition 2 and Result 3 by marking equilibria with $\rho\left(\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}}\right) \geq 1$ using light shades and equilibria with $\rho\left(\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}}\right) < 1$ using dark shades.

As is well-known, not all Nash equilibria of static games are stable under best reply dynamics (see e.g. Fudenberg & Tirole 1991).[16] Since the Pakes & McGuire (1994) algorithm incorporates best reply dynamics, it is reasonable to expect that this limits its usefulness. In the Online Appendix we argue that this is not the case. More precisely, we show that, holding fixed the value of continued play, the best reply dynamics are contractive (under some conditions on the parameters) and therefore converge to a unique fixed point irrespective of the initial guess. The value function iteration also is contractive holding fixed the policy function. What makes it impossible to obtain a "large" subset of equilibria using the Pakes & McGuire (1994) algorithm is thus the combination of value function iteration with best reply dynamics.

Returning to the set of equilibria of our dynamic stochastic game, generally speaking the equilibria fall into one of four categories, as exemplified by the policy functions in Figure 3. The parameter values are $\rho = 0.85$ and $\delta \in \{0, 0.03, 0.08\}$ and represent the median progress ratio across a wide array of empirical studies combined with the cases of no, low, and high organizational forgetting. The graph in the upper left panel of Figure 3 ($\rho = 0.85$ and $\delta = 0$) is typical for what we call a *flat equilibrium without well.* The policy function is very even over the entire state space. In particular, the price that a firm charges in equilibrium is fairly insensitive to its rival's stock of know-how. In a *flat equilibrium with well*, the policy function continues to be very even over most of the state space. However, price competition is intense in a neighborhood of state $(1, 1)$, which manifests itself as a "well" in the policy function (see the upper right panel of Figure 3 for the case of $\rho = 0.85$ and $\delta = 0.03$). The graph in the lower left panel of Figure 3 ($\rho = 0.85$ and $\delta = 0.03$) exemplifies a *trenchy equilibrium.* The policy function is more uneven and exhibits a "trench" along the diagonal

[16] More generally, in static games, Nash equilibria of degree -1 are unstable under any Nash dynamics, i.e., dynamics with rest points that coincide with Nash equilibria, including replicator and smooth ¯ctitious play dynamics (Demichelis & Germano 2002).

of the state space. This trench extends from state $(1, 1)$ beyond the bottom of the learning curve in state (m, m) all the way to state (M, M). Hence, in a trenchy equilibrium, price competition between firms with similar stocks of know-how is extremely intense, but price competition abates once firms become asymmetric. Finally, in an *extra-trenchy equilibrium*, the policy function not only has a diagonal trench, but it also has trenches parallel to the edges of the state space. In an extra-trenchy equilibrium, price competition between symmetric firms is extremely intense. Furthermore, due to the sideways trenches, there are also parts of the state space where the leader competes aggressively with the follower (see the lower right panel of Figure 3 for the case of $\rho = 0.85$ and $\delta = 0.08$).

The value functions in Figure 4 correspond to the policy functions in Figure 3. 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. **Steve { Please revive the classi¯cation algorithm. Add the distinction between °at equilibria with and without well (talk to David about how this may be done). Check the code to make sure that the categories are mutually exclusive and exhaustive. Run it and generate a ¯gure with four subplots corresponding to our four categories of equilibria. In each subplot indicate the presence of an equilibrium of this category in (ρ, δ)-space.**

In sum, accounting for organizational forgetting in a model of learning-by-doing leads to multiple equilibria and a rich array of pricing behaviors. In the following section, we explore them in more detail.

5 Pricing behavior

{PRICING}

Re-writing equation (4) shows that pricing behavior satisfies

$$p^*(\mathsf{e}) = c^*(\mathsf{e}) + \frac{\sigma}{1 - D_1^*(\mathsf{e})}, \qquad (11)$$

{DYNAMICPRICI

where the *virtual* marginal cost

$$c^*(\mathsf{e}) = c(e_1) - \beta\Phi^*(\mathsf{e})$$

is determined in equilibrium and equals the *actual* marginal cost $c(e_1)$ minus the discounted *prize* from winning a sale $\beta\Phi^*(\mathbf{e})$. The prize itself is given by

$$\Phi^*(\mathbf{e}) = \overline{V}_1^*(\mathbf{e}) - \overline{V}_0^*(\mathbf{e}) + \frac{D_2^*(\mathbf{e})}{1 - D_1^*(\mathbf{e})}\left(\overline{V}_0^*(\mathbf{e}) - \overline{V}_2^*(\mathbf{e})\right)$$

and has two components. First, by winning a sale, firm 1 may move further down its learning curve, which in expectation is worth $\overline{V}_1^*(\mathbf{e}) - \overline{V}_0^*(\mathbf{e})$. We call this the *advantage-building* motive. Second, firm 1 may prevent firm 2 from moving further down its learning curve, which is worth $\overline{V}_0^*(\mathbf{e}) - \overline{V}_2^*(\mathbf{e})$ times the probability $\frac{D_2^*(\mathbf{e})}{1-D_1^*(\mathbf{e})}$ that firm 2 wins the sale conditional on firm 1 not winning the sale. We call this the *advantage-defending* motive.

The prize $\Phi^*(\mathbf{e})$ is the wedge that causes dynamic pricing behavior to differ from static pricing behavior. To see this, recall that pricing behavior in a static Nash equilibrium satisfies

$$p^\dagger(\mathbf{e}) = c(e_1) + \frac{\sigma}{1 - D_1^\dagger(\mathbf{e})}. \tag{12}$$ {STATICPRICIN

Clearly, equation (11) reduces to equation (12) if either the firm is myopic, i.e., $\beta = 0$, or its prize is zero, i.e., $\Phi^*(\mathbf{e}) = 0$. The difference in firms' incentives depends on the difference in their virtual marginal costs. This difference, in turn, depends on the difference in their actual marginal costs and the difference in their prizes; Cabral & Riordan (1994) refer to these as the *cost effect* and the *prize effect*, respectively. It is conceivable that a firm has a disadvantage over its rival in terms of actual marginal cost but an advantage in terms of its prize so that, on balance, its virtual marginal cost is lower than that of its rival.

Comparing equation (11) with equation (12), it is seen that equilibrium prices $p^*(\mathbf{e})$ and $p^*(\mathbf{e}^{[2]})$ coincide with the prices that obtain in a static Nash equilibrium with costs set to equal virtual marginal costs $c^*(\mathbf{e})$ and $c^*(\mathbf{e}^{[2]})$. Since in the static Nash equilibrium prices are increasing in either firm's cost (Vives 1999), it follows that as long as both firms' prizes are nonnegative, equilibrium prices are bounded above by the prices that obtain in a static Nash equilibrium with costs set to equal actual marginal costs $c(e_1)$ and $c(e_2)$. More formally, if $\Phi^*(\mathbf{e}) \geq 0$ and $\Phi^*(\mathbf{e}^{[2]}) \geq 0$, then $p^*(\mathbf{e}) \leq p^\dagger(\mathbf{e})$ and $p^*(\mathbf{e}^{[2]}) \leq p^\dagger(\mathbf{e}^{[2]})$.

A sufficient condition for $\Phi^*(\mathbf{e}) \geq 0$ for each state $\mathbf{e}$ is that the value function $V^*(e_1, e_2)$ is nondecreasing in e_1 and nonincreasing in e_2. Intuitively, it cannot hurt firm 1 if it moves down its learning curve and it cannot benefit firm 1 if firm 2 moves down its learning curve. If so, then $\overline{V}_1^*(\mathbf{e}) - \overline{V}_0^*(\mathbf{e}) \geq 0$ and $\overline{V}_0^*(\mathbf{e}) - \overline{V}_2^*(\mathbf{e}) \geq 0$, and the advantage-building and advantage-defending motives work together to ensure a nonnegative prize and thus $p^*(\mathbf{e}) \leq p^\dagger(\mathbf{e})$. While there are no guarantees, in the absence of organizational forgetting equilibrium prices have always been bounded above by the static Nash equilibrium in our computations:

{RESULT11}

Result 4 *If organizational forgetting is absent ($\delta = 0$), then we have $p^*(\mathbf{e}) \leq p^\dagger(\mathbf{e})$ for all $\mathbf{e} \in \{1, \ldots, M\}^2$.*

Result 4 highlights the fundamental economics of learning-by-doing: as long as improvements in competitive position are valuable, firms use price cuts as investments to achieve them.

In the presence of organizational forgetting, by contrast, pricing behavior is much more intricate:

{RESULT12}

Result 5 *If organizational forgetting is present ($\delta > 0$), then we may have $p^*(\mathbf{e}) > p^\dagger(\mathbf{e})$ for some $\mathbf{e} \in \{1, \ldots, M\}^2$.*

Figure 5 illustrates Result 5 by plotting the share of equilibria that violate the price ceiling of Result 4.[17] Darker shades indicate higher shares. As can be seen, the price ceiling remains in effect if organizational forgetting is very weak ($\delta \to 0$). Unsurprisingly, the price ceiling is also effective if learning-by-doing is very weak ($\rho \to 1$). Apart from these extremes (and a region around $\delta = 0.6$ and $\rho = 0.5$), however, at least some, if not all, equilibria fail to obey the price ceiling. This happens commonly for forgetting rates δ in the empirically relevant range below 0.1.

At first glance, Result 5 suggests that organizational forgetting makes firms less aggressive. This is intuitive. After all, why invest in improvements in competitive position when they are transitory? Surprisingly, however, it turns out that organizational forgetting is a source of aggressive pricing behavior. Proposition 3 provides us with a benchmark:

{PROP11}

Proposition 3 *If organizational forgetting is absent ($\delta = 0$), then we have (i) $p^*(\mathbf{e}) = p^\dagger(\mathbf{e}) = p^\dagger(m, m) > c(m)$ for all $\mathbf{e} \in \{m, \ldots, M\}^2$; (ii) $p^*(\mathbf{e}) > c(m)$ for all $e_1 \in \{m, \ldots, M\}$ and $e_2 \in \{1, \ldots, m-1\}$.*

Proof. See Appendix. ■

If $\delta = 0$, then the prize reduces to

$$\Phi^*(\mathbf{e}) = V^*(e_1 + 1, e_2) - V^*(e_1, e_2) + \frac{D_2^*(\mathbf{e})}{1 - D_1^*(\mathbf{e})} \left(V^*(e_1, e_2) - V^*(e_1, e_2 + 1) \right).$$

Once both firms reach the bottom of their learning curves, there are no further improvements in competitive position. In state (m, m), for example, $V^*(m + 1, m) = V^*(m, m) =$

[17] To take into account the limited precision of our computations, we take the price ceiling to be violated if $p^*(\mathbf{e}) > p^\dagger(\mathbf{e}) + \epsilon$ for some $\mathbf{e} \in \{1, \ldots, M\}^2$, where ϵ is positive but small. Speci¯cally, we set $\epsilon = 10^{-2}$, so that if prices are measured in dollars, then the price ceiling must be violated by more than a cent. Given that the homotopy algorithm solves the system of equations up to a maximum absolute error of about 10^{-12}, Figure 5 therefore almost certainly understates the extent of violations.

$V^*(m, m+1)$, and the advantage-building and advantage-defending motives disappear. Consequently, equilibrium prices coincide with prices in the static Nash equilibrium. In a static Nash equilibrium, in turn, price is set above cost.

If one but not the other firm has reached the bottom of its learning curve, then the leader no longer has an advantage-building motive but he continues to have an advantage-defending motive. This raises the possibility that the leader uses price cuts to delay the follower's progress in moving down its learning curve. As part (ii) of Proposition 3 shows, there is a limit to how aggressive the leader can be in defending its advantage: below-cost-pricing is impossible.

By establishing a price floor, Proposition 3 provides a hint that in the absence of organizational forgetting a leader is not overly aggressive in defending its advantage. The next results leaves no doubt:

{RESULT13}

Result 6 *If organizational forgetting is present ($\delta > 0$), then we may have (i) $p^*(\mathbf{e}) < p^\dagger(\mathbf{e})$ for some $\mathbf{e} \in \{m, \ldots, M\}^2$; (ii) $p^*(\mathbf{e}) \leq c(m)$ for some $e_1 \in \{m, \ldots, M\}$ and $e_2 \in \{1, \ldots, m-1\}$.*

Figure 6 illustrates Result 6. It depicts the share of equilibria that violate the price floor. As can be seen, unless organizational forgetting or learning-by-doing is very weak at least some, if not all, equilibria fail to obey the price floor. That is, the leader may be more aggressive in defending its advantage in the presence of organizational forgetting than in its absence.

Because organizational forgetting makes improvements in competitive position from learning-by-doing more transitory, one might expect that it makes firms more reluctant to invest in the acquisition of know-how through price cuts. We reach the opposite conclusion: organizational forgetting tends to make firms more instead of less aggressive. The most dramatic expression of this are the wells and trenches in the policy function (see again Figure 3).

Wells (as seen in the upper right panel of Figure 3) are preemption battles that are fought by firms at the top of their learning curves. A well serves to build a competitive advantage as both firms use price cuts in the hope of being the first to move down the learning curve. Once one firm has moved ahead of the other, both the leader and the follower raise their price. The follower, in fact, surrenders by setting a much higher price than the leader. Yet, once the follower starts to move down its learning curve, the leader makes no attempt to defend its position. The competitive advantage is thus of a transitory nature.

While a leader is not overly aggressive in defending its advantage in the absence of organizational forgetting, the price floor in Proposition 3 does not apply at the top of the

learning curve. Indeed, in our computations the equilibria in case of $\delta = 0$ have always been flat either without or with well depending on the progress ratio. For sufficiently large (about $\rho > 0.8$) or small (about $\rho < 0.3$) progress ratios we obtain flat equilibria without well; for intermediate progress ratios, we obtain flat equilibria with well.

Diagonal trenches (as seen in the lower panels of Figure 3) are price wars that are fought by fairly symmetric firms. Like a well, a diagonal trench serves to build a competitive advantage. Unlike a well, however, a diagonal trench also serves to defend it. Aggressive pricing is not confined to the top of the learning curve. On the contrary, it takes place along the diagonal, even in states at the bottom of the learning curve, as each firm uses price cuts to push the state to "its" side of the diagonal and keep it there.

An immediate consequence of the price floor in part (i) of Proposition 3 is to rule out diagonal trenches (and thus trenchy and extra-trenchy equilibria). Indeed, in the absence of organizational forgetting, prices are flat if both firms are at the bottom of their learning curves. In the presence of organizational forgetting, in contrast, firms may compete fiercely even though they have already exhausted the gains from learning-by-doing. Diagonal trenches are about acquiring and maintaining a permanent competitive advantage.

We can employ backward induction logic to gain intuition about the link between organizational forgetting and diagonal trenches. Consider state (e, e), where $e \geq m$, on the diagonal of the state space at or beyond the bottom of the learning curve. From part (i) of Proposition 3, without organizational forgetting, the advantage-building and advantage-defending motives disappear and equilibrium prices coincide with prices in the static Nash equilibrium. However, with organizational forgetting, the advantage-building and advantage-defending motives continue to operate. The advantage-building motive operates in state (e, e) because by winning a sale, the firm adds a "buffer stock" of know-how that provides insurance against future loses of know-how due to organizational forgetting. The advantage-defending motive operates because by winning a sale, the firm increases the likelihood that its rival slides back up its learning curve. Thus, organizational forgetting predisposes firms to compete fiercely even though they have already exhausted the gains from learning-by-doing.

Next consider the adjacent state $(e, e-1)$ in which firm 1 has a slight lead over firm 2. With organizational forgetting, the leader's prize from winning a sale is likely greater than the follower's. The reason is that winning a sale tends to benefit the leader whereas winning a sale tends to hurt the follower. Suppose the follower wins. In this case the follower may overtake the leader if the industry moves to state $(e-1, e)$, but there is also the possibility that the industry moves back to state (e, e) where there is brutal price competition. Such an "improvement" in competitive position thus tends to hurt the follower as indicated by the diagonal trenches in the value functions in the lower panels of Figure 4. By contrast,

ρ	0.95	0.85	0.75	0.65	0.55	0.35	0.15	0.05
δ	0.17	0.38	0.56	0.67	0.74	0.83	0.86	0.88

Table 1: Critical value of δ given ρ. TAB11

if the leader wins, then the industry is guaranteed not move back to state (e, e), thereby avoiding the brutal price competition in this state. The leader also, of course, avoids the possibility of losing its competitive advantage by precluding movements to states such as $(e-1, e)$ and $(e-1, e-1)$. And the leader may even enhance its competitive advantage by inducing movements to states such as $(e, e-2)$, $(e+1, e-1)$, and $(e+1, e-2)$. Due to this difference in prizes, even though it has at most a small advantage in terms of actual marginal cost, the leader has a large advantage over the follower in terms of virtual marginal cost. Hence, the leader substantially underprices the follower and defends its position. As a result the leadership position in state $(e, e-1)$ is more secure in the presence of organizational forgetting than in its absence.

This has implications for adjacent states. Because the leadership position in state $(e, e-1)$ is so secure, in state $(e-1, e-1)$ both firms fight hard to attain it, thereby intensifying price competition in this state. In addition, the follower in states such as $(e, e-2)$ and $(e, e-3)$ has a weaker incentive to fight hard to improve its competitive position knowing that being in state $(e, e-1)$ is not especially attractive. These effects cascade through the state space and give rise to diagonal trenches.

In sum, diagonal trenches are self-reinforcing mechanisms that lead to market dominance: if the leadership position is aggressively defended, symmetric firms fight a price war to attain it. This provide all the more reason to aggressively defend a competitive advantage because if it is lost and the industry moves back to the diagonal of the state space, then a price war ensues.

There is a limit to the enhancement of price-cutting incentives through organizational forgetting. If the forgetting rate is too large, then organizational forgetting stifles investment in learning-by-doing altogether. Indeed, if δ exceeds the critical value listed in Table 1 for a variety of progress ratios $\rho \in \{0.95, 0.85, 0.75, 0.65, 0.55, 0.35, 0.15, 0.05\}$, then firms cannot expect to make it down their learning curves. In this case, equilibrium prices at the top of the learning curve are close to prices in the static Nash equilibrium. However, price competition at the bottom of the learning curve is extremely intense as both firms seek to reduce the chance of being the first to slide back up the learning curve.

Sideways trenches (as seen in the lower right panel of Figure 3) are price wars that are fought by fairly asymmetric firms. They are triggered when a distant follower starts to move down its learning curve. While diagonal trenches are about fighting an imminent

threat, sideways trenches are about fighting a distant threat. Put more informally, they are an equilibrium manifestation of the dictum "Only the paranoid survive."

Similar to diagonal trenches, sideways trenches are about acquiring and maintaining a permanent competitive advantage. In fact, a sideways trench can be thought of as an endogenously arising mobility barrier (Caves & Porter 1977). If the follower crashes through this mobility barrier, it moves from being a docile competitor to being a viable threat for the leader. To avoid this, the leader stalls the follower.

The intuition for sideways trenches is somewhat different than for diagonal trenches. Consider state (e_1, e_2), where $e_1 > e_2$ so that firm 1 leads and firm 2 follows. As long as the probability of winning the sale is much less than $\delta(e_2)$, the probability of losing a unit of know-how through organizational forgetting, the follower is stalled and has practically no chance of ever making it down its learning curve.[18] Next consider the adjacent state $(e_1, e_2 + 1)$. The leader enjoys a large prize from winning a sale because the industry may move back to state (e_1, e_2) where the follower is stalled. The follower wants to stave off this fate, and so it, too, enjoys a large prize from winning a sale. The result is that both firms have strong incentives to cut prices in state $(e_1, e_2 + 1)$. Again, this is self-reinforcing: since being in $(e_1, e_2 + 1)$ is not especially attractive, the follower has little reason to cut price in state (e_1, e_2) in order to avoid being stalled there in the first place. Finally, if the leader is able to stall the follower in state (e_1, e_2), then the leader is able to stall the follower in state $(e_1 + 1, e_2)$, where it enjoys an even lower marginal cost, so that the trench extends parallel to the edge of the state space.

We close this section with a general point about organizational forgetting. Unlike the model of learning-by-doing without organizational forgetting, a firm can move both forward to a higher state and backward to a lower state whenever $\delta \in (0, 1)$. The possibility of bidirectional movements through the state space opens up new strategic possibilities for firms that work to enhance the advantage-building and advantage-defending motives. By winning a sale, a firm makes itself less vulnerable to future losses from organizational forgetting, thus enhancing the advantage-building motive. At the same time, it makes its rival more vulnerable to future losses from organizational forgetting, thus enhancing the advantage-defending motive. Because these additional benefits (like the benefits from learning-by-doing) are achieved by winning a sale, organizational forgetting creates strong incentives to cut prices and is thus a source of aggressive pricing behavior.

[18] Short of forcing it to exit, stalling the follower is perhaps the best the leader can do.

6 Industry dynamics

{DYNAMICS}

To study industry dynamics, we use the policy function to construct the probability distribution over next period's state $\mathbf{e}'$ given this period's state $\mathbf{e}$, i.e., the transition matrix that characterizes industry dynamics in equilibrium. This allows us to use tools from 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 when both firms were at the top of their learning curves at the outset of the game. In addition, we compute the limiting (or ergodic) distribution over states, $\mu^\infty(\cdot)$.[19] The transient distribution captures short-run dynamics and the limiting distribution captures long-run (or steady-state) dynamics.

Figures 7 and 8 display the transient distribution in period 8 and 32, respectively, and Figure 9 displays the limiting distribution for our four typical cases (as seen in Figures 3 and 4).[20] In the *flat equilibrium without well* ($\rho = 0.85$, $\delta = 0$, see upper left panels), the transient and limiting distributions are unimodal. The most likely industry structure is symmetric (or close to it). For example, the modal state is $(5, 5)$ in period 8, $(9, 9)$ in period 16, the modal states are $(16, 17)$ and $(17, 16)$ in period 32, and the model state is $(30, 30)$ in period 64. Turning from the short run to the long run, the industry ends up in state $(30, 30)$ because, in the absence of organizational forgetting, both firms must eventually reach the bottom of their learning curves. In short, the industry starts symmetric and stays symmetric.

By contrast, in the *flat equilibrium with well* for ($\rho = 0.85$, $\delta = 0.03$, see upper right panels) the transient distributions are first bimodal and then unimodal as is the limiting distribution. The modal states are $(1, 8)$ and $(8, 1)$ in period 8, $(3, 11)$ and $(11, 3)$ in period 16, $(7, 15)$ and $(15, 7)$ in period 32, but the modal state is $(16, 16)$ in period 64 and the modal state of the limiting distribution is $(22, 22)$. That is, the industry evolves first towards an asymmetric structure and then towards a symmetric structure. The reason is that the well serves to build, but not to defend, a competitive advantage, so that asymmetries vanish as time passes.

While the modes of the transient distributions are more separated and pronounced in the *trenchy equilibrium* ($\rho = 0.85$, $\delta = 0.03$, see lower left panels) than in the flat equilibrium with well, the dynamics of the industry are similar at first. Unlike in the flat

[19] Let P be the $M^2 \times M^2$ transition 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 tth matrix power of P. If $\delta \in (0, 1)$, then the Markov process is 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 $\mu^\infty = \mu^\infty P$. If $\delta = 0$ ($\delta = 1$), then there is also a single closed communicating class, but its sole member is state (M, M) ($(1, 1)$).

[20] To avoid clutter, we do not graph states that have probability of less than 10^{-4}.

equilibrium with well, however, the industry continues to evolve towards an asymmetric structure. The modal states are $(12, 19)$ and $(19, 12)$ in period 64 and the modal states of the limiting distribution are $(18, 25)$ and $(25, 18)$. Asymmetries persist as time passes because the diagonal trench serves to build and to defend a competitive advantage.

In the *extra-trenchy equilibrium* ($\rho = 0.85$, $\delta = 0.08$, see lower right panels), due to the sideways trench one firm never makes it down from the top of its learning curve. The transient and limiting distributions are bimodal, and the most likely industry structure extremely asymmetric. The modal states are $(1, 6)$ and $(6, 1)$ in period 8 , $(1, 9)$ and $(9, 1)$ in period 16, $(1, 13)$ and $(13, 1)$ in period 32, $(1, 16)$ and $(16, 1)$ in period 64. The modal states of the limiting distribution are $(1, 19)$ and $(19, 1)$. In short, one firm acquires a competitive advantage early on and maintains it throughout.

We use the transient distribution over states in period t, $\mu^t(\cdot)$, to compute the expected Herfindahl index

$$H^t = \sum_{\mathbf{e}} \frac{D_1^*(\mathsf{e})^2 + D_2^*(\mathsf{e})^2}{(1 - D_0^*(\mathsf{e}))^2} \mu^t(\mathsf{e}).$$

The time path of the expected Herfindahl index summarizes the implications of learning-by-doing and organizational forgetting for the dynamics of the industry. To the extent that the industry evolves in a more or less symmetric fashion, the expected Herfindahl index is close to 0.5 at all times. If, by contrast, asymmetries arise, then the expected Herfindahl index exceeds 0.5. The maximum expected Herfindahl index

$$\widehat{H} = \max_{t \in \{1,\ldots,100\}} H^t$$

is thus a summary measure of short-run industry concentration. Turning to the long run, the limiting expected Herfindahl index H^∞ is a summary measure of industry concentration. Since it is computed from the limiting distribution over states, $\mu^\infty(\cdot)$, H^∞ exceeds 0.5 if asymmetries persist.

Figure 10 presents our summary measures of industry concentration for forgetting rates $\delta \in [0, 1]$ and a variety of progress ratios $\rho \in \{0.95, 0.85, 0.75, 0.65, 0.55, 0.35, 0.15, 0.05\}$. H^∞ is plotted as a solid line, $\widehat{H}$ as a dashed line. Figure 10 illustrates the fundamental economics of organizational forgetting. If organizational forgetting is sufficiently weak ($\delta \to 0$), then asymmetries may arise but they cannot persist, i.e., $\widehat{H} \geq 0.5$ and $H^\infty \approx 0.5$. Moreover, if asymmetries arise in the short run, they are modest. If organizational forgetting is sufficiently strong ($\delta \to 1$), then asymmetries cannot arise in the first place, i.e., $\widehat{H} \approx H^\infty \approx 0.5$. The reason is that organizational forgetting stifles investment in learning-by-doing altogether. By contrast, for intermediate degrees of organizational forgetting, asymmetries arise and persist. Even extreme asymmetries akin to near-monopoly are pos-

sible. This is because organizational forgetting in effect commits the leader to aggressively defending its position against imminent and distant threats. This more than offsets the increased vulnerability to organizational forgetting as the stock of know-how grows and makes the leadership position more secure than it would have been in the absence of organizational forgetting.[21]

Contrary to what one might expect, organizational forgetting does not negate learning-by-doing. Rather, as can be seen in Figure 10, at least over the empirically relevant range of progress ratios ρ between 0.70 and 0.95 and forgetting rates δ below 0.1, learning-by-doing and organizational forgetting reinforce each other: starting from the absence of both learning-by-doing ($\rho = 1$) and organizational forgetting ($\delta = 0$), a steeper learning curve, i.e., a lower progress ratio, tends to lead to a more asymmetric industry structure just as a higher forgetting rate does.

7 Dominance properties {DOMINANCE}

Common intuition suggests that learning-by-doing leads by itself to market dominance by giving a more experienced firm the ability to profitably underprice its less experienced rival. This enables the leader to widen its competitive advantage over time, thereby further enhancing its ability to profitably underprice the follower. Cabral & Riordan (1994) formalize this idea with "two concepts of self-reinforcing market dominance" (p. 1115), increasing dominance (ID) and increasing increasing dominance (IID). The equilibrium exhibits ID if $p^*(\mathbf{e}) < p^*(\mathbf{e}^{[2]})$ whenever $e_1 > e_2$ and IID if $p^*(\mathbf{e}) - p^*(\mathbf{e}^{[2]})$ is decreasing in e_1. If ID holds, the leader charges a lower price than the follower and therefore enjoys a higher probability of making a sale. If IID holds, the gap between the leader's price and the follower's price widens with the length of the lead. Athey & Schmutzler (2001) extend this idea to dynamic games with deterministic state-to-state transitions. Their notion of weak increasing dominance describes the relationship between players' states and their actions and coincides with Cabral & Riordan's (1994) notion of ID.

In the absence of organizational forgetting, Cabral & Riordan (1994) show that ID and IID hold provided that the discount factor β is sufficiently close to 1 (or, alternatively, sufficiently close to 0). Their main result carries over to our parameterization:

{RESULT31}

Result 7 *If organizational forgetting is absent ($\delta = 0$), then IID holds. Thus, ID holds.*

Even though the equilibrium satisfies ID and IID, it is not clear that the industry is inevitably progressing towards monopolization. If the price gap is small, then the effect of

[21] Since the Markov process is irreducible if $\delta \in (0, 1)$, it is inevitable that the follower eventually overtakes the leader. However, as a practical matter, the expected time to a role reversal is so large that this almost never occurs.

ID and IID may be trivial.[22] In such a scenario, the leader charges a *slightly* lower price than the follower and this gap widens *a bit* over time. However, with even a modest degree of horizontal product differentiation, the firms still split sales more or less equally and thus move down the learning curve in tandem. Consequently, ID and IID have no discernible impact on industry structure and dynamics, neither in the short run nor in the long run. This is exactly what happens in the absence of organizational forgetting. For example, the flat equilibrium without well ($\rho = 0.85$, $\delta = 0$) satisfies IID and thus ID. Yet, the industry is likely to be a symmetric duopoly at all times. More generally, as Figure 10 shows, in the absence of organizational forgetting asymmetries are modest if they arise at all. In fact, although ID and IID hold, the maximum expected Herfindahl index is 0.68 (attained at $\rho = 0.60$). Hence, ID and IID are not sufficient for economically meaningful market dominance.

ID and IID are also not necessary for market dominance. To give an example, the extra-trenchy equilibrium ($\rho = 0.85$, $\delta = 0.08$) violates ID and thus IID because the leader coasts if it is far ahead of the follower and charges a higher price. Yet, the industry is likely to be a near-monopoly at all times. More generally, ID and IID may fail in the presence of organizational forgetting:

{RESULT32}

Result 8 *If organizational forgetting is present ($\delta > 0$), then (i) IID may fail; (ii) ID may fail.*

Figure 11 illustrates Result 8 by plotting the share of equilibria that violate IID (upper panel) and ID (lower panel). As can be seen, all equilibria fail to obey IID unless organizational forgetting or learning-by-doing is very weak. Even violations of ID are extremely common, especially for forgetting rates δ in the empirically relevant range below 0.1.

The broad points are this: First, while the empirical studies of Argote et al. (1990), Darr et al. (1995), Benkard (2000), and Thompson (2003) warrant the addition of organizational forgetting to Cabral & Riordan's (1994) model of learning-by-doing, their main result regarding market dominance does not seem to be particularly robust. Second, even if ID and IID hold, they may not tell us as much as we like about the structure and dynamics of an industry. Because they appear to be neither necessary nor sufficient for economically meaningful market dominance, making inferences about industry structure on the basis of ID and IID may be misleading. In contrast, wells and trenches are tell-tale signs of market dominance. However, to determine whether the industry is on a path towards monopolization, examining firms' pricing behavior is plainly not enough. Rather it is crucial to directly examine the dynamics implied by that behavior.

[22] Indeed, Cabral & Riordan (1994) show that $p^*(\mathbf{e}) \rightarrow p^\dagger(m, m)$ for all $\mathbf{e} \in \{1, \ldots, M\}^2$ as $\beta \rightarrow 1$, i.e., both ¯rms price *as if* at the bottom of their learning curves. This suggests that the price gap may be small for \reasonable" discount factors.

	preemption battle (well)	price war triggered by imminent threat (diagonal trench)	price war triggered by distant threat (sideways trench)	short-run market dominance	long-run market dominance
flat eqbm. without well	no	no	no	no	no
flat eqbm. with well	yes	no	no	yes	no
trenchy eqbm.	no	yes	no	yes	yes, modest
extra-trenchy eqbm.	no	yes	yes	yes	yes, extreme

Table 2: Pricing behavior and industry dynamics. TAB41

8 Organizational forgetting and multiple equilibria {MULTIPLICIT

The four different categories of equilibria that we have identified in Section 4 give rise to a rich array of pricing behaviors and, in turn, industry dynamics as discussed in Sections 5 and 6, respectively. Table 2 provides a summary. Generally speaking, "trenchier" equilibria are associated with more aggressive behavior and more concentrated industries, both in the short run and in the long run. Moreover, as we have argued, rather than impeding it, organizational forgetting facilitates aggressive behavior.

While the equilibrium is unique if organizational forgetting is either absent ($\delta = 0$) or certain ($\delta = 1$), there may be multiple equilibria for intermediate forgetting rates. Surprisingly, these equilibria range from "peaceful coexistence" to "trench warfare." Consequently, in addition to the degree of organizational forgetting, the equilibrium by itself is an important determinant of pricing behavior and industry dynamics.

Why do multiple equilibria arise in our model? To explore this question, it is useful to think about the strategic situation faced by firms in setting prices in state $\mathbf{e}$. The value of continued play to firm n is given by the conditional expectation of its value function, $\overline{V}_{n0}(\mathbf{e})$, $\overline{V}_{n1}(\mathbf{e})$, and $\overline{V}_{n2}(\mathbf{e})$. Holding the value of continued play fixed, the strategic situation in state $\mathbf{e}$ is thus akin to a static game. If the reaction functions in this game intersect more than once, then multiple equilibria arise. On the other hand, we say that the model satisfies stagewise uniqueness if the reactions functions of the two firms intersect once irrespective of the value of continued play. This is indeed the case provided that the outside good is sufficiently unattractive:

{PROP31}

Proposition 4 *If the outside good is sufficiently unattractive ($v_0 - c_0 \to -\infty$), then statewise uniqueness holds.*

Proof. See Appendix. ■

Note that the proof of Proposition 4 relies on the functional form of demand. This is reminiscent of the restrictions on demand (e.g., log-concavity) that Caplin & Nalebuff (1991) set forth to guarantee uniqueness of Nash equilibrium in static price-setting games.

If the outside good has a nontrivial share, then stagewise uniqueness cannot be ascertained without knowing the specific magnitudes of $\overline{V}_{n0}(\mathbf{e})$, $\overline{V}_{n1}(\mathbf{e})$, and $\overline{V}_{n2}(\mathbf{e})$ for each firm $n \in \{1, 2\}$. We therefore check for each state $\mathbf{e} \in \{1, \ldots, M\}^2$ that the reaction functions intersect once using the computed value function to determine the value of continued play, i.e., presuming the value of continued play that is attained in equilibrium. The reaction functions always had a unique intersection. **Steve { Could you please verify this.** This suggests that stagewise uniqueness continues to hold even if the outside good has a nontrivial share.

Given that the model satisfies stagewise uniqueness, multiple equilibria must arise from firms' expectations regarding the value of continued play. To see this, consider again state $\mathbf{e}$. The intersection of the reaction functions constitutes a Nash equilibrium in prices in a subgame in which firm n believes that its value of continued play is $\overline{V}_{n0}(\mathbf{e})$, $\overline{V}_{n1}(\mathbf{e})$, and $\overline{V}_{n2}(\mathbf{e})$. If firms have rational expectations, i.e., if the conjectured value of continued play is actually attained, then these prices constitute an equilibrium of our dynamic stochastic game. In other words, taking the value of continued play as given, the reaction functions intersect once, but there is more than one value of continued play that is consistent with rational expectations. Interestingly, in our model multiple equilibria do not arise because of the demand specification; instead, multiplicity is rooted in the dynamics of the model.

The key driver of multiplicity is organizational forgetting. Dynamic competition with learning-by-doing and organizational forgetting can be likened to racing down an upward-moving escalator. Unless a firm makes sales at a rate that exceeds the rate at which it loses know-how through organizational forgetting, its marginal cost is bound to increase. The "inflow" of know-how into the industry is almost one unit per period because the share of the outside good is small, whereas in expectation the "outflow" in state $\mathbf{e}$ is $\delta(e_1) + \delta(e_2)$. Consider state (e, e), where $e \geq m$, on the diagonal of the state space at or beyond the bottom of the learning curve. If $1 \ll 2\delta(e)$, then it is virtually impossible that both firms reach the bottom of their learning curves. Knowing this, firms have no choice but to price aggressively. The result is trench warfare as each firm uses price cuts to push the state to "its" side of the diagonal and keep it there. If $1 \gg 2\delta(e)$, then it is virtually inevitable that both firms reach the bottom of their learning curves, and firms may as well price softly. In both cases, the primitives of the model tie down the equilibrium.

This is no longer the case if $1 \approx 2\delta(e)$, and the stage is set for multiple equilibria as diverse as peaceful coexistence and trench warfare. If firms believe that they cannot

peacefully coexist at the bottom of their learning curves and that instead one firm comes to dominate the market, then both firms cut their prices in the hope of acquiring a competitive advantage early on and maintaining it throughout, thereby leading to trench warfare and market dominance. However, if firms believe that they can peacefully coexist at the bottom of their learning curves, then neither firm cuts its price. Soft pricing, in turn, ensures that the anticipated symmetric industry structure actually emerges. A back-of-the-envelope calculation is reassuring here: recall that $m = 15$ and note that $1 = 2\delta(15)$ implies $\delta = 0.05$, $1 = 2\delta(20)$ implies $\delta = 0.03$, and $1 = 2\delta(30)$ implies $\delta = 0.02$. This is indeed the range of forgetting rates where multiplicity prevails (see again Figure 1).

In general, a sufficient condition for uniqueness of equilibrium in a dynamic stochastic game is that the model satisfies stagewise uniqueness and that movements through the state space are unidirectional. Stagewise uniqueness precludes players' actions from giving rise to multiple equilibria and unidirectional movements preclude their expectations from doing so. To illustrate, recall that in the game at hand a firm can never move backward to a lower state if $\delta = 0$. Hence, once the industry reaches state (M, M), it remains there forever, so that the value of future play in state (M, M) coincides with the value of being in this state. In conjunction with stagewise uniqueness, this uniquely determines the value of being in state (M, M). Next consider states $(M-1, M)$ and $(M, M-1)$. The value of future play in states $(M-1, M)$ and $(M, M-1)$ depends on the value of being in state (M, M). Stagewise uniqueness ensures that firms' prices in states $(M-1, M)$ and $(M, M-1)$ as well as the value of being in these states are uniquely determined. Further, since the value of future play in state $(M-1, M-1)$ depends on the value of being in states $(M-1, M)$ and $(M, M-1)$, firms' prices and the value of being in state $(M-1, M-1)$ are uniquely determined. Continuing to work backwards in this fashion establishes that the equilibrium is unique. If $\delta = 1$, a firm can never move forward to a higher state. A similar argument establishes uniqueness of equilibrium except that in the case of $\delta = 1$ the argument is anchored in state $(1, 1)$ rather than state (M, M).

Steve { Consider two equilibria with value functions V^{*1}(e) and V^{*2}(e), respectively. We say that the ¯rst equilibrium dominates the second i® $V^{*1}(e_1, 1) > V^{*2}(e_1, 1)$ and $V^{*1}(1, e_1) > V^{*2}(1, e_1)$ for all $e_1 \in \{1, \ldots, M\}$. Please plot the points in (δ, ρ)-space where there exists an equilibrium that dominates all other equilibria. (Of course, this is only applicable to points in (δ, ρ)-space where there are multiple equilibria.) Mark { If the results work out, please add your selection story here.

9 Robustness checks {ROBUSTNESS}

In this section, we discuss how the model specification and parameterization affect our results. In the interest of brevity, we confine ourselves to summarizing our robustness checks; details can be found in the Online Appendix.

9.1 Outside good

Eliminating the outside good by setting $v_0 - c_0 = -\infty$ has almost no impact on our results. Compared to our baseline parameterization the value and policy functions remain essentially unchanged as do the measures of short-run and long-run industry concentration. ***** Uli { Maybe cite some distance measures here. ***** In particular, multiple equilibria continue to arise in the empirically relevant range of forgetting rates δ below 0.1. In light of Proposition 1, this strongly indicates that the presence of organizational forgetting rather than the presence of the outside good is responsible for multiple equilibria.

Given our baseline parameterization, the share of the outside good is quite small in general. In the static Nash equilibrium, as the marginal cost of production declines, the share of the outside good declines from 0.63 in state $(1, 1)$ over 0.33 in state $(2, 2)$ and 0.15 in state $(4, 4)$ to 0.03 in state $(15, 15)$. The share of the outside good in these states increases to 0.97, 0.87, 0.67, and 0.30, respectively, as the attractiveness of the outside good increases from $v_0 - c_0 = 0$ to $v_0 - c_0 = 3$. For sufficiently low values of δ, the increase in $v_0 - c_0$ deepens the wells and trenches in the policy function and this more aggressive behavior, in turn, leads to more pronounced asymmetries both in the short-run and in the long-run. For sufficiently high values of δ, however, it leads to less pronounced asymmetries. The reason is that a more attractive outside good *de facto* makes the market smaller, so that it is easier for organizational forgetting to stifle investment in learning-by-doing altogether. Steve { For $v_0 - c_0 = 3$ and $\rho \in \{0.85, 0.15\}$, what is the largest value of δ such that the mode of the ergodic distribution is not $(1, 1)$? ***** Uli { Compare to Table 1. ***** Finally, multiple equilibria arise for lower forgetting rates: ***** This needs to be checked! ***** Because the share of the outside good is large, the inflow of know-how into the industry is much less than one unit per period. Compared to the baseline parameterization, the inflow balances the outflow for lower forgetting rates, thereby setting the stage for multiple equilibria.

9.2 Product differentiation

Our baseline parameterization gives rise to a moderate degree of horizontal product differentiation. In the static Nash equilibrium the own-price elasticity of demand ranges between -10.27 in state $(1, 15)$ and -2.52 in state $(15, 1)$ and the cross-price elasticity of firm 1's

demand with respect to firm 2's price is 0.80 in state $(15, 1)$ and 6.28 in state $(1, 15)$. As σ is decreased from 1 to 0.1, the respective elasticities become -100.38, -2.23, 0.63, and 93.72 and they become -5.40, -2.30, 0.73, and 2.38 as σ is increased to 2. Weaker differentiation implies more asymmetric industry structures. Each firm is tempted to engage in aggressive pricing behavior since undercutting its rival by even a tiny amount dramatically increases its probability of winning the sale. Conversely, stronger differentiation implies more symmetric industry structures. ***** Uli { Maybe cite some Herfs here. *****

9.3 Choke price

In our computations the equilibria in case of $\delta = 0$ have always been flat either without or with well. In neither category of equilibrium does the leader defend its competitive advantage. One might wonder whether this is an artifact of the lack of a choke price with logit demand. As in Cabral & Riordan (1994), a firm always has a positive probability of making a sale, so that both firms must eventually reach the bottom of their learning curves. With a choke price, in contrast, the leader has the opportunity to prevent the follower from making a sale. But if the leader became much more aggressive in defending its competitive advantage, then the equilibrium may be trenchy instead of flat.

To address this concern, we assume that the probability that firm n makes a sale is given by the linear specification

$$D_1(\mathsf{p}) = \min\left(\max\left(\frac{1}{2} - \gamma(p_1 - p_2), 0\right), 1\right), \quad D_2(\mathsf{p}) = 1 - D_1(\mathsf{p}),$$

where we abstract from an outside good in the interest of simplicity and $\gamma > 0$ parameterizes the degree of horizontal product differentiation. To allow for a fair comparison between linear and logit demand, we set $v_0 - c_0 = -\infty$ and let γ be a function of σ such that in the static Nash equilibrium the own-price elasticity of demand in state $(1, 1)$ is the same in the two specifications.

There is little change in the value and policy functions. ***** This needs to be checked! ***** Flat equilibria continue to arise with linear demand. In particular, in case of $\delta = 0$ we computed the equilibrium for a variety of progress ratios $\rho \in \{0.95, 0.85, 0.75, 0.65, 0.55, 0.35, 0.15,$ and degrees of horizontal product differentiation $\sigma \in \{0.1, 1, 2\}$. The equilibrium has always been flat either without or with well. ***** This needs to be checked! ***** Based on this, we are confident that flat equilibria are not an artifact of the lack of a choke price with logit demand.

9.4 Frequency of sales

Following Cabral & Riordan (1994) we take a period to be just long enough for at most one firm to make a sale. One might wonder whether the insights of the analysis are sensitive to this assumption. To explore this issue without fundamentally departing from our modeling framework, we divide a period into $K > 1$ subperiods. Assuming that at most one sale can occur in a subperiod, at most K sales can occur in a period. If r is the discount rate per period, then $\frac{r}{K}$ is the discount rate per subperiod and $\beta = \frac{1}{1+\frac{r}{K}}$ the discount factor. We are careful not to change the properties of learning-by-doing and organizational forgetting by changing the frequency of sales. For example, we want the reduction in marginal cost that is achievable by a period's worth of sales in the original specification to be comparable to the reduction that is achievable by K subperiods' worth of sales in the alternative specification. To accomplish this, we take the state space to be $\{1, \ldots, K(M-1)+1\}^2$. The marginal cost and probability of forgetting of firm n in the alternative specification are given by $c\left(\frac{e_n-1}{K}+1\right)$ and $\delta\left(\frac{e_n-1}{K}+1\right)$. Finally, we take $K(m-1)+1$ to be the stock of know-how at which a firm reaches the bottom of its learning curve.

We have recomputed the equilibria in our four typical cases whilst doubling the frequency of sales. The equilibria for $K = 2$ closely resemble those for $K = 1$. For example, we obtain both a flat equilibrium with well and a trenchy equilibrium for $\delta = 0.03$. ***** This needs to be checked! ***** ***** Uli { Maybe cite some distance measures here. ***** Overall, it appears that our results are not sensitive to the assumption that at most one sale can occur in a period.

9.5 Organizational forgetting

We take the probability $\delta(e_n)$ that firm n loses a unit of know-how through organizational forgetting to be increasing and concave in e_n (as long as $\delta > 0$) to capture the idea that more know-how means more dispersed know-how. This appears to describe the way in which know-how accumulates in management consulting. As a firm handles a wider set of engagements, its capabilities are typically embedded in a larger number of employees. Hence, for a given rate of labor turnover, a firm with more know-how is more vulnerable to organizational forgetting than a firm with less know-how.

An alternative specification takes $\delta(e_n)$ to be constant. This may be appropriate in situations in which there is a leading edge of know-how which, if not continually applied, is at risk of being forgotten. Suppose that through trial and error, a chemical firm discovers a way to economize on raw materials in the production of a particular compound. This trick is internalized by repeating it over and over again. In this example, the probability that the firm loses a unit of know-how through organizational forgetting is likely to be independent

of its stock of know-how, i.e., $\delta(e_n) = \delta$.

Our results carry over to the alternative specification. ***** This needs to be checked! ***** If organizational forgetting is sufficiently weak, then asymmetries may arise but they cannot persist. Moreover, if asymmetries arise in the short run, they are modest. If organizational forgetting is sufficiently strong, then asymmetries cannot arise in the first place. By contrast, for intermediate degrees of organizational forgetting, asymmetries arise and persist. Even extreme asymmetries akin to near-monopoly are possible. ***** Uli { Maybe add some ¯gures here. ***** While the equilibria in the alternative specification continue to fall into our four categories, the main difference to the original specification is that multiplicity no longer arises for forgetting rates below 0.1. Instead, multiplicity prevails for forgetting rates around 0.5. ***** This needs to be checked! ***** This reaffirms our notion that the primitives of the model tie down the equilibrium unless the inflow of know-how into the industry balances the outflow. This happens for forgetting rates around 0.5, and the nature of the equilibrium is therefore governed by firms' expectations regarding to value of continued play.

9.6 State-to-state transitions

Organizational forgetting plays two roles in our model. First, a firm that makes a sale in the current period may not gain know-how and fail to move down its learning curve. Second, a firm that is not making a sale may lose know-how and slide back up its learning curve. In principle, the two roles can be separated. First, suppose that a firm gains know-how for sure whenever it makes a sale. Now a firm makes itself immune to losses from organizational forgetting by winning a sale, thus enhancing the advantage-building motive. Second, suppose that a firm loses know-how for sure whenever it does not make a sale. By winning a sale, a firm now makes its rival susceptible to losses from organizational forgetting, thus enhancing the advantage-defending motive. Both alternative specifications therefore induce more aggressive behavior and deepen the wells and trenches in the policy function. Our computations readily confirm this intuition. ***** This needs to be checked! *****

9.7 Entry and exit

So far we have assumed that the industry is composed of two firms at all times. It is straightforward to allow for entry and exit. The Online Appendix formally derives the general model; here, we briefly sketch it.

We assume that at any point in time there is a total of N firms, each of which can be either an incumbent firm or a potential entrant. Thus, if N^* is the number of incumbent

firms, $N - N^*$ is the number of potential entrants. Once an incumbent firm exits the industry, it perishes and a potential entrant automatically takes its "slot" and has to decide whether or not to enter the industry. Potential entrants are drawn from a large pool. Hence, if a potential entrant chooses not to enter the industry in the current period, it disappears and its slot is given to another potential entrant in the subsequent period. In what follows we focus on the case of $N = 2$. Though alternatives are possible, we specify that an entrant comes into the industry at the top of the learning curve.

To ensure the existence of an equilibrium, we use the approach in Doraszelski & Satterthwaite (2003). In each period, each potential entrant receives a privately observed draw S_n from a uniform distribution of possible set-up costs with support $[3, 6]$ and each incumbent firm receives a privately observed draw X_n from a uniform distribution of possible salvage values with support $[0, 3]$.[23] It is convenient to summarize the entry and exit decisions of firm n using an operating probability $\lambda_n(\mathbf{e})$, where $\mathbf{e} \in \{0, 1, \ldots, M\}^N$ is the state of the industry. If $e_n = 0$, firm n is a potential entrant and $\lambda_n(\mathbf{e})$ is the probability that it enters the industry in state $\mathbf{e}$; if $e_n \neq 0$, firm n is an incumbent firm and $\lambda_n(\mathbf{e})$ is the probability that it remains in the industry. A symmetric and anonymous Markov perfect equilibrium consists of a value function $V^*(\mathbf{e}) = V_1(\mathbf{e})$, a pricing function $p^*(\mathbf{e}) = p_1(\mathbf{e})$, and an operating probability $\lambda^*(\mathbf{e}) = \lambda_1(\mathbf{e})$ for firm 1.

Table 3 presents our summary measures of industry concentration for a progress ratio of $\rho = 0.85$ and a variety of forgetting rates $\delta \in \{0, 0.01, \ldots, 0.1, 0.3, \ldots, 0.9\}$. As a point of comparison, it also shows the limiting expected Herfindahl index H^∞ and the maximum expected Herfindahl index $\widehat{H}$ in the basic model with neither entry nor exit. To ease the computational burden, we restrict attention to equilibria that can be computed using the Pakes & McGuire (1994) algorithm.

Entry and exit do not alter the thrust of our results: organizational forgetting remains a source of aggressive pricing behavior. Indeed, allowing for exit adds another component to the prize from winning a sale because by winning a sale, a firm may move the industry to a state in which its rival is likely to exit. But if the rival exits, then it may be replaced by an entrant that comes into the industry at the top of its learning curve or it may not be replaced at all. As a result behavior is more aggressive than in the basic model with neither entry nor exit. This deepens the wells and trenches in the policy function and leads to more pronounced asymmetries both in the short-run and in the long-run as can be seen in Table 3. For intermediate degrees of organizational forgetting, it is even possible that the industry is monopolized.

[23] This implies that some portion of set-up costs is sunk, thereby eliminating the possibility that a ¯rm enters the industry merely because it hopes to draw a salvage value that exceeds its set-up cost.

	no entry and exit		entry and exit	
δ	$\widehat{H}$	H^{∞}	$\widehat{H}$	H^{∞}
0	0.527	0.500	0.527	0.500
0.01	0.547	0.500	0.547	0.500
0.02	0.600	0.500	0.600	0.500
0.03	0.736, 0.796	0.500, 0.516	0.765, 0.945	0.500, 1.000
0.04	0.825, 0.823	0.502, 0.513	0.857, 0.999	0.502, 1.000
0.05	0.847	0.530	1.000	1.000
0.06	0.870	0.559	1.000	1.000
0.07	0.881	0.744	1.000	1.000
0.08	0.882	0.770	1.000	1.000
0.09	0.882	0.787	1.000	1.000
0.10	0.881	0.803	1.000	1.000
0.30	0.860	0.860	0.997	0.997
0.50	0.597	0.597	0.659	0.597
0.70	0.539	0.539	0.690	0.539
0.90	0.511	0.511	0.728	0.511

Table 3: Limiting expected Herfindahl index H^{∞} and maximum expected Herfindahl index $\widehat{H}$. TAB51

10 Conclusions {CONCLUSIONS

Learning-by-doing and organizational forgetting have been shown to be important in a variety of industrial settings. Using the Markov-perfect equilibrium framework of Ericson & Pakes (1995) this paper provides a general model of dynamic competition that accounts for these economic fundamentals and shows how they shape industry structure and dynamics. We enhance the methodological foundations of this literature in two ways. First, we show that there are equilibria that cannot be computed using the Pakes & McGuire (1994) algorithm. Second, we propose a homotopy algorithm that allows to describe in detail the structure of the set of equilibria of our dynamic stochastic game.

In contrast to the present paper, the existing literature on learning-by-doing has largely ignored organizational forgetting. Moreover, it has mainly focused on firms' pricing behavior and has fallen short of directly examining the impact of learning-by-doing on industry dynamics. By directly examining industry dynamics, we are able to show that ID and IID (Cabral & Riordan 1994, Athey & Schmutzler 2001) may not be sufficient for economically meaningful market dominance. By generalizing the existing models of learning-by-doing through the addition of organizational forgetting, we are able to show that these dominance properties of firms' pricing behavior break down in the presence of even a small degree of organizational forgetting.

A careful analysis of the role of organizational forgetting reveals that learning-by-doing and organizational forgetting are distinct economic forces. Organizational forgetting, in particular, does not simply negate learning-by-doing. The unique role played by organizational forgetting comes about because it makes bidirectional movements through the state space possible. Hence, dynamic competition with learning-by-doing and organizational forgetting is akin to racing down an upward moving escalator. As a consequence, a model with both learning-by-doing and organizational forgetting can give rise to aggressive pricing behavior, market dominance, and multiple equilibria, whereas a model with learning-by-doing alone cannot.

Appendix

Proof of Proposition 1. To simplify the notation, we express the Bellman equations and FOCs in state $\mathbf{e}$ as

$$V_1 = D_1(p_1, p_2)\,(p_1 - c(e_1)) + \beta \sum_{k=0}^{2} D_k(p_1, p_2)\overline{V}_{1k}, \tag{13}$$ {VALUE1}

$$V_2 = D_2(p_1, p_2)\,(p_2 - c(e_2)) + \beta \sum_{k=0}^{2} D_k(p_1, p_2)\overline{V}_{2k}, \tag{14}$$ {VALUE2}

$$0 = \sigma - (1 - D_1(p_1, p_2)\,(p_1 - c(e_1)) - \beta\overline{V}_{11} + \beta \sum_{k=0}^{2} D_k(p_1, p_2)\overline{V}_{1k}, \tag{15}$$ {POLICY1}

$$0 = \sigma - (1 - D_2(p_1, p_2)\,(p_2 - c(e_2)) - \beta\overline{V}_{22} + \beta \sum_{k=0}^{2} D_k(p_1, p_2)\overline{V}_{2k}, \tag{16}$$ {POLICY2}

where V_n is shorthand for $V_n(\mathbf{e})$, p_n for $p_n(\mathbf{e})$, etc.

First suppose $\delta = 0$. The proof proceeds in a number of steps. In step 1, we establish that the equilibrium in state (M, M) is unique. In step 2a, we assume that there is a unique equilibrium in state $(e_1 + 1, M)$, where $e_1 \in \{1, \ldots, M-1\}$, and show that this implies that the equilibrium in state (e_1, M) is unique. In step 2b, we assume that there is a unique equilibrium in state $(M, e_2 + 1)$, where $e_2 \in \{1, \ldots, M-1\}$, and show that this implies that the equilibrium in state (M, e_2) is unique. By induction, steps 1, 2a, and 2b establish uniqueness along the upper edge of the state space. In step 3, we assume that there is a unique equilibrium in states $(e_1 + 1, e_2)$ and $(e_1, e_2 + 1)$, where $e_1 \in \{1, \ldots, M-1\}$ and $e_2 \in \{1, \ldots, M-1\}$, and show that this implies that in the limit as $v_0 - c_0 \to -\infty$ the equilibrium in state (e_1, e_2) is unique. Hence, uniqueness in state $(M-1, M-1)$ follows from uniqueness in states $(M, M-1)$ and $(M-1, M)$, uniqueness in state $(M-2, M-1)$ from uniqueness in states $(M-1, M-1)$ and $(M-2, M)$, etc. Working backwards gives uniqueness in states $(e_1, M-1)$, where $e_1 \in \{1, \ldots, M-1\}$. This, in turn, gives uniqueness in states $(e_1, M-2)$, where $e_1 \in \{1, \ldots, M-1\}$, etc.

Step 1: Consider state $\mathbf{e} = (M, M)$. From the definition of the state-to-state transitions

in Section 2, we have

$$\overline{V}_{10} = \overline{V}_{11} = \overline{V}_{12} = V_1, \quad \overline{V}_{20} = \overline{V}_{21} = \overline{V}_{22} = V_2.$$

Solving equations (13) and (14) for V_1 and V_2, respectively, yields

$$V_1 = \frac{D_1(p_1, p_2)(p_1 - c(e_1))}{1 - \beta}, \tag{17} \quad \{\texttt{VALUE1MM}\}$$

$$V_2 = \frac{D_2(p_1, p_2)(p_2 - c(e_2))}{1 - \beta}. \tag{18} \quad \{\texttt{VALUE2MM}\}$$

Substituting equations (17) and (18) into equations (15) and (16) and dividing through by $-(1 - D_1(p_1, p_2))$ and $-(1 - D_2(p_1, p_2))$, respectively, yields

$$0 = p_1 - c(e_1) - \frac{\sigma}{1 - D_1(p_1, p_2)} = F_1(p_1, p_2), \tag{19} \quad \{\texttt{POLICY1MM}\}$$

$$0 = p_2 - c(e_2) - \frac{\sigma}{1 - D_2(p_1, p_2)} = F_2(p_1, p_2). \tag{20} \quad \{\texttt{POLICY2MM}\}$$

The system of equations (19) and (20) determines equilibrium prices. Once we have established that there is a unique solution for p_1 and p_2, equations (17) and (18) immediately ascertain that V_1 and V_2 are unique.

Let $p_1^\natural(p_2)$ and $p_2^\natural(p_1)$ be defined by

$$F_1(p_1^\natural(p_2), p_2) = 0, \quad F_2(p_1, p_2^\natural(p_1)) = 0$$

and let $F(p_1) = p_1 - p_1^\natural(p_2^\natural(p_1))$. The p_1 that solves the system of equations (19) and (20) is the solution to $F(p_1) = 0$, and this solution is unique provided that $F(p_1)$ is strictly monotone. The implicit function theorem yields

$$F'(p_1) = 1 - \frac{\left(-\frac{\partial F_1}{\partial p_2}\right)}{\frac{\partial F_1}{\partial p_1}} \frac{\left(-\frac{\partial F_2}{\partial p_1}\right)}{\frac{\partial F_2}{\partial p_2}}.$$

Straightforward differentiation shows that

$$\frac{\left(-\frac{\partial F_1}{\partial p_2}\right)}{\frac{\partial F_1}{\partial p_1}} = \frac{D_1(p_1, p_2) D_2(p_1, p_2)}{1 - D_1(p_1, p_2)} \in (0, 1),$$

$$\frac{\left(-\frac{\partial F_2}{\partial p_1}\right)}{\frac{\partial F_2}{\partial p_2}} = \frac{D_1(p_1, p_2) D_2(p_1, p_2)}{1 - D_2(p_1, p_2)} \in (0, 1).$$

It follows that $F'(p_1) > 0$.

Step 2a: Consider state $\mathbf{e} = (e_1, M)$, where $e_1 \in \{1, \ldots, M - 1\}$. We have

$$\overline{V}_{10} = \overline{V}_{12} = V_1, \quad \overline{V}_{20} = \overline{V}_{22} = V_2.$$

Solving equations (13) and (14) for V_1 and V_2, respectively, yields

$$V_1 = \frac{D_1(p_1,p_2)(p_1 - c(e_1) + \beta\overline{V}_{11})}{1-\beta(1-D_1(p_1,p_2))}, \tag{21}$$
{VALUE1E1M}

$$V_2 = \frac{D_2(p_1,p_2)(p_2 - c(e_2)) + \beta D_1(p_1,p_2)\overline{V}_{21}}{1-\beta(1-D_1(p_1,p_2))}. \tag{22}$$
{VALUE2E1M}

Substituting equations (21) and (22) into equations (15) and (16) and dividing through by $-\frac{(1-\beta)(1-D_1(p_1,p_2))}{1-\beta(1-D_1(p_1,p_2))}$ and $-\frac{1-D_2(p_1,p_2)-\beta D_0(p_1,p_2)}{1-\beta(1-D_1(p_1,p_2))}$, respectively, yields

$$0 = p_1 - c(e_1) + \beta\overline{V}_{11} - \frac{(1-\beta(1-D_1(p_1,p_2)))\sigma}{(1-\beta)(1-D_1(p_1,p_2))} = G_1(p_1,p_2), \tag{23}$$
{POLICY1E1M}

$$0 = p_2 - c(e_2) - \frac{\beta(1-\beta)D_1(p_1,p_2)\overline{V}_{21} + (1-\beta(1-D_1(p_1,p_2)))\sigma}{1-D_2(p_1,p_2)-\beta D_0(p_1,p_2)} = G_2(p_1,p_2). \tag{24}$$
{POLICY2E1M}

David { While they look somewhat di®erent, the above equations are identical to equations (B.5a) and (B5.b) (modulo the typos) in the earlier version! The system of equations (23) and (24) determines equilibrium prices as a function of $\overline{V}_{11}$ and $\overline{V}_{21}$. These are given by $V_1(e_1+1,M)$ and $V_2(e_1+1,M)$, respectively, and are unique by hypothesis. As in step 1, it thus remains to establish that there is a unique solution for p_1 and p_2.

Proceeding as in step 1, we have to show that $G(p_1) = p_1 - p_1^\natural(p_2^\natural(p_1))$, where $p_1^\natural(p_2)$ and $p_2^\natural(p_1)$ are defined by $G_1(p_1^\natural(p_2),p_2) = 0$ and $G_2(p_1,p_2^\natural(p_1)) = 0$, respectively, is strictly monotone. Straightforward differentiation shows that

$$\frac{\left(-\frac{\partial G_1}{\partial p_2}\right)}{\frac{\partial G_1}{\partial p_1}} = \frac{D_1(p_1,p_2)D_2(p_1,p_2)}{(1-D_1(p_1,p_2))(1-\beta(1-D_1(p_1,p_2)))} \in (0,1),$$

$$\frac{\left(-\frac{\partial G_2}{\partial p_1}\right)}{\frac{\partial G_2}{\partial p_2}} = \frac{D_1(p_1,p_2)\left(\sigma(1-\beta)D_2(p_1,p_2) - \beta(1-\beta)^2 D_0(p_1,p_2)\overline{V}_{21}\right)}{\sigma(1-\beta(1-D_1(p_1,p_2)))(1-D_2(p_1,p_2)-\beta D_0(p_1,p_2))}.$$

David { The above expressions are identical to the expressions on p. 22 and p. 25 of your notes! Regarding the second expression, there are two cases two consider. (i) Suppose that $\overline{V}_{21} \geq \frac{\sigma D_2(p_1,p_2)}{\beta(1-\beta)D_0(p_1,p_2)}$. We have $\frac{-\frac{\partial G_2}{\partial p_1}}{\frac{\partial G_2}{\partial p_2}} \leq 0$. It follows that $G'(p_1) \geq 1$. (ii) Suppose that $\overline{V}_{21} < \frac{\sigma D_2(p_1,p_2)}{\beta(1-\beta)D_0(p_1,p_2)}$. Using the fact that in equilibrium $\overline{V}_{21} \geq 0$,[24] we have

$$\frac{\left(-\frac{\partial G_2}{\partial p_1}\right)}{\frac{\partial G_2}{\partial p_2}} \leq \left(\frac{D_1(p_1,p_2)}{1-D_2(p_1,p_2)-\beta D_0(p_1,p_2)}\right)\left(\frac{(1-\beta)D_2(p_1,p_2)}{1-\beta(1-D_1(p_1,p_2))}\right) \in (0,1)$$

because both terms in brackets are seen to be in the interval $(0,1)$. It follows that $G'(p_1) > 0$.

[24] To see this, note that a ¯rm is free to set price equal to marginal cost in each state in order to guarantee itself zero pro¯ts in perpetuity.

Step 2b: Consider state $\mathbf{e} = (M, e_2)$, where $e_2 \in \{1, \ldots, M-1\}$. We have

$$\overline{V}_{10} = \overline{V}_{11} = V_1, \quad \overline{V}_{20} = \overline{V}_{21} = V_2.$$

The argument is completely symmetric to the argument in step 2a and therefore omitted.

Step 3: Consider state $\mathbf{e} = (e_1, e_2)$, where $e_1 \in \{1, \ldots, M-1\}$ and $e_2 \in \{1, \ldots, M-1\}$. We have

$$\overline{V}_{10} = V_1, \quad \overline{V}_{20} = V_2.$$

Solving equations (13) and (14) for V_1 and V_2, respectively, yields

$$V_1 = \frac{D_1(p_1,p_2)(p_1 - c(e_1) + \beta \overline{V}_{11}) + \beta D_2(p_1,p_2)\overline{V}_{12}}{1 - \beta D_0(p_1,p_2)}, \qquad (25) \quad \{\texttt{VALUE1E1E2}\}$$

$$V_2 = \frac{D_2(p_1,p_2)(p_2 - c(e_2) + \beta \overline{V}_{22}) + \beta D_1(p_1,p_2)\overline{V}_{21}}{1 - \beta D_0(p_1,p_2)}. \qquad (26) \quad \{\texttt{VALUE2E1E2}\}$$

Substituting equations (25) and (26) into equations (15) and (16) and dividing through by $-\frac{(1-\beta)D_1(p_1,p_2)+\beta D_2(p_1,p_2)}{1-\beta D_0(p_1,p_2)}$ and $-\frac{\beta D_1(p_1,p_2)+(1-\beta)D_2(p_1,p_2)}{1-\beta D_0(p_1,p_2)}$, respectively, yields

$$0 = p_1 - c(e_1) + \beta \overline{V}_{11} - \frac{\beta D_2(p_1,p_2)\overline{V}_{12} + (1 - \beta D_0(p_1,p_2))\sigma}{1 - D_1(p_1,p_2) - \beta D_0(p_1,p_2)} = H_1(p_1,p_2), \qquad (27) \quad \{\texttt{POLICY1E1E2}\}$$

$$0 = p_2 - c(e_2) + \beta \overline{V}_{22} - \frac{\beta D_1(p_1,p_2)\overline{V}_{21} + (1 - \beta D_0(p_1,p_2))\sigma}{1 - D_2(p_1,p_2) - \beta D_0(p_1,p_2)} = H_2(p_1,p_2). \qquad (28) \quad \{\texttt{POLICY2E1E2}\}$$

David { While they look somewhat di®erent, the above equations are identical to equations (B.4a) and (B4.b) in the earlier version! The system of equations (27) and (28) determines equilibrium prices as a function of $\overline{V}_{11}$, $\overline{V}_{12}$, $\overline{V}_{21}$, and $\overline{V}_{22}$. These are given by $V_1(e_1+1, e_2)$, $V_1(e_1, e_2+1)$, $V_2(e_1+1, e_2)$, and $V_2(e_1, e_2+1)$, respectively, and are unique by hypothesis. As in step 1, it thus remains to establish that there is a unique solution for p_1 and p_2.

Proceeding as in step 1, we have to show that $H(p_1) = p_1 - p_1^\natural(p_2^\natural(p_1))$, where $p_1^\natural(p_2)$ and $p_2^\natural(p_1)$ are defined by $H_1(p_1^\natural(p_2), p_2) = 0$ and $H_2(p_1, p_2^\natural(p_1)) = 0$, respectively, is strictly monotone. Straightforward differentiation shows that

$$\frac{\left(-\frac{\partial H_1}{\partial p_2}\right)}{\frac{\partial H_1}{\partial p_1}} = \frac{D_2(p_1,p_2)\left(\sigma D_1(p_1,p_2) - \beta(1-\beta)D_0(p_1,p_2)\overline{V}_{12}\right)}{\sigma(1-\beta D_0(p_1,p_2))(1 - D_1(p_1,p_2) - \beta D_0(p_1,p_2))},$$

$$\frac{\left(-\frac{\partial H_2}{\partial p_1}\right)}{\frac{\partial H_2}{\partial p_2}} = \frac{D_1(p_1,p_2)\left(\sigma D_2(p_1,p_2) - \beta(1-\beta)D_0(p_1,p_2)\overline{V}_{21}\right)}{\sigma(1-\beta D_0(p_1,p_2))(1 - D_2(p_1,p_2) - \beta D_0(p_1,p_2))}.$$

David { We seem to di®er here. On p. 34 of your notes you get

$$\frac{\left(-\frac{\partial H_1}{\partial p_2}\right)}{\frac{\partial H_1}{\partial p_1}} = \frac{D_2(p_1,p_2)\left(\sigma D_1(p_1,p_2) - \beta(1-\beta)D_0(p_1,p_2)(\overline{V}_{12} - 1)\right)}{\sigma(1-\beta D_0(p_1,p_2))(1 - D_1(p_1,p_2) - \beta D_0(p_1,p_2))}$$

instead of my ¯rst expression. I tried to go over your calculations. Here is what I don't get: On p. 32 you de¯ne

$$h_1 = -\frac{\beta\eta}{1-\eta} V^*(e_1, e_2+1) + \ldots,$$

where, of course, $V^*(e_1, e_2+1) = \overline{V}_{12}$. On p. 32B you establish that

$$\frac{\partial}{\partial p_2} \frac{\eta}{1-\eta} = -\frac{\eta}{\sigma(1-\eta)}.$$

Hence, it follows that

$$\frac{\partial h_1}{\partial p_2} = \frac{\beta\eta}{\sigma(1-\eta)} V^*(e_1, e_2+1) + \ldots.$$

This last expression indeed appears on the last line of p. 33. Yet, the line above it has

$$\frac{\partial h_1}{\partial p_2} = \underbrace{\frac{\beta(1-\beta)\eta(1-\eta)}{\sigma(1-\beta(1-\eta))^2} V^*(e_1, e_2+1)}_{=z \text{ on p. 34}} + \ldots.$$

p I(d(o(e1n1Z't seeThi.75116(gov)3sseer(33.72(P.9Terhapssee)-383(y)31(o.75184(ldsee)-2double-c3(y)

The proof is completed by recalling a basic result from linear algebra: Let A be an arbitrary matrix and $\sigma(A)$ its spectrum. Then $\sigma(A+I) = \sigma(A)+1$. Hence, because $\frac{\partial F(\mathbf{x}(s),\delta(s))}{\partial \mathbf{x}}$ has at least one real nonnegative eigenvalue, it follows from equation (10) that $\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}}$ has at least one real eigenvalue equal to or bigger than unity. ■

Proof of Proposition 3. *Part (i):* Consider the static Nash equilibrium. Equations (5) and (6) imply $p_1^{\dagger}(\mathbf{e}) > c(e_1)$ and $p_2^{\dagger}(\mathbf{e}) > c(e_2)$ and thus in particular $p^{\dagger}(m,m) > c(m)$. In addition, $p^{\dagger}(\mathbf{e}) = p^{\dagger}(m,m)$ because $c(e_1) = c(e_2) = c(m)$ for all $\mathbf{e} \in \{m,\dots,M\}^2$.

Turning to our dynamic stochastic game, suppose that $\delta = 0$. The proof of part (i) proceeds in a number of steps, similar to the proof of Proposition 1. In step 1, we establish that equilibrium prices in state (M,M) coincide with the static Nash equilibrium. In step 2a, we assume that the equilibrium in state (e_1+1,M), where $e_1 \in \{m,\dots,M-1\}$, coincides with the equilibrium in state (M,M) and show that this implies that the equilibrium in state (e_1,M) does the same. In step 2b, we assume that the equilibrium in state (M,e_2+1), where $e_2 \in \{m,\dots,M-1\}$, coincides with the equilibrium in state (M,M) and show that this implies that the equilibrium in state (M,e_2) does the same. In step 3, we assume that the equilibrium in states (e_1+1,e_2) and (e_1,e_2+1), where $e_1 \in \{m,\dots,M-1\}$ and $e_2 \in \{m,\dots,M-1\}$, coincides with the equilibrium in state (M,M) and show that this implies that the equilibrium in state (e_1,e_2) does the same. Also similar to the proof of Proposition 1, we continue to use V_n as shorthand for $V_n(\mathbf{e})$, p_n for $p_n(\mathbf{e})$, etc.

Step 1: Consider state $\mathbf{e} = (M,M)$. From the proof of Proposition 1, equilibrium prices are determined by the system of equations (19) and (20). Since equations (19) and (20) are identical to equations (5) and (6), equilibrium prices are $p_1 = p_1^{\dagger}$ and $p_2 = p_2^{\dagger}$. Substituting equation (19) into (17) and equation (20) into (18) yields equilibrium values

$$V_1 = \frac{\sigma D_1(p_1,p_2)}{(1-\beta)(1-D_1(p_1,p_2))}, \tag{29}$$ {EQVALUE1MM}

$$V_2 = \frac{\sigma D_2(p_1,p_2)}{(1-\beta)(1-D_2(p_1,p_2))}. \tag{30}$$ {EQVALUE2MM}

Step 2a: Consider state $\mathbf{e} = (e_1,M)$, where $e_1 \in \{m,\dots,M-1\}$. Equilibrium prices are determined by the system of equations (23) and (24). Given $\overline{V}_{11} = V_1(e_1+1,M) = V_1(M,M)$ and $\overline{V}_{21} = V_2(e_1+1,M) = V_2(M,M)$, it is easy to see that $p_1 = p_1(M,M)$ and $p_2 = p_2(M,M)$ are a solution. Substituting equation (23) into (21) and equation (24) into (22) yields equilibrium values $V_1 = V_1(M,M)$ and $V_2 = V_2(M,M)$ as given by equations (29) and (30).

Step 2b: Consider state $\mathbf{e} = (M,e_2)$, where $e_2 \in \{m,\dots,M-1\}$. The argument is completely symmetric to the argument in step 2a and therefore omitted.

Step 3: Consider state $\mathbf{e} = (e_1,e_2)$, where $e_1 \in \{m,\dots,M-1\}$ and $e_2 \in \{m,\dots,M-1\}$. Equilibrium prices are determined by the system of equations (27) and (28). Given $\overline{V}_{11} = V_1(e_1+1,e_2) = V_1(M,M)$, $\overline{V}_{12} = V_1(e_1,e_2+1) = V_1(M,M)$, $\overline{V}_{21} = V_2(e_1+1,e_2) = V_2(M,M)$, and $\overline{V}_{22} = V_2(e_1,e_2+1) = V_2(M,M)$, it is easy to see that $p_1 = p_1(M,M)$ and $p_2 = p_2(M,M)$ are a solution. Substituting equation (27) into (25) and equation (28) into (26) yields equilibrium values $V_1 = V_1(M,M)$ and $V_2 = V_2(M,M)$ as given by equations (29) and (30).

Part (ii): We show that $p_2(\mathbf{e}) > c(m)$ for all $e_1 \in \{1,\dots,m-1\}$ and $e_2 \in \{m,\dots,M\}$.

The claim follows because $p^*(\mathsf{e}) = p_2(\mathsf{e}^{[2]})$.

The proof of part (ii) proceeds in two steps. In step 1, we establish that the equilibrium price of firm 2 in state (e_1, M), where $e_1 \in \{1, \ldots, m-1\}$, exceeds $c(m)$. In step 2, we assume that the equilibrium in state (e_1, e_2+1), where $e_1 \in \{1, \ldots, m-1\}$ and $e_2 \in \{m, \ldots, M-1\}$, coincides with the equilibrium in state (e_1, M) and show that this implies that the equilibrium in state (e_1, e_2) does the same.

Step 1: Consider state $\mathsf{e} = (e_1, M)$, where $e_1 \in \{1, \ldots, m-1\}$. From the proof of Proposition 1, equilibrium prices are determined by the system of equations (23) and (24). Using the fact that in equilibrium $\overline{V}_{21} \geq 0$, equation (24) implies $p_2 > c(m)$. Substituting equation (23) into (21) and equation (24) into (22) yields equilibrium values

$$V_1 = \frac{\sigma D_1(p_1, p_2)}{(1-\beta)(1 - D_1(p_1, p_2))}, \tag{31}$$ {EQVALUE1E1M}

$$V_2 = \frac{\sigma D_2(p_1, p_2) + \beta D_1(p_1, p_2)\overline{V}_{21}}{1 - D_2(p_1, p_2) - \beta D_0(p_1, p_2)}. \tag{32}$$ {EQVALUE2E1M}

Step 2: Consider state $\mathsf{e} = (e_1, e_2)$, where $e_1 \in \{1, \ldots, m-1\}$ and $e_2 \in \{m, \ldots, M-1\}$. Equilibrium prices are determined by the system of equations (27) and (28). Given $\overline{V}_{12} = V_1(e_1, e_2+1) = V_1(e_1, M)$ and $\overline{V}_{22} = V_2(e_1, e_2+1) = V_2(e_1, M)$, it is easy to see that $p_1 = p_1(e_1, M)$ and $p_2 = p_2(e_1, M)$ are a solution. Substituting equation (27) into (25) and equation (28) into (26) yields equilibrium values $V_1 = V_1(e_1, M)$ and $V_2 = V_2(e_1, M)$ as given by equations (31) and (32). ■

Proof of Proposition 4. To simplify the notation, we express the FOCs in state e as

$$0 = \sigma - (1 - D_1(p_1, p_2)\,(p_1 - c(e_1)) - \beta\overline{V}_{11} + \beta\sum_{k=0}^{2} D_k(p_1, p_2)\overline{V}_{1k},$$

$$0 = \sigma - (1 - D_2(p_1, p_2)\,(p_2 - c(e_2)) - \beta\overline{V}_{22} + \beta\sum_{k=0}^{2} D_k(p_1, p_2)\overline{V}_{2k},$$

where V_n is shorthand for $V_n(\mathsf{e})$, p_n for $p_n(\mathsf{e})$, etc. Dividing through by $-(1 - D_1(p_1, p_2))$ and $-(1 - D_2(p_1, p_2))$, respectively, yields

$$0 = p_1 - c(e_1) + \beta\overline{V}_{11} - \frac{\sigma + \beta\left(D_0(p_1, p_2)\overline{V}_{10} + D_2(p_1, p_2)\overline{V}_{12}\right)}{1 - D_1(p_1, p_2)} = H_1(p_1, p_2), \tag{33}$$ {AGAINPOLICY1

$$0 = p_2 - c(e_2) + \beta\overline{V}_{22} - \frac{\sigma + \beta\left(D_0(p_1, p_2)\overline{V}_{20} + D_1(p_1, p_2)\overline{V}_{21}\right)}{1 - D_2(p_1, p_2)} = H_2(p_1, p_2). \tag{34}$$ {AGAINPOLICY2

The system of equations (33) and (34) determines equilibrium prices. We have to establish that there is a unique solution for p_1 and p_2 irrespective of $\overline{V}_{10}$, $\overline{V}_{11}$, $\overline{V}_{12}$, $\overline{V}_{20}$, $\overline{V}_{21}$, and $\overline{V}_{22}$.

Proceeding as in step 3 of the proof of Proposition 1, we have to show that $H(p_1) = p_1 - p_1^\natural(p_2^\natural(p_1))$, where $p_1^\natural(p_2)$ and $p_2^\natural(p_1)$ are defined by $H_1(p_1^\natural(p_2), p_2) = 0$ and $H_2(p_1, p_2^\natural(p_1)) = 0$,

respectively, is strictly monotone. Straightforward differentiation shows that

$$\frac{\left(-\frac{\partial H_1}{\partial p_2}\right)}{\frac{\partial H_1}{\partial p_1}} = \frac{D_2(p_1,p_2)\left(\sigma D_1(p_1,p_2) + \beta D_0(p_1,p_2)\left(\overline{V}_{10} - \overline{V}_{12}\right)\right)}{\sigma(1 - D_1(p_1,p_2))},$$

$$\frac{\left(-\frac{\partial H_2}{\partial p_1}\right)}{\frac{\partial H_2}{\partial p_2}} = \frac{D_1(p_1,p_2)\left(\sigma D_2(p_1,p_2) + \beta D_0(p_1,p_2)\left(\overline{V}_{20} - \overline{V}_{21}\right)\right)}{\sigma(1 - D_2(p_1,p_2))}.$$

As $v_0 - c_0 \to -\infty$, $D_0(p_1,p_2) \to 0$ and $D_1(p_1,p_2) + D_2(p_1,p_2) \to 1$. Hence,

$$\frac{\left(-\frac{\partial H_1}{\partial p_2}\right)}{\frac{\partial H_1}{\partial p_1}} \to D_1(p_1,p_2) \in (0,1),$$

$$\frac{\left(-\frac{\partial H_2}{\partial p_1}\right)}{\frac{\partial H_2}{\partial p_2}} \to D_2(p_1,p_2) \in (0,1).$$

It follows that $H'(p_1) > 0$. David and Mark { Just as in the proof of Proposition 1 this argument begs the question whether we need conditions to ensure that the solution to the limit of the system is equal to the limit of the solution. What do you think about the following way to proceed?

As $v_0 - c_0 \to -\infty$, $D_0(p_1,p_2) \to 0$. Hence, if the outside good is su±ciently unattractive, then

$$\begin{aligned} D_0(p_1,p_2) \quad < \quad \min\Bigg\{ & \frac{\sigma\left(1 - D_1(p_1,p_2)(1 + D_2(p_1,p_2))\right)}{\beta D_2(p_1,p_2)\left|\overline{V}_{10} - \overline{V}_{12}\right|}, \frac{\sigma D_1(p_1,p_2)}{\beta\left|\overline{V}_{10} - \overline{V}_{12}\right|}, \\ & \frac{\sigma\left(1 - D_2(p_1,p_2)(1 + D_1(p_1,p_2))\right)}{\beta D_1(p_1,p_2)\left|\overline{V}_{20} - \overline{V}_{21}\right|}, \frac{\sigma D_2(p_1,p_2)}{\beta\left|\overline{V}_{20} - \overline{V}_{21}\right|} \Bigg\}. \end{aligned}$$

(To see that this is well-de¯ned, note that $D_1(p_1,p_2) \le 1 - D_2(p_1,p_2)$ implies $D_1(p_1,p_2)(1+D_2(p_1,p_2)) \le (1-D_2(p_1,p_2))(1+D_2(p_1,p_2)) < 1$ and thus $1-D_1(p_1,p_2)(1+D_2(p_1,p_2)) > 0$.) There are three cases to consider depending on the sign of $\overline{V}_{10}-\overline{V}_{12}$. (i) Suppose that $\overline{V}_{10}-\overline{V}_{12} = 0$. We have $\frac{-\frac{\partial H_1}{\partial p_2}}{\frac{\partial H_1}{\partial p_1}} = \frac{D_2(p_1,p_2)D_1(p_1,p_2)}{1-D_1(p_1,p_2)} \in (0,1)$. (ii) Suppose that $\overline{V}_{10}-\overline{V}_{12} > 0$. We have $\frac{-\frac{\partial H_1}{\partial p_2}}{\frac{\partial H_1}{\partial p_1}} > 0$. Moreover, since $D_0(p_1,p_2) < \frac{\sigma(1-D_1(p_1,p_2)(1+D_2(p_1,p_2)))}{\beta D_2(p_1,p_2)\left|\overline{V}_{10}-\overline{V}_{12}\right|}$, we have $\frac{-\frac{\partial H_1}{\partial p_2}}{\frac{\partial H_1}{\partial p_1}} < 1$. (iii) Suppose that $\overline{V}_{10} - \overline{V}_{12} < 0$. We have $\frac{-\frac{\partial H_1}{\partial p_2}}{\frac{\partial H_1}{\partial p_1}} < 1$. Moreover, since $D_0(p_1,p_2) < \frac{\sigma D_1(p_1,p_2)}{\beta\left|\overline{V}_{10}-\overline{V}_{12}\right|}$, we have $\frac{-\frac{\partial H_1}{\partial p_2}}{\frac{\partial H_1}{\partial p_1}} > 0$. Similarly, there are three cases to consider depending on the sign of $\overline{V}_{20} - \overline{V}_{21}$. The argument is completely symmetric and therefore omitted. It follows that $H'(p_1) > 0$. ■

References

Alchian, A. (1963), 'Reliability of progress curves in airframe production', *Econometrica* **31**(4), 679–693.

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

Athey, S. & Schmutzler, A. (2001), 'Investment and market dominance', *Rand Journal of Economics* **32**(1), 1–26.

Bailey, C. (1989), 'Forgetting and the learning curve: A laboratory study', *Management Science* **35**(3), 340–352.

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

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

Cabral, L. & Riordan, M. (1994), 'The learning curve, market dominance, and predatory pricing', *Econometrica* **62**(2), 1115–1140.

Cabral, L. & Riordan, M. (1997), 'The learning curve, predation, antitrust, and welfare', *Journal of Industrial Economics* **45**(2), 155–169.

Caplin, A. & Nalebuff, B. (1991), 'Aggregation and imperfect competition: On the existence of equilibrium', *Econometrica* **59**(1), 26–59.

Caves, R. & Porter, M. (1977), 'From entry barriers to mobility barriers: Conjectural decisions and contrived deterrence to new competition', *Quarterly Journal of Economics* **91**(2), 241–261.

Darr, E., Argote, L. & Epple, D. (1995), 'The acquisition, transfer, and depreciation of knowledge in service organizations: Productivity in franchises', *Management Science* **41**(11), 1750–1762.

Demichelis, S. & Germano, F. (2002), 'On (un)knots and dynamics in games', *Games and Economic Behavior* **41**, 46–60.

Doraszelski, U. & Satterthwaite, M. (2003), Foundations of Markov-perfect industry dynamics: Existence, purification, and multiplicity, Working paper, Hoover Institution, Stanford.

Dutton, J. & Thomas, A. (1984), 'Treating progress functions as a managerial opportunity', *Academy of Management Review* **9**(2), 235–247.

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

Fudenberg, D. & Tirole, J. (1983), 'Learning-by-doing and market performance', *Bell Journal of Economics* **14**(2), 522–530.

Fudenberg, D. & Tirole, J. (1991), *Game theory*, MIT Press, Cambridge.

Ghemawat, P. & Spence, M. (1985), 'Learning curve spillovers and market performance', *Quarterly Journal of Economics* **100**(Supplement), 839–852.

Herings, J. & Peeters, R. (2004), 'Stationary equilibria in stochastic games: Structure, selection, and computation', *Journal of Economic Theory* **118**, 32–60.

Irwin, D. & Klenow, P. (1994), 'Learning-by-doing spillovers in the seminconductor industry', *Journal of Political Economy* **102**(6), 1200–1227.

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

Judd, K. & Schmedders, K. (2004), A computational approach to proving uniqueness in dynamic games, Working paper, Hoover Institution, Stanford.

Lewis, T. & Yildirim, H. (2002), 'Managing dynamic competition', *American Economic Review* **92**(4), 779–797.

Lieberman, M. (1984), 'The learning curve and pricing in the chemical processing industries', *Rand Journal of Economics* **15**(2), 213–228.

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., Ostrovsky, M. & Berry, S. (2004), Simple estimators for the parameters of discrete dynamic games (with entry/exit examples), Working paper no. 10506, NBER, Cambridge.

Ross, D. (1986), 'Learning to dominate', *Journal of Industrial Economics* **34**(4), 337–353.

Schmedders, K. (1998), 'Computing equilibria in the general equilibrium model with incomplete asset markets', *Journal of Economic Dynamics and Control* **22**, 1375–1401.

Schmedders, K. (1999), 'A homotopy algorithm and an index theorem for the general equilibrium model with incomplete asset markets', *Journal of Mathematical Economics* **32**(2), 225–241.

Spence, M. (1981), 'The learning curve and competition', *Bell Journal of Economics* **12**(1), 49–70.

Thompson, P. (2001), 'How much did the Liberty shipbuilders learn? new evidence for an old case study', *Journal of Political Economy* **109**(1), 103–137.

Thompson, P. (2003), How much did the Liberty shipbuilders forget?, Working paper, Florida International University, Miami.

Thornton, R. & Thompson, P. (2001), 'Learning from experience and learning from others: An exploration of learning and spillovers in wartime shipbuilding', *American Economic Review* **91**(5), 1350–1368.

Vives, X. (1999), *Oligopoly pricing: Old ideas and new tools*, MIT Press, Cambridge.

Watson, L., Billups, S. & Morgan, A. (1987), 'HOMPACK: A suite of codes for globally convergent homotopy algorithms', *ACM Transcations on Mathematical Software* **13**, 281–310.

Wixted, M. & Ebbesen, E. (1991), 'On the form of forgetting', *Psychological Science* **2**(6), 409–415.

Zangwill, W. & Garcia, C. (1981), *Pathways to solutions, fixed points, and equilibria*, Prentice Hall, Englewood Cliffs.

Zimmerman, M. (1982), 'Learning effects and the commercialization of new energy technologies: The case of nuclear power', *Bell Journal of Economics* **13**(2), 297–310.

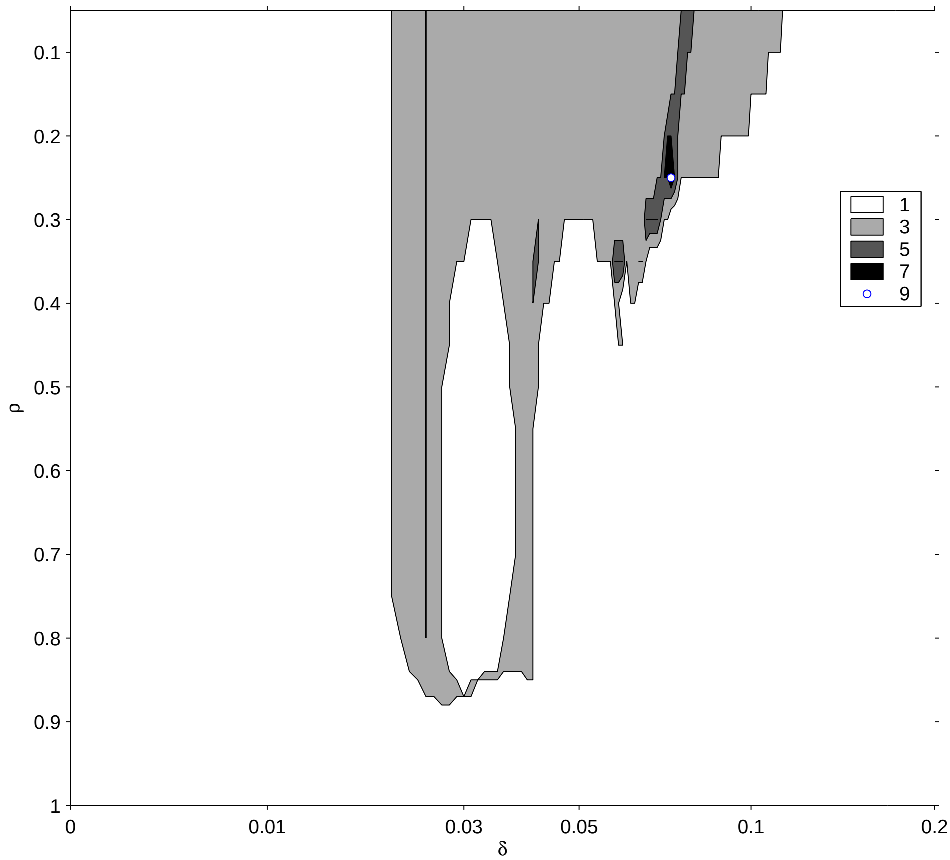

FIG1
Figure 1: Number of equilibria.

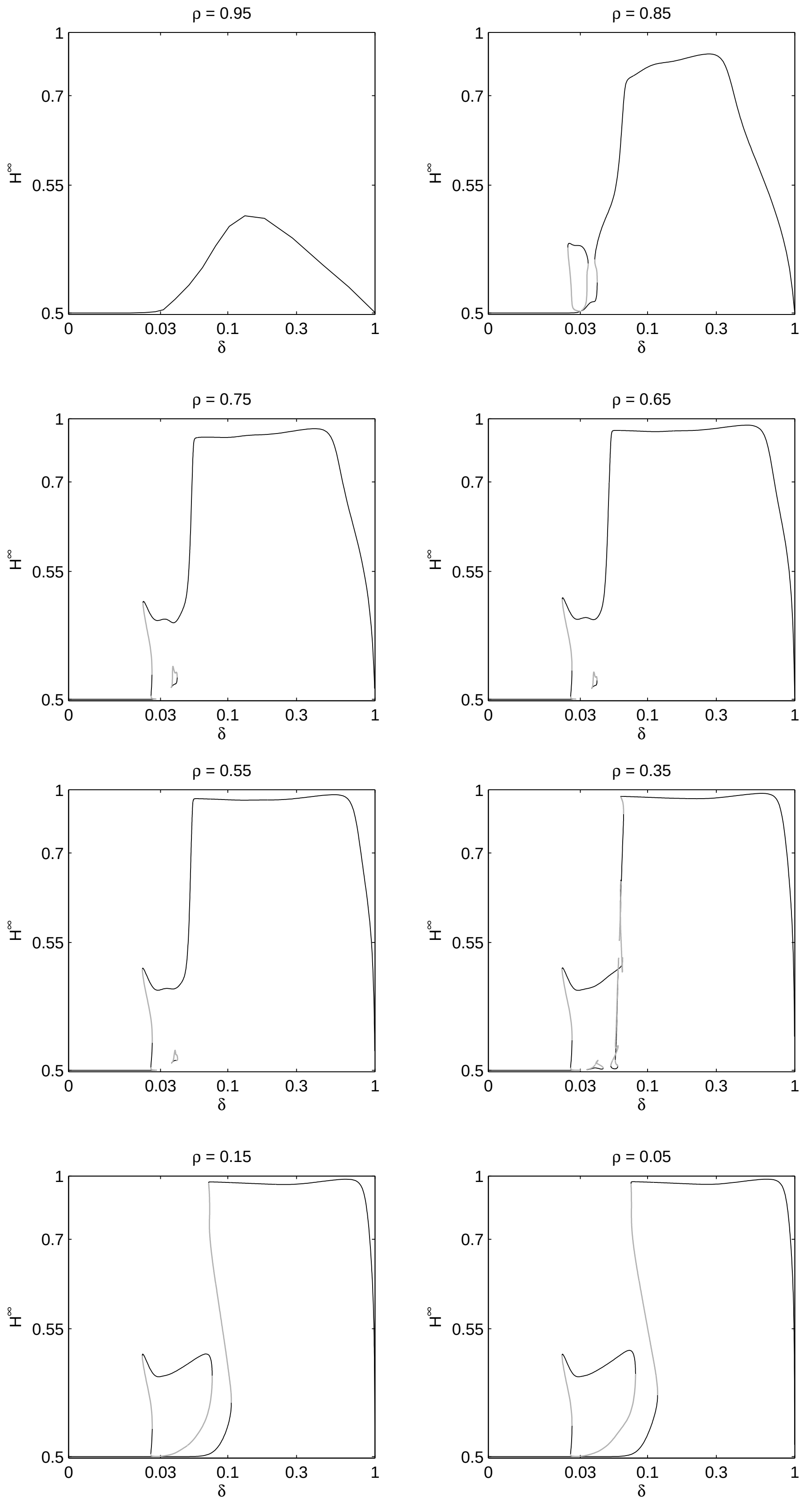

FIG2

Figure 2: Limiting expected Herfindahl index H^{∞}. Dark shades correspond to equilibria with $\rho\left(\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}}\right) < 1$, light shades to equilibria with $\rho\left(\frac{\partial G(\mathbf{x}(s))}{\partial \mathbf{x}}\right) \geq 1$.

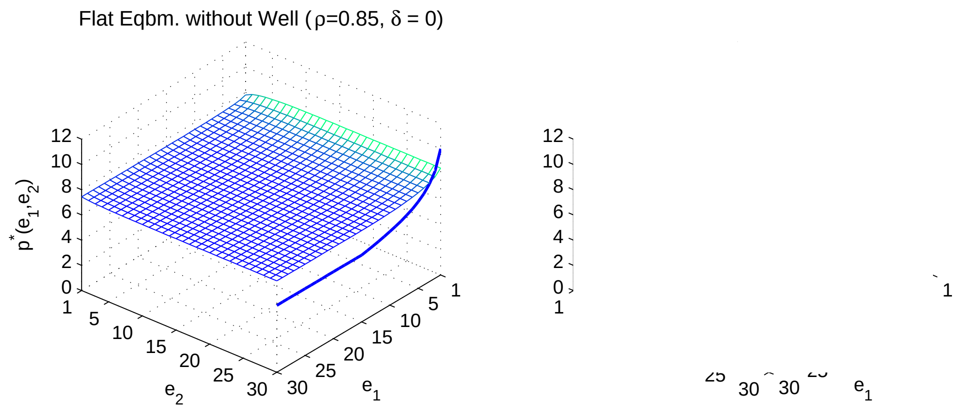

FIG3
Figure 3: Policy function $p^*(e_1, e_2)$. Marginal cost $c(e_1)$ (solid line in $e_2 = 30$-plane).

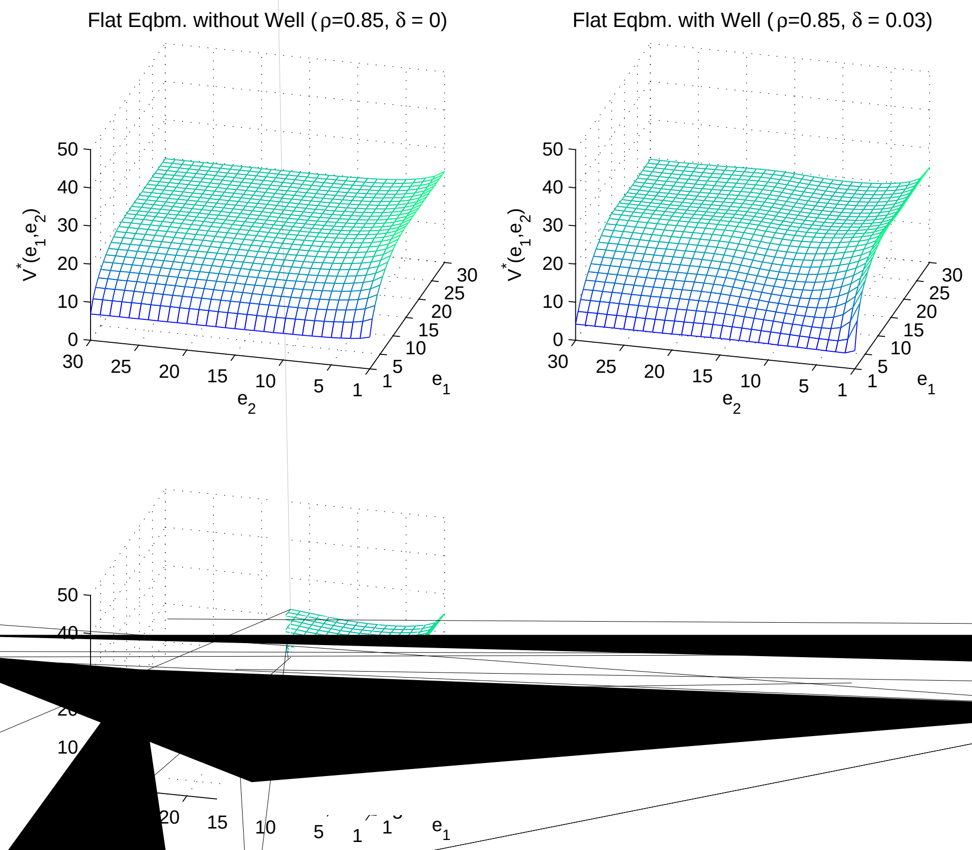

FIG4

Figure 4: Value function $V^*(e_1, e_2)$.

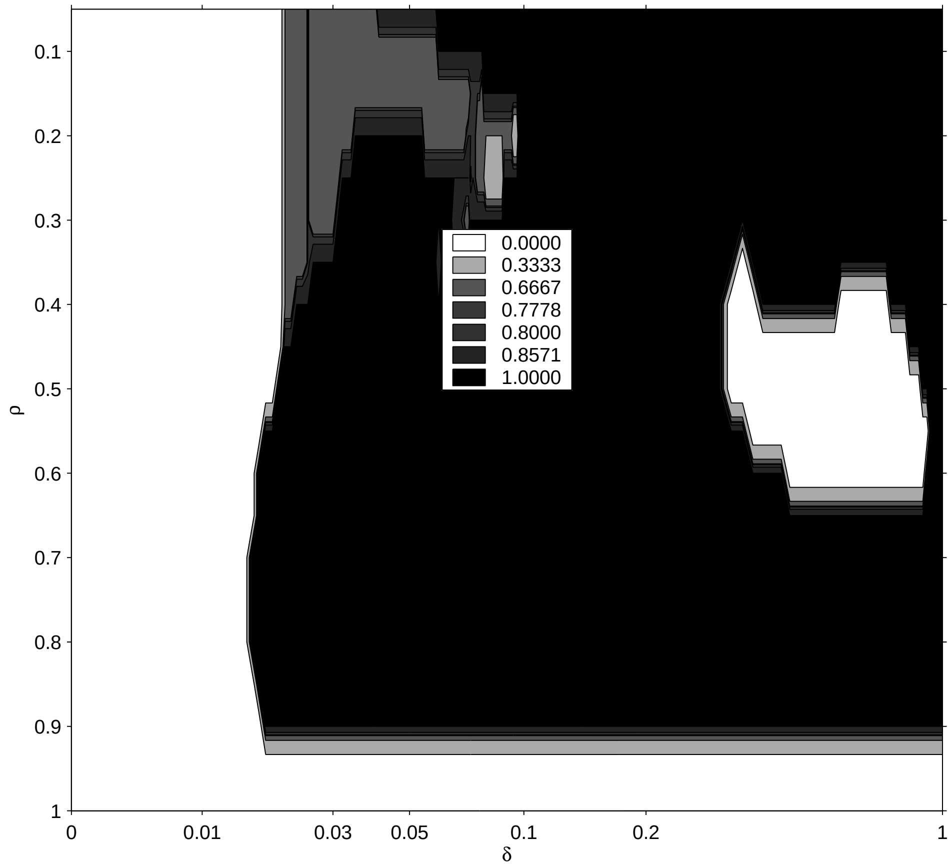

FIG11

Figure 5: Share of equilibria violating price ceiling (Result 4).

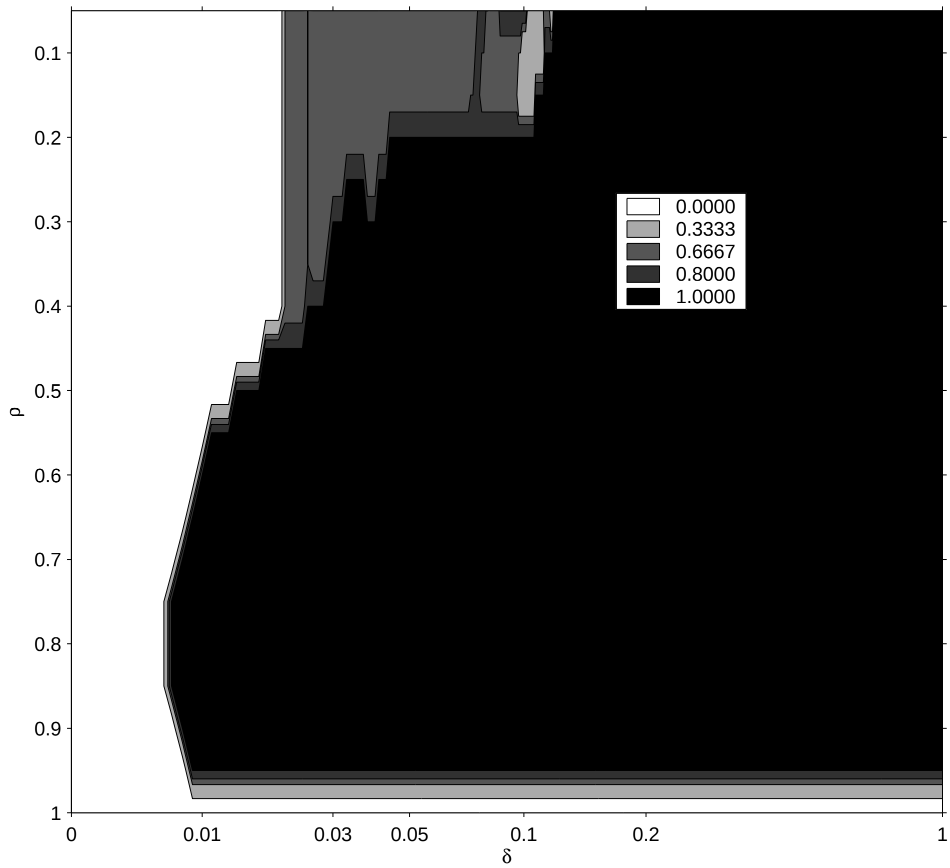

FIG12
Figure 6: Share of equilibria violating price floor (Proposition 3).

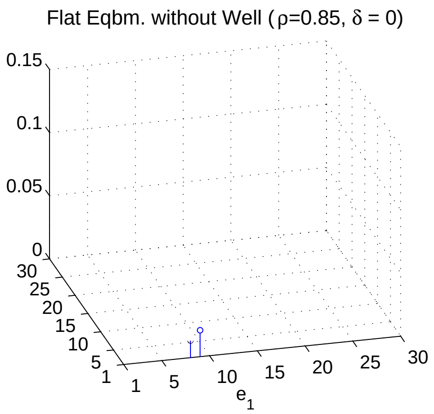

FIG31

Figure 7: Transient distribution over states in period 8 given initial state $(1, 1)$.

[redacted]

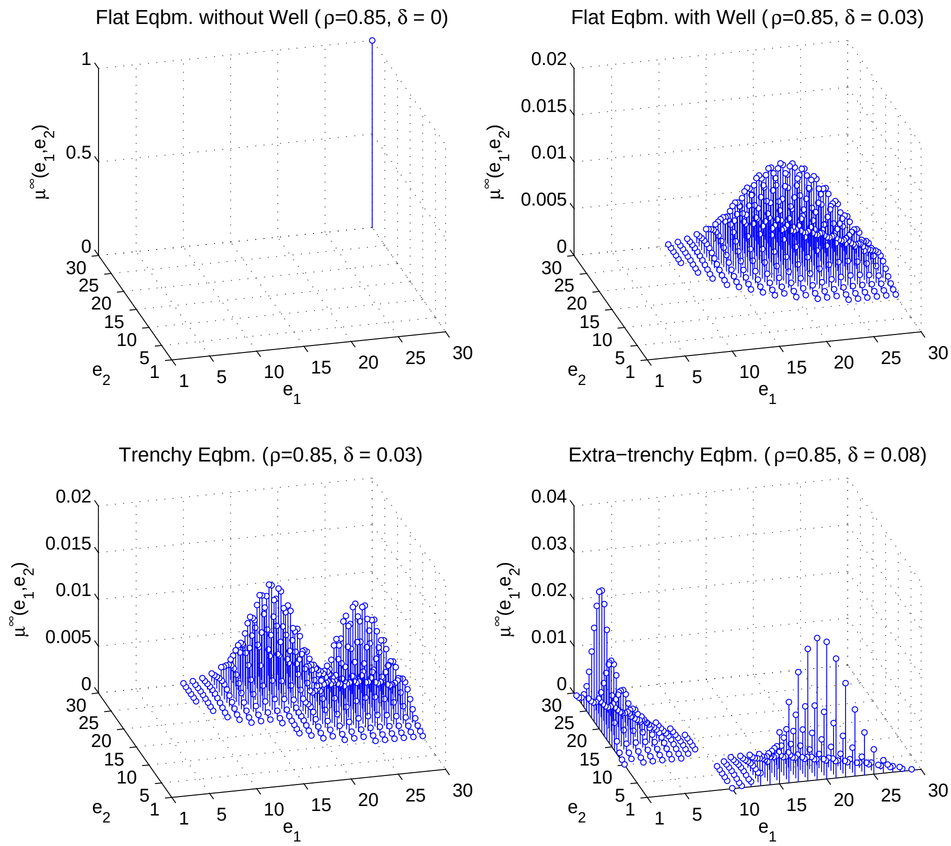

FIG33
Figure 9: Limiting distribution over states.

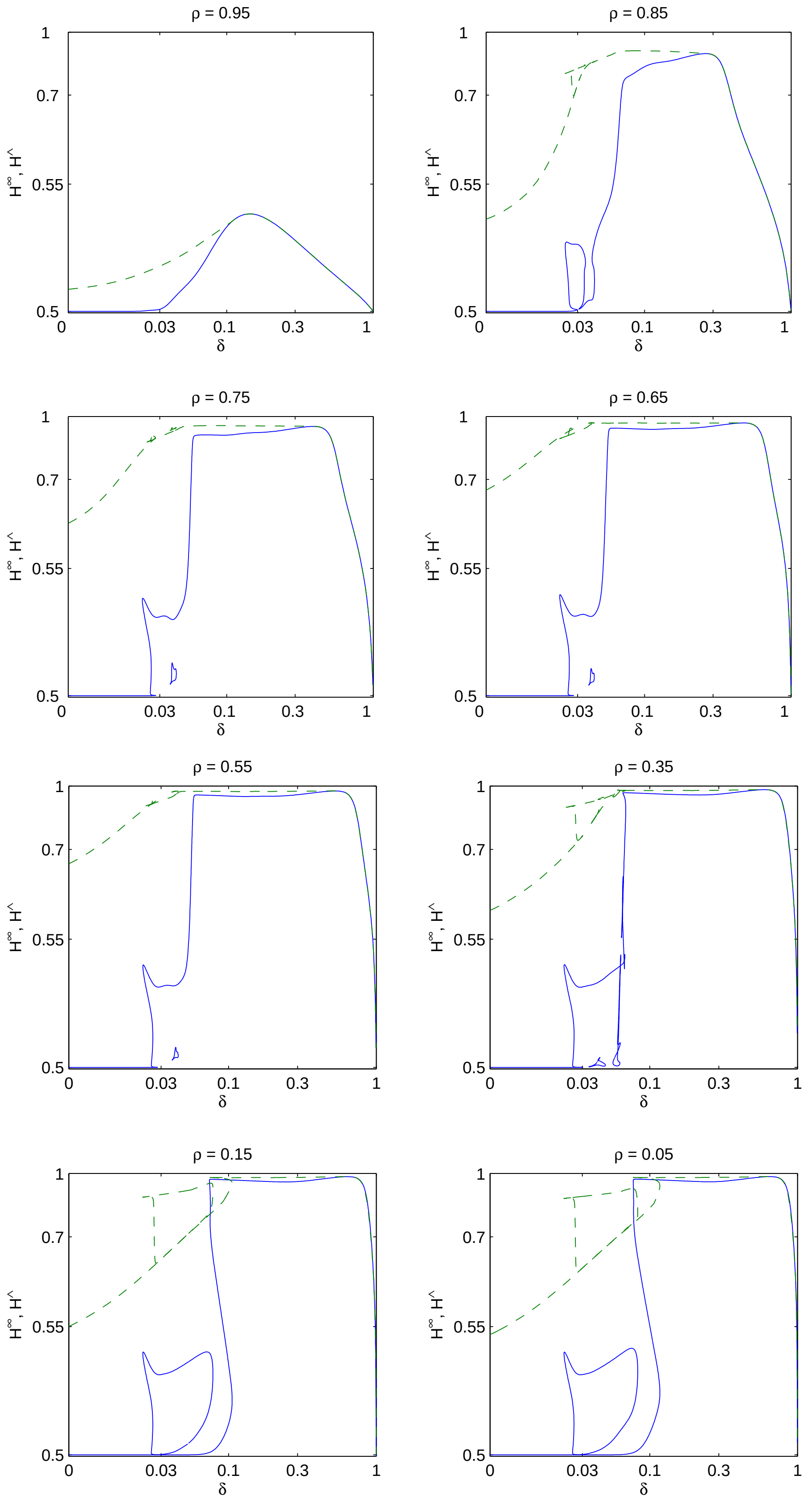

FIG35

Figure 10: Limiting expected Herfindahl index H^{∞} (solid line) and maximum expected Herfindahl index $\widehat{H}$ (dashed line).

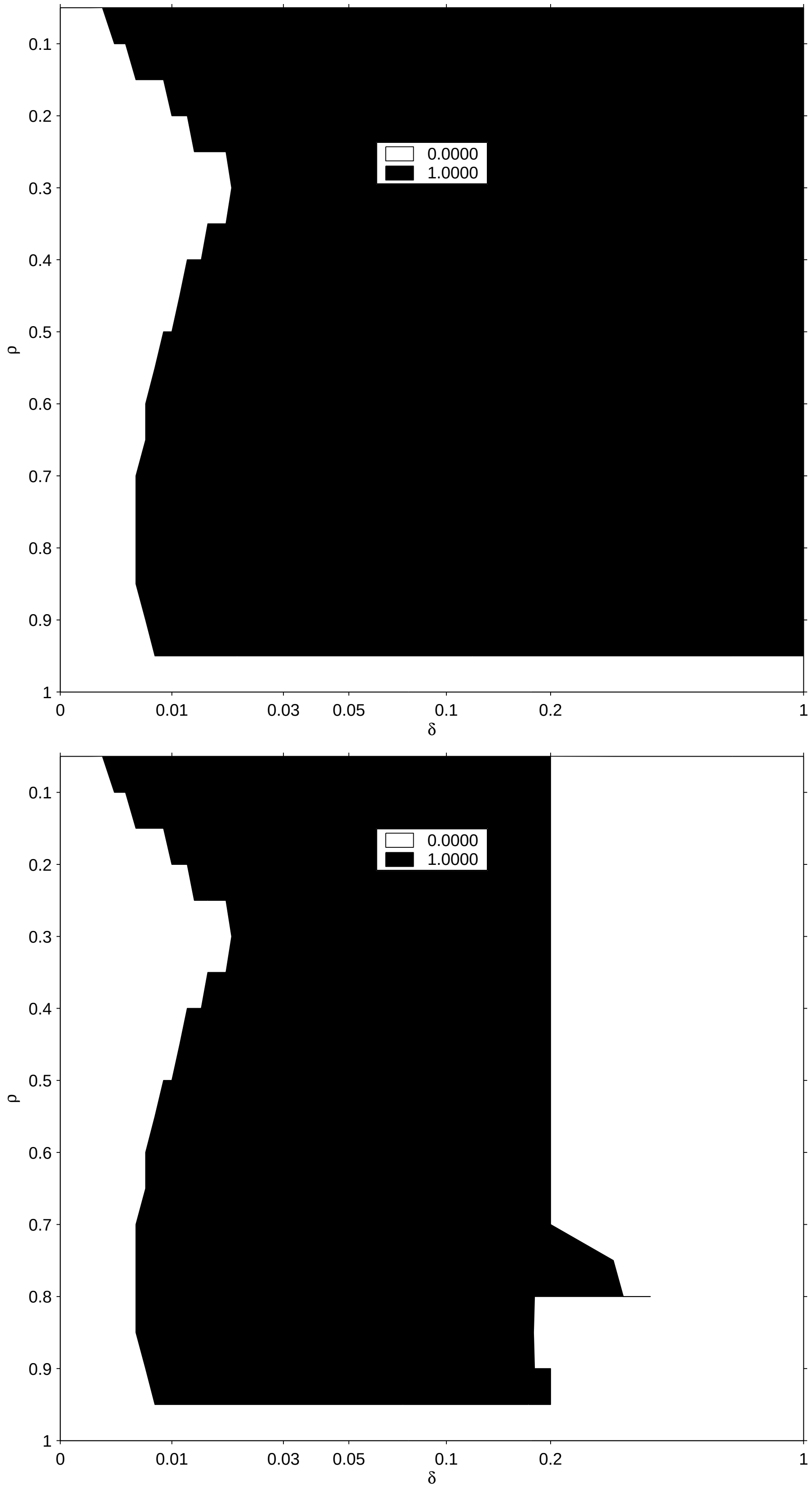

FIG36

Figure 11: Share of equilibria violating IID (part (i) of Result 7, upper panel) and share of equilibria violating ID (part (ii) of Result 7, lower panel).