11-30-2016

Is Dynamic Competition Socially Beneficial? The Case of Price as Investment

David Besanko

Ulrich Doraszelski
University of Pennsylvania

Yaroslav Kryukov

Recommended Citation

Besanko, D., Doraszelski, U., & Kryukov, Y. (2016). Is Dynamic Competition Socially Beneficial? The Case of Price as Investment. Retrieved from https://repository.upenn.edu/marketing_papers/325

This is an unpublished manuscript.

Is Dynamic Competition Socially Beneficial? The Case of Price as Investment

Abstract

We study industries where prices are not limited to their allocative and distributive roles, but also serve as an investment into lower costs or higher demand. While our model focuses on learning-by-doing and the cost advantage that it implies, our conclusions also apply to industries driven by network externalities. Existing literature does not have a clear verdict on whether the investment role of prices benefits or hurts the overall welfare, as there are a number of economic forces at work, e.g. motivation to move down the learning curve faster could be offset by the ease of driving a weaker rival out of the market. We compute both market equilibrium and first-best solution. The resulting deadweight loss appears small, in the sense that eliminating the investment motive from pricing decisions leads to much worse outcomes. Further investigation into components of deadweight loss shows that while pricing distortions are the most important driver of the deadweight loss, these distortions can be fairly small. Entry-exit distortions that arise from duplicated set-up and fixed opportunity costs also contribute to the deadweight loss, but these distortions are partially offset by more beneficial industry structure, as the market equilibrium tends to result in more active firms than the first-best solution.

Disciplines

Behavioral Economics | Business | Business Administration, Management, and Operations | Business Intelligence | Marketing | Sales and Merchandising

Comments

This is an unpublished manuscript.

Is Dynamic Competition Socially Beneficial? The Case of Price as Investment*

David Besanko[†] Ulrich Doraszelski[‡] Yaroslav Kryukov[§]

November 30, 2016
— Preliminary and Incomplete. Do not cite. —

Abstract

We study industries where prices are not limited to their allocative and distributive roles, but also serve as an investment into lower costs or higher demand. While our model focuses on learning-by-doing and the cost advantage that it implies, our conclusions also apply to industries driven by network externalities. Existing literature does not have a clear verdict on whether the investment role of prices benefits or hurts the overall welfare, as there are a number of economic forces at work, e.g. motivation to move down the learning curve faster could be offset by the ease of driving a weaker rival out of the market. We compute both market equilibrium and first-best solution. The resulting deadweight loss appears small, in the sense that eliminating the investment motive from pricing decisions leads to much worse outcomes. Further investigation into components of deadweight loss shows that while pricing distortions are the most important driver of the deadweight loss, these distortions can be fairly small. Entry-exit distortions that arise from duplicated set-up and fixed opportunity costs also contribute to the deadweight loss, but these distortions are partially offset by more beneficial industry structure, as the market equilibrium tends to result in more active firms than the first-best solution.

1 Introduction

In perfectly competitive markets, prices play two roles: allocative and distributive. Prices serve as incentive devices that shape how scarce resources are allocated within and across markets. Prices are also transfers between buyers and sellers, thus determining the distribution of consumer and producer surplus. Prices play these same roles when firms have market power, but the distributive and allocative roles are typically in tension with each other in this case.

*We thank Guy Arie, Ariel Pakes, Robert Porter, Michael Raith, Mike Riordan, and Mike Whinston for helpful discussions and suggestions. We also thank participants at seminars at Carnegie Mellon University, DG Comp at the European Commission, Helsinki Center for Economic Research, Johns Hopkins University, Toulouse School of Economics, Tulane University, and University of Rochester for their useful questions and comments.

[†]Kellogg School of Management, Northwestern University, Evanston, IL 60208, d-besanko@kellogg.northwestern.edu.

[‡]Wharton School, University of Pennsylvania, Philadelphia, PA 19104, doraszelski@wharton.upenn.edu.

[§]Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, kryukov@cmu.edu.

In many interesting settings, though, there is third role for prices: *investment*. The investment role arises when firms jostle for competitive advantage through the prices they set. Examples include competition to build cumulative experience on a learning curve or customer base in markets with network effects, switching costs, or brand loyalty. In these situations, competition is dynamic as current demand translates into lower cost or higher demand in the future, which in turn implies that current prices determine the future evolution of market structure.[1] Settings in which firms jostle for advantage through pricing are a feature of both "new economy" industries (e.g., Amazon versus Barnes and Noble in e-book readers or Sony versus Microsoft in gaming consoles) and "old economy" industries (e.g., airframes, where each new generation of aircraft entails a learning curve).

In their review of the literature on network effects and switching costs, Farrell & Klemperer (2007) point out that the investment role of prices opens up a second dimension of competition, namely competition for the market:

> *For a firm, it makes market share a valuable asset, and encourages a competitive focus on affecting expectations and on signing up pivotal (notably early) customers, which is reflected in strategies such as penetration pricing; competition is shifted from textbook competition in the market to a form of Schumpeterian competition for the market in which firms struggle for dominance.*

The central question of this paper is whether this dynamic competition is socially beneficial.

At first glance, it seems obvious that dynamic competition for advantage based on a valuable asset such as cumulative know-how or installed base, even involving just a few firms, would be extremely efficient. This point can be made by drawing a contrast with rent-seeking models. In rent-seeking models firms compete for dominant market position by engaging in socially wasteful activities (e.g., lobbying). In settings where price serves as an investment, firms also compete for dominant market position, but not through socially wasteful activities. Instead they do it by *transferring* surplus to consumers through low prices (as they engage in competition for the market) and by creating positive social value through, for example, the generation of demand-side economies of scale or learning economies. There may indeed be some distortions due to deviations from first-best pricing, but we would expect that these would be minimal in light of the investment value created by competition.

Though compelling, this intuition is incomplete. First, just because the investment role of prices leads to prices that are less than static equilibrium prices does not imply that deadweight losses (DWLs) will be lessened relative to a static setting. Pricing to gain advantage could lead to periods in which prices are less than first-best prices, resulting in a DWL from *over*production. Second, when firms use price to jostle for advantage, Besanko et al. (2014) show that firms have an *advantage-denying* motive—they choose low prices (in part) to slow the pace at which rivals build their advantages and perhaps make it more likely that they exit the industry. This motive can give rise to equilibria with predation-like behavior and long-run market structures involving monopolization by a single firm. Monopolization could, of course, lead to long-run distortions due to monopoly pricing, but it could also lead to welfare losses from insufficient product variety when consumers have a taste for variety. Third,

[1] Some models of non-price competition have a similar structure; see Besanko, Doraszelski & Kryukov (2014) for details and a list of these models.

even though competition for dominant market position when price serves as an investment does not involve socially wasteful expenditures on activities such as lobbying, it could lead to distortions in entry conduct analogous to the coordination breakdowns in natural monopoly markets highlighted in Bolton & Farrell (1990) or distortions in exit conduct due to war of attrition dynamics (Maynard Smith 1974, Tirole 1988, Bulow & Klemperer 1999).

The question of whether imperfect competition comes close to achieving first-best welfare levels when price has an investment role does not seem to have an "obvious" answer. The goal of this paper is to answer this question. We do so by analyzing a model of learning-by-doing along the lines of Cabral & Riordan (1994), Besanko, Doraszelski, Kryukov & Satterthwaite (2010), and Besanko et al. (2014) (hereafter BDK) that involves dynamic pricing competition in a differentiated product market with entry and exit. We compute Markov perfect equilibria (MPE) and compare them to the solution to a social planner's first-best (FB) problem in which the planner chooses prices and entry and exit decisions to maximize the discounted present value of total surplus. We use the solution to the FB problem to calculate the DWL associated with a market equilibrium: the difference between the discounted present value of total surplus under the FB solution and the MPE.

We show that dynamic price competition does indeed tend to lead to low DWLs. DWLs are rarely greater than 20 percent of the industry's value added, and are they less than 15 percent in 80 percent of the parameterizations for which we computed equilibria.[2] Moreover, the DWLs are also much lower than those that would arise in a counterfactual in which we "strip out" the investment role of prices. This counterfactual, which in a setting without a learning curve and product differentiation corresponds to Tirole's (1988) presentation of the war of attrition, gives rise to DWLs that are on average four times larger than under the MPE.

We then point out that the low DWLs do not arise because equilibrium conduct and market structure are similar to the conduct and structure implied by the FB solution. Indeed, for many parameterizations, equilibrium pricing and entry-exit behavior differ markedly from what the social planner would do. We also show analytically, using a "stripped-down" version of our model, that there is nothing about dynamic competition when price is an investment that precludes potentially costly entry-conduct or exit-conduct distortions. In fact, the entry-exit distortions that can arise in our model are even potentially worse than those that can arise in a rent-seeking model because not only can there be duplication of setup and fixed opportunity costs (analogous to wasteful expenditures in a rent-seeking model), but there can also be losses of surplus because too few firms enter the market in the first place.[3]

We next seek to identify the mechanisms by which dynamic competition leads to low DWLs. To do so, we decompose the DWL under dynamic competition into three components, which add up to the total DWL: the *pricing conduct distortion* that reflects deviations of MPE market shares from FB market shares in each possible state the industry could be in; the *entry-exit conduct distortion* that reflects the extra setup costs and fixed opportunity costs due to excess capacity in the industry; and the *market structure distortion*, that reflects the likelihood that the MPE ends up in less favorable states (higher costs, less product variety) than does the FB solution.

[2]Added value is the difference between total surpluses of the first-best solution and an empty industry. The reasons for this choice of scaling are explained in Section 3.2.

[3]In our model, the fixed opportunity cost is the foregone scrap value when a firm does not exit.

The low DWLs under competition boil down to three regularities involving these components that were exhibited over a wide range of parameterizations. First, pricing conduct distortions are the largest contributor to DWL, but in the "best" equilibria they are quite low. Second, the DWLs under the "worst" equilibria are "not so bad." Third, the sum of the exit-entry distortion and the market-structure distortion—which we refer to as the *non-pricing distortion*—tends to be relatively small relative to the pricing distortions in all equilibria. Each of these regularities is ultimately rooted in the presence of the learning curve. The first regularity, for example, occurs because the learning-based cost advantage that firms achieve in equilibrium marginalizes the outside (substitute) good, decreasing the price elasticity of aggregate market demand, thus minimizing the distortion from imperfectly competitive pricing. One the reasons for the third regularity (though not the only one) is that the MPE typically involve excess entry and insufficient exit, which means that the MPE is often more likely in the long run to provide consumers with greater product variety than would arise in the FB solution.[4] When learning economies are present, the enhanced variety under the MPE is especially valuable and can go a long way in offsetting the excess setup costs from overentry and the foregone scrap values from underexit.

All in all, we find that dynamic competition when price serves an investment role works remarkably well—not because competition for the market is a "magic bullet" that achieves full social efficiency, but because the components that contribute to DWL are either small or partly offset each other. And this, in turn, happens because of the learning curve itself. Learning economies, it turns out, tend to make the components of DWL under dynamic competition fairly "forgiving."

Our paper is related to a large literature of models of dynamic competition in which price serves as an investment. Besides the aforementioned models of learning-by-doing, Dasgupta & Stiglitz (1988) and Cabral & Riordan (1997) also study price competition when there is learning by doing. Price also serves as an investment in the models of network effects in Mitchell & Skrzypacz (2006), Chen, Doraszelski & Harrington (2009), Dube, Hitsch & Chintagunta (2010), and Cabral (2011); habit formation in Bergemann & Välimäki (2006); and switching costs in Dube, Hitsch & Rossi (2009) and Chen (2011). This work has generally focused on characterizing the properties of equilibria rather than anatomizing the welfare properties of competition.

Though not explicitly studying price as an investment, Segal & Whinston (2007) study a model in which two firms engage in Schumpeterian competition for the market. They investigate the impact on social welfare of antitrust policies that affect how an incumbent can behave toward an entrant during the period in which a new entrant has just entered the market with a disruptive innovation. They show that antitrust policy that protects new entrants and the expense of incumbents can have the salutary effect of increasing the overall rate of innovation. The paper thus highlights that there need not be a tension between competition for the market and competition in the market. Our paper relates to Segal & Whinston (2007) in its focus on the welfare effects of dynamic competition and on teasing out the dynamic consequences of reducing sources of static welfare losses. But unlike our paper, Segal & Whinston (2007) do not explicitly model dynamic price competition, nor do

[4] Interestingly, this latter finding contrasts with findings from static models of imperfect competition with free entry (Dixit & Stiglitz 1977, Koenker & Perry 1981, Besanko, Perry & Spady 1990, Anderson, de Palma & Thisse 1992).

they diagnose the sources of welfare losses or gains from dynamic competition.

Finally, our paper has commonalities with Pakes & McGuire's (1994) welfare analysis of MPE in a quality ladder model. Like us, they find that equilibrium welfare losses under competition can be quite low even though equilibrium conduct and market structure may differ greatly from what a social planner would choose. Unlike our analysis, however, which explores large swaths of parameter space, this result pertains to a single parameterization, so it is not clear to what extent it generalizes. In addition, in the quality ladder model price does not serve as an investment, so the mechanism driving the low DWLs in this model could be different from the mechanisms uncovered here.

The organization of the remainder of the paper is as follows. Section 2 sets up the model and characterizes the equilibrium. Section 3 analyzes the social planner's first-best problem and presents the relevant welfare metrics we use in our analysis. Section 4 provides an analytical characterization of the equilibrium in our model and the associated deadweight losses for a special case using a two-step learning curve and price inelastic demand. Sections 5 presents computations of deadweight losses over a wide range of parameter space and characterizes the regularities we observe. To put these computations in perspective, 6 compares our equilibrium deadweight losses to those that arise in a counterfactual in which the investment role of pricing is removed. Section 7 seeks to identify why deadweight losses under competition tend to be low by decomposing the DWL into the three components discussed above and analyzing the behavior of the decomposition terms in our computations. Section 8 ties together the insights from the preceding sections and offers a summary explanation of why the deadweight losses in our model are small.

Throughout the paper we distinguish between propositions that are established through formal arguments and results. A result either establishes a possibility through a numerical example or summarizes a regularity through a systematic exploration of the parameter space. Unless indicated otherwise, proofs of propositions are in the Appendix. There will also be an Online Appendix that will contain detail and background to support the analysis presented in the paper.

2 Model {Section: Mod

We study a discrete-time, infinite-horizon dynamic stochastic game between two firms in an industry that is characterized by learning-by-doing. Our model is a special case of the dynamic pricing model with endogenous competitive advantage and industry structure in Besanko et al. (2014) which, in turn, builds on the learning-by-doing models in Cabral & Riordan (1994) and Besanko et al. (2010).

At any point in time, firm $n \in \{1, 2\}$ is described by its state $e_n \in \{0, 1, \dots, M\}$. A firm can be either an incumbent firm that actively produces or a potential entrant. State $e_n = 0$ indicates a potential entrant. States $e_n \in \{1, \dots, M\}$ indicate the cumulative experience or stock of know-how of an incumbent firm. By making a sale in the current period, an incumbent firm can add to its stock of know-how and, through learning-by-doing, lower its production cost in the subsequent period. Competitive advantage and industry leadership are therefore determined endogenously in our model. The industry's state is the vector of firms' states $\mathbf{e} = (e_1, e_2)$. It completely describes the number of incumbent firms—and therefore the extent of product variety—along with their cost positions. If $e_1 > e_2$ ($e_1 < e_2$), then we refer to firm 1 as the leader (follower) and to firm 2 as the follower (leader).

In each period, firms first set prices and then decide on exit and entry. During the price-setting phase, the state changes from $\mathbf{e}$ to $\mathbf{e}'$ depending on the outcome of the pricing game between the incumbent firms. In particular, if firm 1 makes the sale and adds to its stock of know-how, the state changes to $\mathbf{e}' = \mathbf{e}^{1+} = (\min\{e_1 + 1, M\}, e_2)$; if firm 2 makes the sale, the state changes to $\mathbf{e}' = \mathbf{e}^{2+} = (e_1, \min\{e_2 + 1, M\})$.

During the exit-entry phase, the state then changes from $\mathbf{e}'$ to $\mathbf{e}''$ depending on the exit decisions of the incumbent firms and the entry decisions of the potential entrants. We model the entry of firm n as a transition from state $e_n' = 0$ to state $e_n'' = 1$ and exit as a transition from state $e_n' \geq 1$ to state $e_n'' = 0$. As the exit of an incumbent firm creates an opportunity for a potential entrant to enter the industry, re-entry is possible. The state at the end of the current period finally becomes the state at the beginning of the subsequent period.

Before analyzing firms' decisions and the equilibrium of our dynamic stochastic game, we describe the remaining primitives.

Learning-by-doing and production cost. Incumbent firm n's marginal cost of production $c(e_n)$ depends on its stock of know-how through a learning curve with a progress ratio $\rho \in [0, 1]$:

$$c(e_n) = \begin{cases} \kappa\rho^{\log_2 e_n} & \text{if} \quad 1 \leq e_n < m, \\ \kappa\rho^{\log_2 m} & \text{if} \quad m \leq e_n \leq M. \end{cases} \tag{1}$$ {curve}

Because marginal cost decreases by $100(1-\rho)\%$ as the stock of know-how doubles, a lower progress ratio implies stronger learning economies.

The marginal cost for a firm without prior experience, $c(1)$, is $\kappa > 0$. Once the firm reaches state m, the learning curve "bottoms out," and there are no further experience-based cost reductions. We accordingly refer to an industry in state $\mathbf{e}$ as a mature duopoly if $e_1 \geq m$ and $e_2 \geq m$ and as a mature monopoly if either $e_1 \geq m$ and $e_2 = 0$ or $e_1 = 0$ and $e_2 \geq m$.

Demand. The industry draws customers from a large pool of potential buyers. One buyer enters the market each period and purchases one unit of either one of the "inside goods" that are offered by the incumbent firms at prices $\mathbf{p} = (p_1, p_2)$ or an "outside good" at an exogenously given price p_0. The probability that firm n makes the sale is given by the logit specification

$$D_n(\mathbf{p}) = \frac{\exp(\frac{v-p_n}{\sigma})}{\sum_{k=0}^{2}\exp(\frac{v-p_k}{\sigma})} = \frac{\exp(\frac{-p_n}{\sigma})}{\sum_{k=0}^{2}\exp(\frac{-p_k}{\sigma})},$$

where v is gross utility and $\sigma > 0$ is a scale parameter that governs the degree of product differentiation. As $\sigma \to 0$, goods become homogeneous and the firm that sets the lowest price makes the sale for sure.[5] If firm n is a potential entrant, then we set its price to infinity so that $D_n(\mathbf{p}) = 0$.

Throughout we assume that the outside good is supplied competitively and priced at its marginal cost of production $c_0 \geq 0$. The price of the outside good $p_0 = c_0$ determines the overall level of demand for the inside goods. As it decreases, the market becomes smaller, and ultimately the industry is no longer viable.

[5] If there is more than one such firm, each of them makes the sale with equal probability.

Scrap values and setup costs. To facilitate the subsequent computations, we "purify" mixed exit and entry strategies. If incumbent firm n exits the industry, it receives a scrap value X_n drawn from a symmetric triangular distribution $F_X(\cdot)$ with support $[\overline{X}-\Delta_X, \overline{X}+\Delta_X]$, where $E_X(X_n) = \overline{X}$ and $\Delta_X > 0$ is a scale parameter. If potential entrant n enters the industry, it incurs a setup cost S_n drawn from a symmetric triangular distribution $F_S(\cdot)$ with support $[\overline{S}-\Delta_S, \overline{S}+\Delta_S]$, where $E_S(S_n) = \overline{S}$ and $\Delta_S > 0$ is a scale parameter. Scrap values and setup costs are independently and identically distributed across firms and periods, and their realization is observed by the firm but not its rival.

Although in our model a firm formally follows a pure strategy in making its exit or entry decision, the dependence of its decision on its randomly drawn, privately known scrap value, respectively, setup cost implies that its rival perceives the firm *as if* it was following a mixed strategy. As $\Delta_X \to 0$ and $\Delta_S \to 0$, the scrap value is fixed at $\overline{X}$ and the setup cost at $\overline{S}$ and we revert to mixed exit and entry strategies (Doraszelski & Satterthwaite 2010, Doraszelski & Escobar 2010).

2.1 Firms' decisions

To analyze the pricing and exit decisions of incumbent firms and the entry decisions of potential entrants, we work backwards from the exit-entry phase to the price-setting phase. Combining exit and entry decisions, we let $\phi_n(\mathbf{e}')$ denote the probability, as viewed from the perspective of its rival, that firm n decides *not* to operate in state $\mathbf{e}'$: if $e_n \neq 0$ so that firm n is an incumbent, then $\phi_n(\mathbf{e}')$ is the probability of exiting; if $e'_n = 0$ so that firm n is an entrant, then $\phi_n(\mathbf{e}')$ is the probability of not entering.

We use $V_n(\mathbf{e})$ to denote the expected net present value (NPV) of future cash flows to firm n in state $\mathbf{e}$ at the beginning of the period and $U_n(\mathbf{e}')$ to denote the expected NPV of future cash flows to firm n in state $\mathbf{e}'$ *after* pricing decisions but *before* exit and entry decisions are made. The price-setting phase determines the value function $\mathbf{V}_n$ along with the policy function $\mathbf{p}_n$ with typical element $V_n(\mathbf{e})$, respectively, $p_n(\mathbf{e})$; the exit-entry phase determines the value function $\mathbf{U}_n$ along with the policy function $\boldsymbol{\phi}_n$ with typical element $U_n(\mathbf{e}')$, respectively, $\phi_n(\mathbf{e}')$.

Exit decision of incumbent firm. To simplify the exposition, we focus on firm 1; the derivations for firm 2 are analogous. If incumbent firm 1 exits the industry, it receives the scrap value X_1 in the current period and perishes. If it does not exit, its expected NPV is

$$\widehat{X}_1(\mathbf{e}') = \beta\left[V_1(\mathbf{e}')(1-\phi_2(\mathbf{e}')) + V_1(e'_1, 0)\phi_2(\mathbf{e}')\right],$$

where $\beta \in [0,1)$ is the discount factor. The probability of incumbent firm 1 exiting the industry in state $\mathbf{e}'$ is therefore $\phi_1(\mathbf{e}') = E_X\left[1\left[X_1 \geq \widehat{X}_1(\mathbf{e}')\right]\right] = 1 - F_X(\widehat{X}_1(\mathbf{e}'))$, where $1[\cdot]$ is the indicator function and $\widehat{X}_1(\mathbf{e}')$ is the critical level of the scrap value above which exit occurs. Moreover, the expected NPV of incumbent firm 1 in the exit-entry phase is given by the Bellman equation

$$U_1(\mathbf{e}') = E_X\left[\max\left\{\widehat{X}_1(\mathbf{e}'), X_1\right\}\right]$$
$$= (1-\phi_1(\mathbf{e}'))\beta\left[V_1(\mathbf{e}')(1-\phi_2(\mathbf{e}')) + V_1(e'_1, 0)\phi_2(\mathbf{e}')\right] + \phi_1(\mathbf{e}')E_X\left[X_1 | X_1 \geq \widehat{X}_1(\mathbf{e}')\right], \quad (2) \quad \text{\{INCUMBENT VALU}$$

where $E_X\left[X_1|X_1 \geq \widehat{X}_1(\mathbf{e}')\right]$ is the expectation of the scrap value conditional on exiting the industry.

Entry decision of potential entrant. There is a large queue of potential entrants. Depending on the number of incumbent firms, up to two potential entrants can enter the industry in each period. If a potential entrant does not enter, it perishes. If it enters, it becomes an incumbent firm without prior experience in the subsequent period. Hence, upon entry, the expected NPV of potential entrant 1 is

$$\widehat{S}_1(\mathbf{e}') = \beta\left[V_1(1, e_2')(1-\phi_2(\mathbf{e}')) + V_1(1,0)\phi_2(\mathbf{e}')\right].$$

In addition, potential entrant 1 incurs the setup cost S_1 in the current period. The probability of potential entrant 1 *not* entering the industry in state $\mathbf{e}'$ is therefore $\phi_1(\mathbf{e}') = E_S\left[1\left[S_1 \geq \widehat{S}_1(\mathbf{e}')\right]\right] = 1 - F_S(\widehat{S}_1(\mathbf{e}'))$, where $\widehat{S}_1(\mathbf{e}')$ is the critical level of the setup cost below which entry occurs. Moreover, the expected NPV of potential entrant 1 in the exit-entry phase is given by the Bellman equation

$$\begin{aligned} U_1(\mathbf{e}') &= E_S\left[\max\left\{\widehat{S}_1(\mathbf{e}') - S_1, 0\right\}\right] \\ &= (1-\phi_1(\mathbf{e}'))\left\{\beta[V_1(1,e_2')(1-\phi_2(\mathbf{e}')) + V_1(1,0)\phi_2(\mathbf{e}')] - E_S\left[S_1|S_1 \leq \widehat{S}_1(\mathbf{e}')\right]\right\}, \end{aligned} \quad (3)$$

{ENTRANT VALUE

where $E_S\left[S_1|S_1 \leq \widehat{S}_1(\mathbf{e}')\right]$ is the expectation of the setup cost conditional on entering the industry.[6]

Pricing decision of incumbent firm. In the price-setting phase, the expected NPV of incumbent firm 1 is

$$\begin{aligned} V_1(\mathbf{e}) &= \max_{p_1} D_1(p_1, p_2(\mathbf{e}))(p_1 - c(e_1)) + \sum_{n=0}^{2} D_n(p_1, p_2(\mathbf{e}))U_1\left(\mathbf{e}^{n+}\right) \\ &= \max_{p_1} D_1(p_1, p_2(\mathbf{e}))(p_1 - c(e_1)) + U_1(\mathbf{e}) + \sum_{n=1}^{2} D_n(p_1, p_2(\mathbf{e}))\left[U_1\left(\mathbf{e}^{n+}\right) - U_1(\mathbf{e})\right], \end{aligned} \quad (4)$$

{BELLMAN EQUATI

where we let $\mathbf{e}^{0+} = \mathbf{e}$ and use the fact that $\sum_{n=0}^{2} D_n(\mathbf{p}) = 1$. Because the maximand on the right-hand side of Bellman equation (4) is strictly quasiconcave in p_1 (given $p_2(\mathbf{e})$ and $\mathbf{e}$), the pricing decision $p_1(\mathbf{e})$ is uniquely determined by the first-order condition

$$p_1(\mathbf{e}) - \frac{\sigma}{1 - D_1(\mathbf{p}(\mathbf{e}))} - c(e_1) + \left[U_1\left(\mathbf{e}^{1+}\right) - U_1(\mathbf{e})\right] + \Upsilon(p_2(\mathbf{e}))\left[U_1(\mathbf{e}) - U_1\left(\mathbf{e}^{2+}\right)\right] = 0, \quad (5)$$

{LBDFOC}

where $\mathbf{p}(\mathbf{e}) = (p_1(\mathbf{e}), p_2(\mathbf{e}))$ and

$$\Upsilon(p_2(\mathbf{e})) = \frac{D_2(\mathbf{p}(\mathbf{e}))}{1 - D_1(\mathbf{p}(\mathbf{e}))} = \frac{\exp\left(-\frac{p_2(\mathbf{e})}{\sigma}\right)}{\exp\left(-\frac{p_0}{\sigma}\right) + \exp\left(-\frac{p_2(\mathbf{e})}{\sigma}\right)}$$

[6] See Appendix A for closed-form expressions for $E_X\left[X_1|X_1 \geq \widehat{X}_1(\mathbf{e}')\right]$ in equation (2) and $E_S\left[S_1|S_1 \leq \widehat{S}_1(\mathbf{e}')\right]$ in equation (3).

is the probability of firm 2 making a sale conditional on firm 1 not making a sale.

As discussed in Besanko et al. (2014), the pricing decision $p_1(\mathbf{e})$ impounds two distinct goals beyond static profit $D_1(\mathbf{p}(\mathbf{e}))(p_1(\mathbf{e})-c(e_1))$: the *advantage-building motive* $U_1\left(\mathbf{e}^{1+}\right) - U_1(\mathbf{e})$ and the *advantage-denying motive* $U_1(\mathbf{e}) - U_1\left(\mathbf{e}^{2+}\right)$. The advantage-building motive is the reward that the firm receives *by winning* a sale and moving down its learning curve. The advantage-denying motive is the penalty that the firm avoids *by preventing its rival from winning* the sale and moving down its learning curve. The advantage-building and advantage-denying motives arise in a broad class of dynamic models and together capture the investment role of price.

2.2 Equilibrium and industry dynamics

Because the demand and cost specification is symmetric, we restrict ourselves to symmetric Markov perfect equilibria (MPE) in pure strategies of our learning-by-doing model. Existence follows from the arguments in Doraszelski & Satterthwaite (2010). In a symmetric equilibrium, the decisions taken by firm 2 in state $\mathbf{e}$ are identical to the decisions taken by firm 1 in state (e_2, e_1). More formally, we have $V_2(\mathbf{e}) = V_1(e_2, e_1)$, $U_2(\mathbf{e}) = U_1(e_2, e_1)$, $p_2(\mathbf{e}) = p_1(e_2, e_1)$, and $\phi_2(\mathbf{e}) = \phi_1(e_2, e_1)$. It therefore suffices to determine the value and policy functions $\mathbf{V}_1$, $\mathbf{U}_1$, $\mathbf{p}_1$, and $\boldsymbol{\phi}_1$ of firm 1.

Despite the restriction to symmetric equilibria, there is a substantial amount of multiplicity (as in Besanko et al. 2010, Besanko et al. 2014). Because the literature offers little guidance regarding equilibrium selection, we make no attempt to do so and thus view all equilibria that arise for a fixed set of primitives as equally likely.

To study the evolution of the industry under a particular equilibrium, we use the policy functions $\mathbf{p}_1$ and $\boldsymbol{\phi}_1$ to construct the matrix of state-to-state transition probabilities that characterizes the Markov process of industry dynamics. From this, we compute the transient distribution over states in period t, $\boldsymbol{\mu}_t$, starting from state $(0, 0)$ (the empty industry with an outside good but without the inside goods) in period 0.[7] The typical element $\mu_t(\mathbf{e})$ is the probability that the industry is in state $\mathbf{e}$ in period t. Depending on t, the transient distributions can capture short-run or long-run (steady-state) dynamics. We think of period 500 as the long run and, with a slight abuse of notation, denote $\boldsymbol{\mu}_{500}$ by $\boldsymbol{\mu}_{\infty}$. We use the transient distribution in period 500 rather than the limiting (or ergodic) distribution to capture long-run dynamics because the Markov process implied by the equilibrium may have multiple closed communicating classes.

[7]By starting from state $(0, 0)$ we take an *ex ante* perspective. We have in mind a setting in which two firms have developed versions of a new product that can potentially draw customers away from an established product (the outside good) but which have not yet been brought to market. This is an interesting setting in its own right: the jostle for competitive advantage by sellers of next-generation products is a pervasive feature of the business landscape, and one where the investment role of price is particularly salient. In addition, starting from state $(0, 0)$ "stacks the deck" against finding small deadweight losses by fully recognizing any distortions in the entry process (see Section 4).

3 First-best planner, welfare, and deadweight loss

{Section: Fir

3.1 First-best planner

Our welfare benchmark is a first-best planner who makes pricing, exit, and entry decisions to maximize the expected NPV of total surplus (consumer plus producer surplus).[8] In contrast to the market, the planner centralizes and coordinates decisions across firms as in Bolton & Farrell (1990). To "stack the deck" against finding small deadweight losses, we assume an omniscient planner that has access to privately known scrap values and setup costs.[9]

Combining exit and entry decisions, we let $\psi^{FB}_{1,1}(\mathbf{e}')$ denote the probability that the planner in state $\mathbf{e}'$ decides to operate both firms in the subsequent period, $\psi^{FB}_{1,0}(\mathbf{e}')$ the probability that the planner decides to operate only firm 1, $\psi^{FB}_{0,1}(\mathbf{e}')$ the probability that the planner decides to operate only firm 2, and $\psi^{FB}_{0,0}(\mathbf{e}')$ the probability that the planner decides to operate neither firm.

We use $V^{FB}(\mathbf{e})$ to denote the expected NPV of total surplus in state $\mathbf{e}$ at the beginning of the period and $U^{FB}(\mathbf{e}')$ the expected NPV of total surplus in state $\mathbf{e}'$ *after* pricing decisions but *before* exit and entry decisions are made. The price-setting phase determines the value function $\mathbf{V}^{FB}$ along with the policy functions $\mathbf{p}^{FB}_n$ for $n \in \{1,2\}$; the entry-exit phase determines the value function $\mathbf{U}^{FB}$ along with the policy functions $\boldsymbol{\psi}^{FB}_{\boldsymbol{\iota}}$ for $\boldsymbol{\iota} \in \{0,1\}^2$. We refer to $\boldsymbol{\iota} = (\iota_1, \iota_2)$ as the operating decisions of the first-best planner. Note that $\sum_{\boldsymbol{\iota}\in\{0,1\}^2} \psi^{FB}_{\boldsymbol{\iota}}(\mathbf{e}') = 1$ and that the probability that firm 1 does *not* operate in state $\mathbf{e}'$ is $\phi^{FB}_1(\mathbf{e}') = \sum_{\iota_2=0}^{1} \psi^{FB}_{0,\iota_2}(\mathbf{e}')$.

Operating decisions. Define

$$U^{FB}_{\boldsymbol{\iota}}(\mathbf{e}',\mathbf{X},\mathbf{S}) = \begin{cases} \beta V^{FB}(e_1'\iota_1, e_2'\iota_2) + X_1(1-\iota_1) + X_2(1-\iota_2) & \text{if} \quad e_1' \neq 0, e_2' \neq 0, \\ \beta V^{FB}(\iota_1, e_2'\iota_2) - S_1\iota_1 + X_2(1-\iota_2) & \text{if} \quad e_1' = 0, e_2' \neq 0, \\ \beta V^{FB}(e_1'\iota_1, \iota_2) + X_1(1-\iota_1) - S_2\iota_2 & \text{if} \quad e_1' \neq 0, e_2' = 0, \\ \beta V^{FB}(\iota_1, \iota_2) - S_1\iota_1 - S_2\iota_2 & \text{if} \quad e_1' = 0, e_2' = 0 \end{cases} \tag{6}$$

{EEPLANNER}

to be the expected NPV of total surplus in state $\mathbf{e}'$ given operating decisions $\boldsymbol{\iota} \in \{0,1\}^2$, scrap values $\mathbf{X} = (X_1, X_2)$, and setup costs $\mathbf{S} = (S_1, S_2)$. Equation (6) distinguishes between firm n actively producing in the current period ($e_n' \neq 0$) and it being inactive ($e_n' = 0$). If firm n is active, then the first-best planner receives the scrap value X_n upon deciding *not* to operate it in the subsequent period ($\iota_n = 0$); if firm n is inactive, then the planner incurs the setup cost S_n upon deciding to operate it ($\iota_n = 1$). The optimal operating decisions are

$$U^{FB}\left(\mathbf{e}',\mathbf{X},\mathbf{S}\right) = \max_{\boldsymbol{\iota}\in\{0,1\}^2} U^{FB}_{\boldsymbol{\iota}}(\mathbf{e}',\mathbf{X},\mathbf{S}),$$

with associated operating probabilities

$$\psi^{FB}_{\boldsymbol{\iota}}\left(\mathbf{e}'\right) = E_{\mathbf{X},\mathbf{S}}\left[1\left[U^{FB}\left(\mathbf{e}',\mathbf{X},\mathbf{S}\right) = U^{FB}_{\boldsymbol{\iota}}(\mathbf{e}',\mathbf{X},\mathbf{S})\right]\right] \tag{7}$$

{eq:FBopprob}

[8] Mermelstein, Nocke, Satterthwaite & Whinston (2014) consider the expected NPV of total surplus and, to a lesser extent, also the expected NPV of consumer surplus as possible objectives of an antitrust authority. We follow them in using the same discount factor for firms and the planner.

[9] As Bolton & Farrell (1990) discuss, a central authority may often have more limited information.

for $\boldsymbol{\iota} \in \{0,1\}^2$. Finally, the Bellman equation in the exit-entry phase is

$$U^{FB}(\mathbf{e}') = E_{\mathbf{X},\mathbf{S}}\left[U^{FB}\left(\mathbf{e}', \mathbf{X}, \mathbf{S}\right)\right]. \tag{8}$$ {eq:FBbell}

Pricing decisions. In the price-setting phase, the expected NPV of total surplus is

$$V^{FB}(\mathbf{e}) = \max_{\mathbf{p}} CS(\mathbf{p}) + \sum_{n=1}^{2} D_n(\mathbf{p})(p_n - c(e_n)) + \sum_{n=0}^{2} D_n(\mathbf{p}) U^{FB}\left(\mathbf{e}^{n+}\right), \tag{9}$$ {eq:FullW}

where the first term

$$CS(\mathbf{p}) = \sigma \ln\left(\sum_{n=0}^{2} \exp\left(\frac{v - p_n}{\sigma}\right)\right) = v + \sigma \ln\left(\sum_{n=0}^{2} \exp\left(\frac{-p_n}{\sigma}\right)\right) \tag{10}$$ {eq: consumer

is consumer surplus and the second term is the static profit of incumbent firms.[10] Because the outside good is priced at cost, its profit is zero.

The solution to the maximization problem on the right-hand side of Bellman equation (9) can be shown to exist and to be unique and is given by

$$p_n^{FB}(\mathbf{e}) = c(e_n) - \left[U^{FB}\left(\mathbf{e}^{n+}\right) - U^{FB}(\mathbf{e})\right]$$

for $n \in \{1,2\}$. The pricing decision $p_n^{FB}(\mathbf{e})$ reflects the marginal cost of production $c(e_n)$ of incumbent firm n net of the marginal benefit to society of moving the firm down its learning curve.

Solution and industry dynamics. The solution to the first-best planner problem exists and is unique from the contraction mapping theorem. It can be shown that the solution is symmetric in that $V^{FB}(\mathbf{e}) = V^{FB}(e_2, e_1)$, $U^{FB}(\mathbf{e}) = U^{FB}(e_2, e_1)$, $p_1^{FB}(\mathbf{e}) = p_2^{FB}(e_2, e_1)$, and $\psi_{\boldsymbol{\iota}}^{FB}(\mathbf{e}) = \psi_{\iota_2,\iota_1}^{FB}(e_2, e_1)$ for $\boldsymbol{\iota} \in \{0,1\}^2$.

We again use the policy functions to construct the matrix of state-to-state transition probabilities that characterizes the Markov process of industry dynamics and compute the transient distribution over states in period t, $\boldsymbol{\mu}_t^{FB}$, starting from state $(0,0)$ in period 0.

3.2 Welfare and deadweight loss {Section: Welf

To capture both short-run and long-run dynamics, our welfare metric is the expected NPV of total surplus. Under a particular equilibrium, total surplus in state $\mathbf{e}$ is the sum of consumer and producer surplus:

$$TS(\mathbf{e}) = CS(\mathbf{e}) + PS(\mathbf{e}),$$

where, with a slight abuse of notation, we denote $CS(\mathbf{p}(\mathbf{e}))$ by $CS(\mathbf{e})$, and $PS(\mathbf{e})$ includes the static profit $\Pi(\mathbf{e}) = \sum_{n=1}^{2} D_n(\mathbf{p}(\mathbf{e}))(p_n(\mathbf{e}) - c(e_n))$ of incumbent firms as well as their

[10] If firm n is inactive, then we again set its price to infinity so that $D_n(\mathbf{p}) = 0$ and its contribution to $CS(\mathbf{p})$ is zero.

expected scrap values and the expected setup costs of potential entrants.[11] The expected NPV of total surplus is

$$TS_\beta = \sum_{t=0}^{\infty} \beta^t \sum_{\mathbf{e}} \mu_t(\mathbf{e}) \, TS(\mathbf{e}), \tag{11}$$ {eq:defTS}

where, recall, $\mu_t(\mathbf{e})$ is the probability that the industry is in state $\mathbf{e}$ in period t.

Under the first-best planner solution, we define the expected NPV of total surplus TS_β^{FB} analogously. By construction, $TS_\beta^{FB} = V^{FB}(0,0)$. The deadweight loss arising in equilibrium is therefore the difference

$$DWL_\beta = TS_\beta^{FB} - TS_\beta. \tag{12}$$ {eq: DWL equat

Because DWL_β is measured in arbitrary monetary units, we normalize it to better gauge its size and to make it more readily comparable across parameterizations. While it seems natural to express DWL_β as a percentage of TS_β^{FB}, in our learning-by-doing model both TS_β^{FB} and TS_β vary linearly with gross utility v (because consumer surplus does; see equation (10)). Because v cancels out of DWL_β, we can therefore choose v to make DWL_β any desired percentage of TS_β^{FB}. Moreover, this does not affect the behavior of industry participants in any way.

To avoid this issue, we normalize DWL_β by the *maximum value added* by the industry:

$$VA_\beta = TS_\beta^{FB} - TS_\beta^{\varnothing},$$

where $TS_\beta^{\varnothing} = \frac{v - p_0}{1-\beta}$ is the expected NPV of total surplus if the industry remains empty forever with an outside good but without the inside goods. VA_β can be interpreted as a bound on the contribution of the inside goods to the expected NPV of total surplus. Similar to DWL_β, VA_β does not depend on v. We henceforth refer to $\frac{DWL_\beta}{VA_\beta}$ as the *relative* deadweight loss.

4 Is dynamic competition necessarily fully efficient? {Section Fully

In contrast to rent-seeking models, firms in our learning-by-doing model jostle for competitive advantage by pricing aggressively rather than by engaging in socially wasteful activities. To the extent that rents can be efficiently transferred from firms to consumers, one may thus conjecture that dynamic competition is necessarily fully efficient. This conjecture, however, overlooks that dynamic competition extends beyond pricing into exit and entry. Even if pricing is efficient, exit and entry may not be. Distortions in exit and entry can take the form of over-exit (too much or early exit), under-exit (too little or late exit), over-entry (too much or early entry), under-entry (too little or late entry), and cost-inefficient exit (lower-cost firm exits but higher-cost firm does not).

We highlight distortions in exit and entry and demonstrate that dynamic competition is not necessarily fully efficient in an analytically tractable special case of our model with a two-step learning curve, homogeneous goods, and mixed exit and entry strategies:

{ASS1}

Assumption 1 (Two-step learning curve)

[11] See Appendix A for the expression for $PS(\mathbf{e})$ and its counterpart $PS^{FB}(\mathbf{e})$ under the first-best planner solution.

1. $M = m = 2$;
2. $\sigma = 0$;
3. $\Delta_X = \Delta_S = 0$.

Because goods are homogeneous by part (2) of Assumption 1, the firm that sets the lowest price makes the sale. Moreover, aggregate demand for the inside goods is inelastic at prices below p_0. There are therefore no distortions in pricing.

We assume:

{ASS2}

Assumption 2 (Parameter restrictions)

1. $p_0 \geq \kappa$;
2. $\overline{S} > \overline{X} \geq 0$;
3. $\beta \left(p_0 - \kappa + \frac{\beta}{1-\beta} (p_0 - \rho\kappa) \right) > \overline{S}$;

By part (1) of Assumption 2, the marginal cost of the outside good $p_0 = c_0$ is at least as high as the marginal cost $c(1) = \kappa$ of an incumbent firm at the top of its learning curve. This rules out that the first-best planner opts for an empty industry. By part (2) the setup cost is positive and partially sunk and the scrap value is nonnegative. Part (3) implies that operating a single firm is socially beneficial.

The first-best planner solution is straightforward. Because goods are homogeneous and product variety is not socially beneficial, the planner operates the industry as a natural monopoly. In state $(0,0)$ in period 0, the planner decides to operate a single firm (say firm 1) in the subsequent period. In state $(1,0)$ in period 1, firm 1 charges any price below p_0, makes the sale, and moves down its learning curve. The industry remains in state $(2,0)$ in period $t \geq 2$ and firm 1 again makes the sale. The expected NPV of total surplus is thus[12]

$$TS_\beta^{FB} = v - p_0 + \beta (v - \kappa) + \frac{\beta^2}{1-\beta} (v - \kappa\rho) - \overline{S} = \frac{v - p_0}{1-\beta} + \beta \left(p_0 - \kappa + \frac{\beta}{1-\beta} (p_0 - \rho\kappa) \right) - \overline{S},$$

and the maximum value added by the industry is

$$VA_\beta = \beta \left(p_0 - \kappa + \frac{\beta}{1-\beta} (p_0 - \rho\kappa) \right) - \overline{S}.$$

{Proposition:

Proposition 1 (Two-step learning curve) *Under Assumptions 1 and 2, there exists the equilibrium shown in Table 1. The deadweight loss is*

$$DWL_\beta = \frac{\phi_1(0,0)(1-\beta)}{1-\beta\phi_1(0,0)^2} VA_\beta + \frac{(1-\phi_1(0,0))^2}{1-\beta\phi_1(0,0)^2} \left(\overline{S} - \beta\overline{X} \right) \quad (13)$$ {eq: DWL}

and the relative deadweight loss is

$$\frac{DWL_\beta}{VA_\beta} = \frac{\phi_1(0,0) - \beta\phi_1(0,0)^2}{1-\beta\phi_1(0,0)^2}. \quad (14)$$ {eq: percentag

[12] The term $v - p_0$ arises because the consumer purchases the outside good in state $(0,0)$.

Moreover, $\frac{d\left(1-\phi_1(0,0)^2\right)}{d\rho} < 0$ *and* $\frac{d(DWL_\beta/VA_\beta)}{d\rho} > 0$*: as learning economies strengthen, the probability* $1-\phi_1(0,0)^2$ *that the industry "takes off" increases and the relative deadweight loss* $\frac{DWL_\beta}{VA_\beta}$ *decreases.*

The deadweight loss arises because the entry process is decentralized and uncoordinated. The industry can therefore suffer from over-entry and under-entry. To illustrate, we sketch out the evolution of the industry in the equilibrium shown in Table 1. In state $(0,0)$ in period 0, a single firm enters the industry with probability $2(1-\phi_1(0,0))\phi_1(0,0)$; both firms enter with probability $(1-\phi_1(0,0))^2$, and no firms enter with probability $\phi_1(0,0)^2$. The industry continues to evolve as follows:

- *Case 1.* If a single firm (say firm 1) enters, then in state $(1,0)$ in period 1 it charges a price just below the price of the outside good p_0, makes the sale, and moves down its learning curve. In state $(2,0)$ firm 1 remains in the industry ($\phi_1(2,0)=0$) and firm 2 does not enter ($\phi_1(0,2)=1$). The industry remains in state $(2,0)$ in period $t \geq 2$, and firm 1 again makes the sale.

- *Case 2: Over-entry.* If both firms enter, then in state $(1,1)$ in period 1 they charge a price less than static marginal cost κ. One of the firms (say firm 1) makes the sale and moves down its learning curve. In state $(2,1)$, the leader (firm 1) remains in the industry ($\phi_1(2,1)=0$) and the follower (firm 2) exits ($\phi_1(1,2)=1$). The industry moves to—and remains in—state $(2,0)$ in period $t \geq 2$. Note that pricing in state $(1,1)$ is so aggressive that both firms incur a loss of $-\left(\frac{\beta}{1-\beta}(p_0-\rho\kappa)-\overline{X}\right)$ that fully dissipates any future gains from monopolizing the industry.

- *Case 3: Under-entry.* If no firm enters, then the above process repeats itself in state $(0,0)$ in period 1.

In short, the intuition that dynamic competition is necessarily fully efficient is incomplete. In the equilibrium shown in Table 1, while the industry evolves towards the monopolistic structure that the first-best planner operates, this may happen slowly over time due to either over-entry or under-entry.[13] Wasteful duplication and delay (Bolton & Farrell 1990) are both integral parts of the equilibrium.

The equilibrium shown in Table 1 in state $(2,2)$ entails a war of attrition (Maynard Smith 1974, Tirole 1988, Bulow & Klemperer 1999), although state $(2,2)$ is off the equilibrium path starting from state $(0,0)$. The war of attrition arises because a firm is better off staying in the industry if its rival exits but worse off if its rival stays. The resulting non-operating probability is $\phi_1(2,2) = \frac{(1-\beta)\overline{X}}{\frac{\beta}{1-\beta}(p_0-\rho\kappa)-\beta\overline{X}} \in (0,1)$. In contrast, the first-best planner ceases to operate one of the two firms in state $(2,2)$.

Proposition 1 describes one equilibrium in the two-step version of our model. Given Assumptions 1 and 2, there are two other equilibria that can arise in addition to the one in

[13]The first term in equation (13) is due to under-entry and the "discount factor" $\frac{\phi_1(0,0)(1-\beta)}{1-\beta\phi_1(0,0)^2}$ captures the stochastic length of time over which under-entry may occur; the second term is due to over-entry and the "discount factor" $\frac{(1-\phi_1(0,0))^2}{1-\beta\phi_1(0,0)^2}$ captures the stochastic length of time over which over-entry can occur after potentially many periods of under-entry.

$\mathbf{e}$	$p_1(\mathbf{e})$	$\phi_1(\mathbf{e})$	$V_1(\mathbf{e})$	$U_1(\mathbf{e})$
$(0,0)$	∞	$\frac{\overline{S}-\beta\overline{X}}{\beta\left(p_0-\kappa+\frac{\beta}{1-\beta}(p_0-\rho\kappa)\right)-\beta\overline{X}}$	–	0
$(0,1)$	∞	1	–	0
$(0,2)$	∞	1	–	0
$(1,0)$	p_0^-	0	$p_0-\kappa+\frac{\beta}{1-\beta}(p_0-\rho\kappa)$	$\beta\left(p_0-\kappa+\frac{\beta}{1-\beta}(p_0-\rho\kappa)\right)$
$(1,1)$	$\kappa-\left(\frac{\beta}{1-\beta}(p_0-\rho\kappa)-\overline{X}\right)$	$\frac{(1-\beta)\overline{X}}{\beta\left(p_0-\kappa+\frac{\beta}{1-\beta}(p_0-\rho\kappa)\right)-\beta\overline{X}}$	$\overline{X}$	$\overline{X}$
$(1,2)$	κ	1	$\overline{X}$	$\overline{X}$
$(2,0)$	p_0^-	0	$\frac{p_0-\rho\kappa}{1-\beta}$	$\frac{\beta}{1-\beta}(p_0-\rho\kappa)$
$(2,1)$	κ^-	0	$(1-\rho)\kappa+\frac{\beta}{1-\beta}(p_0-\rho\kappa)$	$\frac{\beta}{1-\beta}(p_0-\rho\kappa)$
$(2,2)$	$\rho\kappa$	$\frac{(1-\beta)\overline{X}}{\frac{\beta}{1-\beta}(p_0-\rho\kappa)-\beta\overline{X}}$	$\overline{X}$	$\overline{X}$

Table 1: Equilibrium. Two-step learning curve. In column labelled $p_1(\mathbf{e})$, superscript $-$ indicates that firm 1 charges just below the price stated.

{tbl: MPE in s

Table 1. One is the same as the one in Table 1 except that there is positive probability of the leader exiting in state $(1,0)$ and a positive probability of a new firm coming in.[14] The other is the same as Table 1's equilibrium except the leader exits with certainty and the is replaced by a new firm that enters with probability 1.[15] It is straightforward to see that if we start at state $(0,0)$, the entry-exit phase of state $(1,0)$ is never reached in these equilibria, and thus the equilibrium paths starting with an empty industry in these two additional equilibria are identical to that in the equilibrium in Table 1. Thus, these two additional equilibria give rise to the deadweight loss given by (13) and (14).

Under other parameter conditions additional equilibria can arise that create the possibility of different equilibrium paths and additional sources of deadweight loss. For example, if in addition to Assumption 2 we have $\overline{X} \geq \frac{\beta}{1-\beta}\kappa(1-\rho)$, there is an equilibrium in which both firms have a positive probability of exit in state $(2,1)$:[16]

$$\phi(1,2) = \frac{X - \beta(\kappa - \rho\kappa) - \beta X}{\beta\left(p_0 - \kappa + \frac{\beta}{1-\beta}(p_0 - \rho\kappa)\right) - \beta X} \qquad (15) \quad \{\text{eq: phi (1,2)}$$

$$\phi(2,1) = \frac{X - \beta X}{\beta\left(p_0 - \kappa + \frac{\beta}{1-\beta}(p_0 - \rho\kappa)\right) - \beta X}. \qquad (16) \quad \{\text{eq: phi(2,1)}$$

This gives rise to the possibility of over-exit (in this case, no firms as opposed to one). This equilibrium also worsens the potential welfare losses from over-entry because, in contrast to the equilibrium in Table 1, there is the possibility that two firms remain in the industry for at least an additional period. Finally, this equilibrium also creates the possibility of cost-inefficient exit in states $(2,1)$ and $(1,2)$. If this happens, there would be at least one period in which total industry production costs are higher than they would have been under the solution to the planner's problem. In fact, in this particular equilibrium, not only is there a *possibility* that the low-cost firm exits and the high cost firm does not, but the probability that the low-cost firm exits is higher than the probability that the high-cost firm exits, i.e., $\phi(2,1) > \phi(1,2)$. Thus, we have cost-inefficient exit in both an *ex post* sense and an *ex ante* sense.

The special case of a two-step learning curve relies on extreme values of key parameters. In doing so, it assumes away a meaningful role for product variety and competition from the outside good that can be a source of distortions in pricing. Unfortunately, analytic tractability rapidly declines beyond the two-step learning curve. Moreover, theoretical analysis seems ill-suited to answer the question of how efficient dynamic competition is. We therefore turn to numerical analysis.

---

[14] Specifically, in this equilibrium, the entries in Table 1 are replaced with $\phi_1(1,0) = \frac{\overline{S} - \beta\overline{X}}{\beta\left(p_0 - \kappa + \frac{\beta}{1-\beta}(p_0 - \rho\kappa)\right) - \beta\overline{X}}$,
$\phi(0,1) = \frac{(1-\beta)\overline{X}}{\beta\left(p_0 - \kappa + \frac{\beta}{1-\beta}(p_0 - \rho\kappa)\right) - \beta\overline{X}}$, and $U(1,0) = \overline{X}$. The proof that this is an equilibrium will be in the Online Appendix.

[15] Specifically, in this equilibrium, the entries in Table 1 are replaced with $\phi_1(1,0) = 1$, $\phi(0,1) = 0$, $U(1,0) = \overline{X}$, and $U(0,1) = \beta\left(p_0 - \kappa + \frac{\beta}{1-\beta}(p_0 - \rho\kappa)\right) - S$. The proof that this is an equilibrium will be in the Online Appendix.

[16] The Online Appendix will provide details and the proof. Note that part (3) of Assumption 2 can be shown to imply that $\overline{X} \leq \frac{\beta}{1-\beta}(p_0 - \rho\kappa)$, so this equilibrium arises if $\overline{X} \in \left[\frac{\beta}{1-\beta}\kappa(1-\rho), \frac{\beta}{1-\beta}(p_0 - \rho\kappa)\right]$.

parameter	value	grid
maximum stock of know-how M	30	
cost at top of learning curve κ	10	
bottom of learning curve m	15	
progress ratio ρ	0.75	$\rho \in \{0, 0.05, \ldots, 1\}$
gross utility v	10	
product differentiation σ	1	$\sigma \in \{0.2, 0.3, \ldots, 1, 1.3, 1.6, 2, 2.5, 3.2, 4, 5, 6.3, 7.9, 10\}$
price of outside good p_0	10	$p_0 \in \{0, 1, \ldots, 20\}$
scrap value $\overline{X}$, Δ_X	1.5, 1.5	$\overline{X} \in \{-1.5, -1, \ldots, 7.5\}$
setup cost $\overline{S}$, Δ_S	4.5, 1.5	
discount factor β	0.95	

Table 2: Baseline parameterization and grid points. {baseparms}

5 Numerical analysis and equilibrium {Section numer

5.1 Computation and parameterization

To thoroughly explore the equilibrium correspondence and search for multiple equilibria in a systematic fashion, we use the homotopy or path-following method in Besanko et al. (2010).[17] We caution that the homotopy algorithm cannot be guaranteed to find all equilibria and refer to reader to Besanko et al. (2010) and Borkovsky, Doraszelski & Kryukov (2010, 2012) for additional discussion. We solve the first-best planner problem using value function iteration combined with quasi-Monte Carlo integration (Halton sequences of length $10,000$) to evaluate the operating probabilities in equation (7) and the Bellman equation (8).

Our learning-by-doing model has four key parameters: the progress ratio $\rho \in [0, 1]$, the degree of product differentiation $\sigma > 0$, the price of the outside good $p_0 = c_0 \geq 0$, and the expected scrap value $\overline{X} \in [-\Delta_X, \overline{S} + \Delta_S + \Delta_X]$.[18] To explore how the equilibria vary with these parameters, we compute six two-dimensional slices through the equilibrium correspondence along (ρ, σ), (ρ, p_0), $(\rho, \overline{X})$, (σ, p_0), $(\sigma, \overline{X})$, and $(\overline{X}, p_0)$. We choose sufficiently large upper bounds for σ and p_0 so that beyond them "things don't change much anymore." Back-of-the-envelope calculations yield $\sigma \leq 10$ and $p_0 \leq 20$. Throughout we hold the remaining parameters fixed at the values in the second column of Table 2. While this baseline parameterization is not intended to be representative of any particular industry, it is neither entirely unrepresentative nor extreme.

An industry without firms is unlikely to attract the attention of a central authority. We therefore exclude extreme parameterizations for which the industry is not viable in the sense that the probability $1 - \phi_1(0, 0)^2$ that the industry "takes off" is below 0.01. Unsurprisingly, these parameterizations involve a highly attractive outside good with $p_0 < 5$.

[17] The equilibrium correspondence is $\mathbf{H}^{-1}(\boldsymbol{\omega}) = \{\mathbf{x} | \mathbf{H}(\mathbf{x}, \boldsymbol{\omega}) = 0\}$, where $\boldsymbol{\omega} = (\rho, \sigma, p_0, \overline{X}, \ldots)$ are the parameters of the model, $\mathbf{x} = (\mathbf{V}_1, \mathbf{U}_1, \mathbf{p}_1, \boldsymbol{\phi}_1)$ are the value and policy functions, and $\mathbf{H}(\mathbf{x}, \boldsymbol{\omega}) = 0$ are the Bellman equations and optimality conditions that define an equilibrium.

[18] The bounds on $\overline{X}$ follow from the economic requirement that upon exit a firm's assets are valuable ($X_n \geq 0$) but that their value is limited by the firm's initial outlay at the time of its inception ($X_n \leq S_n$).

Due to the large number of parameterizations and multiplicity of equilibria, we require a way to summarize them. In a first step, we average an outcome of interest over the equilibria at a parameterization. This random sampling is in line with our decision to refrain from equilibrium selection and ensures that parameterizations with many equilibria carry the same weight as parameterizations with few equilibria.

In a second step, we randomly sample parameterizations. To make this practical, we represent a two-dimensional slice through the equilibrium correspondence with a grid of values for the parameters spanning the slice. The third column of Table 2 lists the grid points we use for the four key parameters. We mostly use uniformly spaced grid points, except for $\sigma > 1$, where the grid points approximate a log scale in order to explore very high degrees of product differentiation. We associate each point in a two-dimensional grid with the corresponding average over equilibria. We then pool the points on the six slices through the equilibrium correspondence along (ρ, σ), (ρ, p_0), $(\rho, \overline{X})$, (σ, p_0), $(\sigma, \overline{X})$, and $(\overline{X}, p_0)$ and obtain the distribution of the outcome of interest.

5.2 Equilibrium and first-best planner solution

{Section Showca

To illustrate the types of behavior that can emerge in our learning-by-doing model, we examine the equilibria that arise at the baseline parameterization in Table 2. For two of these three equilibria Figure 1 shows the pricing decision of firm 1, the non-operating probability of firm 2, and the time path of the probability distribution over industry structures (empty, monopoly, and duopoly).[19]

The upper panels of Figure 1 exemplify what Besanko et al. (2014) call an *aggressive equilibrium*. The pricing decision in the upper left panel exhibits a deep well in state $(1, 1)$ with $p_1(1, 1) = -34.78$. A *well* is a preemption battle where firms vie to be the first to move down from the top of their learning curves. Such a battle is likely to ensue because $\phi_1(0, 0) = 0.04$ implies that the probability that both firms enter the industry in period 0 is 0.92. After the industry has emerged from the preemption battle in state (1,1), the leader (say firm 1) continues to price aggressively ($p_1(2, 1) = 0.08$). Indeed, the pricing decision exhibits a deep trench along the e_1-axis with $p_1(e_1, 1)$ ranging from 0.08 to 1.24 for $e_1 \in \{2, \dots, 30\}$.[20] A *trench* is a price war that the leader wages against the follower. We can think of a trench is an endogenous mobility barrier in the sense of Caves & Porter (1977). In the trench the follower (firm 2) exits the industry with a positive probability of $\phi_2(e_1, 1) = 0.22$ for $e_1 \in \{2, \dots, 30\}$ as the upper middle panel shows. The follower remains in this exit zone as long as it does not win a sale. Once the follower exits, the leader raises its price and the industry becomes an entrenched monopoly.[21] This sequence of events resembles conventional notions of predatory pricing.[22] The industry may also evolve into a

[19] The third equilibrium is essentially intermediate between the two shown in Figure 1.

[20] Because prices are strategic complements, there is also a shallow trench along the e_2-axis with $p_1(1, e_2)$ ranging from 3.63 to 4.90 for $e_2 \in \{2, \dots, 30\}$.

[21] While our model allows for re-entry, whether it actually occurs depends on how a potential entrant assesses its prospects in the industry. In this particular equilibrium, $\phi_2(e_1, 0) = 1.00$ for $e_1 \in \{2, \dots, 30\}$, so that the potential entrant does not enter if the incumbent firm has moved down from the top of its learning curve.

[22] Besanko et al. (2014) formalize the notion of predatory pricing in a dynamic pricing model and disentangle it from mere competition for efficiency.

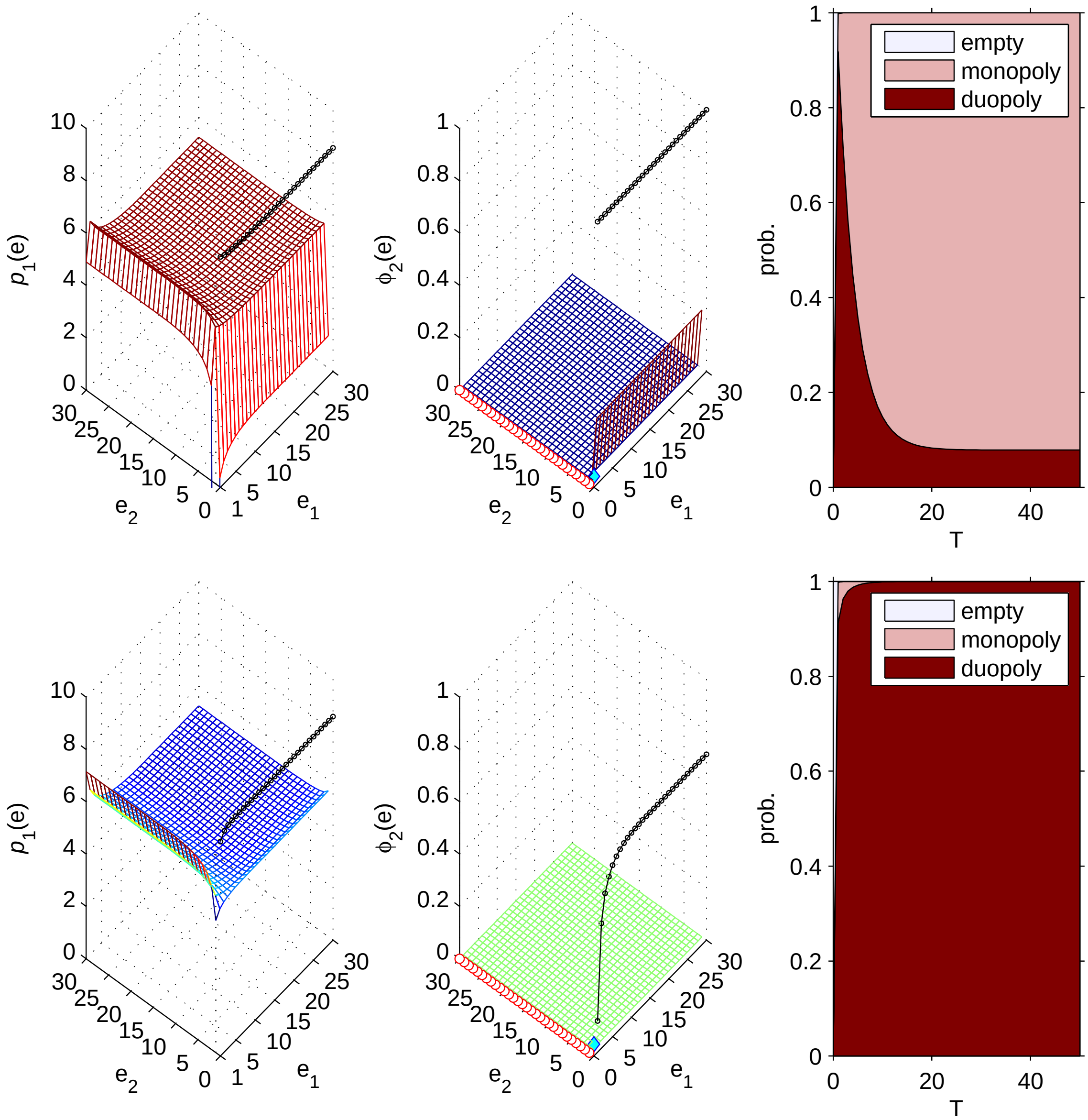

Figure 1: Aggressive (upper panels) and accommodative (lower panels) equilibrium. Pricing decision of firm 1 (left panels), non-operating probability of firm 2 (middle panels), and time path of probability distribution over industry structures (right panels). Dots above the surface in left panels are $p_1(e_1, 0)$ for $e_1 > 0$ and dots in middle panels are $\phi_2(0, e_2)$ for $e_2 > 0$ and $\phi_2(e_1, 0)$ for $e_1 \geq 0$. Baseline parameterization. {fig:showcases}

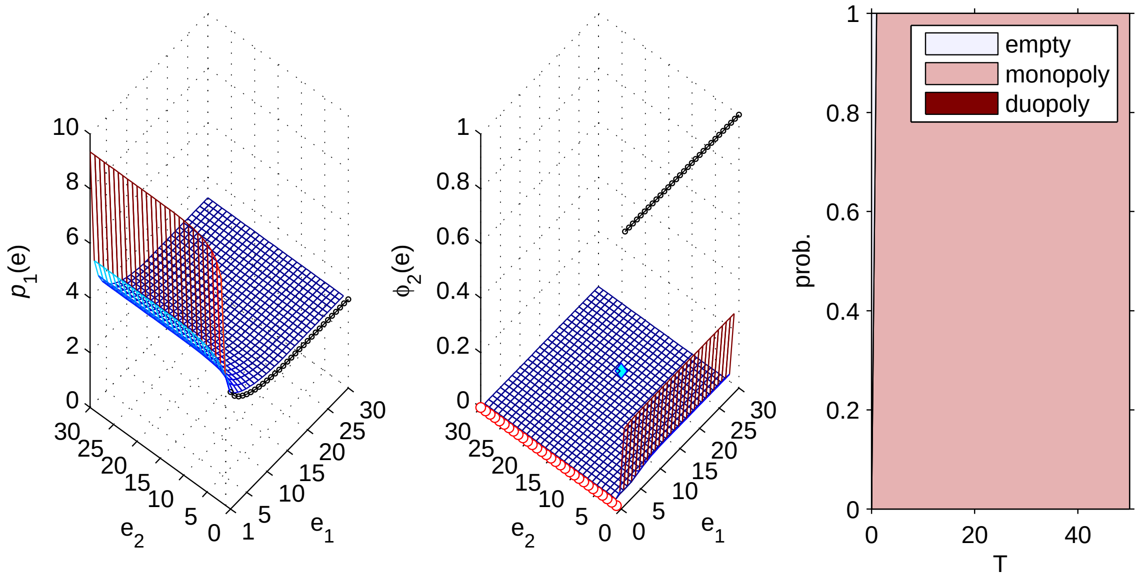

Figure 2: First-best planner solution. Pricing decision of firm 1 (left panel), non-operating probability of firm 2 (middle panel), and time path of probability distribution over industry structures (right panel). Dots beside the surface in left panel are $p_1(e_1, 0)$ for $e_1 > 0$ and dots in middle panel are $\phi_2(0, e_2)$ for $e_2 > 0$ and $\phi_2(e_1, 0)$ for $e_1 \geq 0$. Baseline parameterization. {fig:showcaseFC

mature duopoly if the follower manages to crash through the mobility barrier by winning a sale but, as the upper right panel of Figure 1 shows, this is far less likely than an entrenched monopoly.

The lower panels of Figure 1 are typical for what Besanko et al. (2014) call an *accommodative equilibrium*. There is a shallow well in state $(1, 1)$ with $p_1(1, 1) = 5.05$ as the lower left panel shows. A preemption battle is again likely to ensue because $\phi_1(0, 0) = 0.05$ implies that the probability that both firms enter the industry in period 0 is 0.91. After the industry has emerged from the preemption battle in state $(1, 1)$, the leader enjoys a competitive advantage over the follower. Without mobility barriers in the form of trenches, however, this advantage is temporary and the industry evolves into a mature duopoly as the lower right panel shows.

First-best planner solution. Figure 2 is analogous to Figure 1 and illustrates the first-best planner solution. In state $(0, 0)$ in period 0, the planner decides to operate a single firm (say firm 1) in the subsequent period since $\psi_{1,0}^{FB}(0, 0) = \psi_{0,1}^{FB}(0, 0) = 0.5$. In period $t \geq 1$, the planner marches firm 1 down its learning curve. As the left panel shows, $p_1^{FB}(\mathbf{e}) = 3.25$ if $e_1 \in \{15, \ldots, 30\}$ so that at the bottom of its learning curve firm 1 charges a price equal to marginal cost. In short, the planner operates the industry as a natural monopoly.

As the middle panel shows, there is an exit zone somewhat similar to the one in the aggressive equilibrium. Although state $(1, 1)$ is off the equilibrium path starting from state $(0, 0)$, $\psi_{1,0}^{FB}(1, 1) = \psi_{0,1}^{FB}(1, 1) = 0.04$ implies that if both firms are at the top of their learning curves, then the first-best planner ceases to operate one of them with probability 0.07 to receive the scrap value. On the other hand, if both firms are part of the way down their learning curves, then $\psi_{1,1}^{FB}(\mathbf{e}) = 1$ for $\mathbf{e} \geq (3,3)$ implies that the planner continues to operate both to secure the social benefit of product variety.

Outside the baseline parameterization in Table 2, the first-best planner does not necessarily operate the industry as a natural monopoly. In particular, if the degree of product differentiation is sufficiently large, then the planner immediately decides to operate both firms and continues to do so as they move down their learning curves.

	aggr. eqbm.	accom. eqbm.	planner solution	counter-factual
structure:				
expected short-run number of firms N_1	1.92	1.91	1.00	2.00
expected long-run number of firms N_∞	1.08	2.00	1.00	2.00
conduct:				
expected long-run average price $\overline{p}_\infty$	8.28	5.24	3.25	5.24
expected time to maturity T^m	19.09	37.54	15.02	53.91
performance:				
expected NPV of consumer surplus CS_β	93.87	103.29	131.66	56.88
expected NPV of total surplus TS_β	96.02	105.45	110.45	92.02
deadweight loss DWL_β	14.43	5.01	–	18.43
relative deadweight loss $\frac{DWL_\beta}{VA_\beta}$	13.06%	4.54%	–	16.69%

Table 3: Industry structure, conduct, and performance. Aggressive and accommodate equilibrium, first-best planner solution, and static non-cooperative pricing counterfactual. Baseline parameterization. {Table Industry

Industry structure, conduct, and performance. To succinctly describe an equilibrium and compare it to the first-best planner solution, we use several metrics of industry structure, conduct, and performance.[23] The second, third, and fourth columns of Table 3 show these metrics for the aggressive and accommodative equilibrium and the planner solution. (We discuss the last column of the table below.)

The expected short-run number of firms N_1 is just above 1.90 in both equilibria, compared to $N_1^{FB} = 1.00$ in the first-best planner solution. In the aggressive equilibrium, the expected long-run number of firms N_∞ is 1.08, quite close to the planner solution. In contrast, in the accommodative equilibrium, $N_\infty = 2.00$. The aggressive equilibrium therefore mainly involves over-entry and the accommodative equilibrium involves both over-entry and under-exit.

The expected long-run average price $\overline{p}_\infty^{FB} = 3.25$ in the first-best planner solution is equal to marginal cost at the bottom of the learning curve. It is much higher in both equilibria. In the aggressive equilibrium, in particular, $\overline{p}_\infty = 8.28$ reflects the fact that the industry most likely evolves into an entrenched monopoly.

[23]See Appendix A for formal definitions.

The expected time to maturity T^m is the expected time until the industry first becomes either a mature monopoly or a mature duopoly; it measures the speed at which firms move down their learning curves. Learning economies are exhausted fastest in the first-best planner solution with $T^{m,FB} = 15.02$, followed by the aggressive equilibrium with $T^m = 19.10$ and the accommodative equilibrium with $T^m = 37.50$. This large gap arises because sales are split between the inside goods in the accommodative equilibrium, as well as at least initially with the outside good.

As the industry is substantially more likely to be monopolized in the aggressive equilibrium than in the accommodative equilibrium, the expected NPV of consumer surplus CS_β is lower, as is the expected NPV of total surplus TS_β. Consequently, the deadweight loss DWL_β is higher in the aggressive equilibrium than in the accommodative equilibrium. However, the relative deadweight loss $\frac{DWL_\beta}{VA_\beta}$ seems modest, with 13.06% of the maximum value added by the industry in the aggressive equilibrium and 4.54% in the accommodative equilibrium.

6 Does dynamic competition lead to low deadweight loss? {Section: DWL

The relative deadweight loss $\frac{DWL_\beta}{VA_\beta}$ is modest more generally. Summarizing a large number

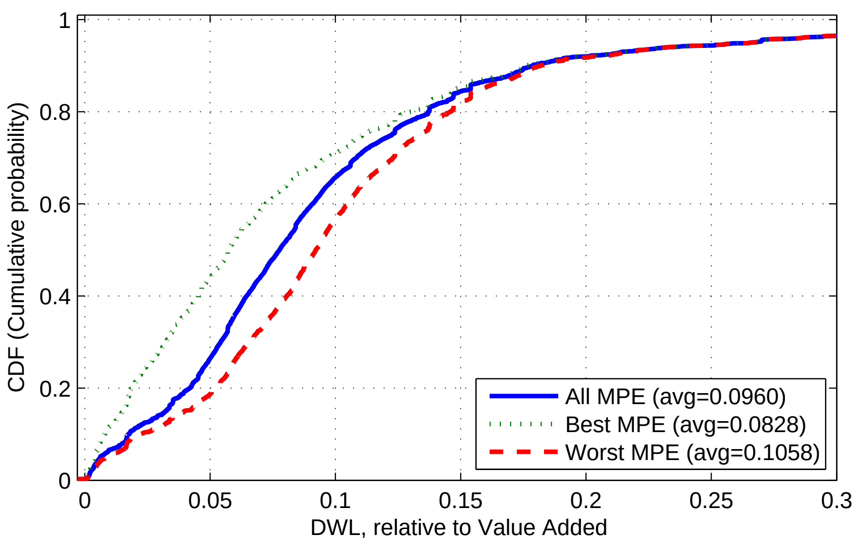

Figure 3: Distribution of relative deadweight loss $\frac{DWL_\beta}{VA_\beta}$. All equilibria (solid line), best equilibrium (dotted line), and worst equilibrium (dashed line). Parameterizations and equilibria within parameterizations weighted equally. {fig:DWLcdf}

of parameterizations and equilibria, Figure 3 shows the cumulative distribution function (CDF) of $\frac{DWL_\beta}{VA_\beta}$ as a solid line. Result 1 highlights some findings:

{Result: DWL u

Result 1 *The relative deadweight loss* $\frac{DWL_\beta}{VA_\beta}$ *is less than 5%, 10%, and 20% in 26.4%, 65.8%, and 92.0% of parameterizations, respectively. The median of* $\frac{DWL_\beta}{VA_\beta}$ *is 7.8%.*

There is a large relative deadweight loss $\frac{DWL_\beta}{VA_\beta}$ in a small number of parameterizations that occur when the industry almost fails to take off (in the sense that $1-\phi_1(0,0)^2 \approx 0.01$) because the outside good is highly attractive. Near this "cusp of viability" the contribution of the inside goods to the expected NPV of total surplus is small and thus $VA_\beta \approx 0$.

Recall that we average over equilibria at a given parameterization to obtain the distribution of $\frac{DWL_\beta}{VA_\beta}$. To look behind these averages, we consider the *best equilibrium* with the highest value of TS_β at a given parameterization as well as the *worst equilibrium* with the lowest value of TS_β. Figure 3 shows the resulting distributions of $\frac{DWL_\beta}{VA_\beta}$ using a dotted line for the best equilibrium and a dashed line for the worst equilibrium, and Result 2 summarizes:

{Result: DWL u

Result 2 *(1) For the best equilibrium, the relative deadweight loss* $\frac{DWL_\beta}{VA_\beta}$ *is less than 5%, 10%, and 20% in 44.2%, 71.1%, and 92.1% of parameterizations, respectively. The median of* $\frac{DWL_\beta}{VA_\beta}$ *is 5.7%. (2) For the worst equilibrium, the relative deadweight loss* $\frac{DWL_\beta}{VA_\beta}$ *is less than 5%, 10%, and 20% in 18.7%, 56.4%, and 91.8% of parameterizations, respectively. The median of* $\frac{DWL_\beta}{VA_\beta}$ *is 9.2%.*

Hence, even in the worst equilibria the relative deadweight loss $\frac{DWL_\beta}{VA_\beta}$ is modest for a wide range of parameterizations.

Closer inspection shows that the best equilibrium is often accommodative in nature whereas the worst equilibrium is often aggressive. In Appendix C we offer formal definitions of aggressive and accommodative equilibria and show that they are closely linked to the worst, respectively, best equilibria. To facilitate the exposition and build intuition, in what follows we therefore identify the best equilibrium with an accommodative equilibrium and the worst equilibrium—to the extent that it differs from the best equilibrium—with an aggressive equilibrium.[24] If the equilibrium is unique, then we identify it with an accommodative equilibrium.

6.1 Deadweight loss in perspective: static non-cooperative pricing counterfactual

Is a relative deadweight loss $\frac{DWL_\beta}{VA_\beta}$ of 10% of the maximum value added by the industry "small" and a relative deadweight loss of 30% "large"? To help put these percentages in perspective, we show that the deadweight loss is lower than expected in view of the two traditional roles of price (allocative and distributional). To this end, we shut down the investment role of price. In the price-setting phase, incumbent firm 1 is thus left to maximize static profit:

$$\max_{p_1} D_1(p_1, p_2^{SN}(\mathbf{e}))(p_1 - c(e_1)).$$

Hence, the pricing decision $p_1^{SN}(\mathbf{e})$ in this static non-cooperative pricing counterfactual is uniquely determined by the first-order condition

$$p_1^{SN}(\mathbf{e}) = c(e_1) + \frac{\sigma}{1 - D_1(p_1^{SN}(\mathbf{e}), p_2^{SN}(\mathbf{e}))}.$$

[24] This association is not perfect. For some parameterizations (e.g., those with weak product differentiation), there are multiple equilibria all of which are aggressive and none of which are accomodative.

The expected NPV of incumbent firm 1 is

$$V_1^{SN}(\mathbf{e}) = D_1(p_1^{SN}(\mathbf{e}), p_2^{SN}(\mathbf{e}))(p_1^{SN}(\mathbf{e}) - c(e_1))$$
$$+U_1^{SN}(\mathbf{e}) + \sum_{n=1}^{2} D_n(p_1^{SN}(\mathbf{e}), p_2^{SN}(\mathbf{e})) \left[U_1^{SN}\left(\mathbf{e}^{n+}\right) - U_1^{SN}(\mathbf{e})\right]$$

and, in contrast to the pricing decision, accounts for the impact of a sale on the value of continued play. Finally, the exit-entry phase remains unchanged.[25] Our computations always led to a unique solution.

The fifth column of Table 3 shows our metrics for industry structure, conduct, and performance for the static non-cooperative pricing counterfactual at the baseline parameterization. Similar to the accommodative equilibrium, the counterfactual involves both over-entry and under-exit ($N_1^{SN} = 2.00$ and $N_\infty^{SN} = 2.00$). Learning economies are exhausted even more slowly than in the accommodative equilibrium ($T^{m,SN} = 53.91 > 37.45 = T^m$) because firms ignore the investment role of price in making their pricing decisions. The deadweight loss DWL_β increases more than threefold relative to the accommodative equilibrium and by more than a quarter relative to the aggressive equilibrium.

The investment role of price is socially beneficial more generally. Figure 4 shows the

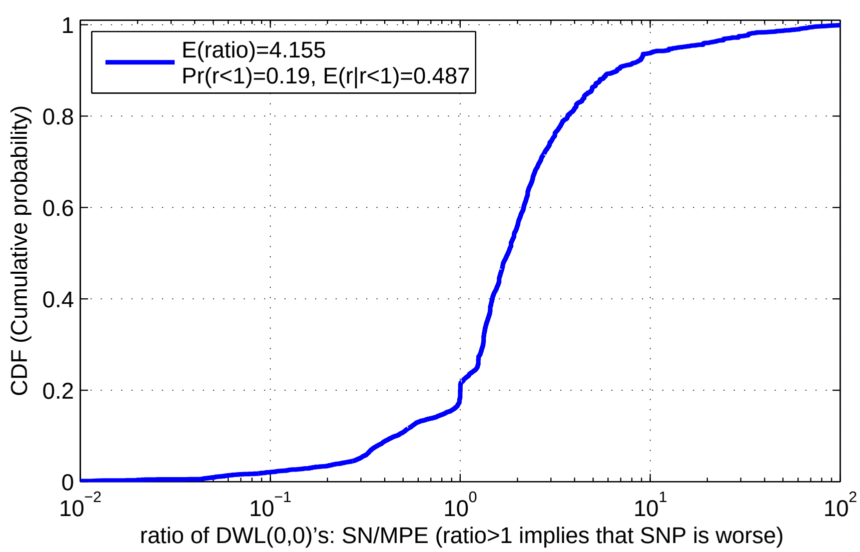

Figure 4: Distribution of deadweight loss ratio $\frac{DWL_\beta^{SN}}{DWL_\beta}$. Log scale. Parameterizations and equilibria within parameterizations weighted equally. {fig:SNDWLgap}

distribution of the deadweight loss ratio $\frac{DWL_\beta^{SN}}{DWL_\beta}$. Note that $\frac{DWL_\beta^{SN}}{DWL_\beta}$ is independent of our normalization by VA_β. Result 3 summarizes: {Result: diffe

Result 3 *DWL_β^{SN} is at least as large as DWL_β in 81% of cases, at least twice as large in 44% of cases, and at least five times as large in 14% of cases. The median of $\frac{DWL_\beta^{SN}}{DWL_\beta}$ is 11.*

[25] Our static non-cooperative pricing counterfactual loosely corresponds to the version of the war of attrition presented in Tirole (1988), with the addition of learning-by-doing and product differentiation.

DWL^{SN}_{β} is smaller than DWL_{β} in a number of parameterizations that mostly involve an unattractive outside good ($p_0 \geq 15$). Because the outside good constrains pricing decisions and profitability much more in a monopolistic than in a duopolistic industry, a less attractive outside good sharpens the incentive to monopolize the industry in equilibrium. But if firms ignore the investment role of price in the static non-cooperative pricing counterfactual, then a duopolistic industry with a lower deadweight loss emerges.

Generally speaking, we conclude that dynamic competition leads to low deadweight loss. The deadweight loss is low not only relative to the maximum value added by the industry, it is also smaller than the deadweight loss that arises if firms ignore the investment role of price.[26] Put differently, the investment role of price is by and large socially beneficial. The static non-cooperative pricing counterfactual also shows that a low deadweight loss is almost certainly not hardwired into the primitives of our learning-by-doing model. Instead, there is something in the nature of the investment role of price and dynamic competition that in equilibrium leads to low deadweight loss.

6.2 Differences between equilibria and first-best planner solution

Dynamic competition leads to low deadweight loss despite distortions in pricing, exit, and entry. Indeed, as we next show, there are typically substantial differences between the equilibria and the first-best planner solution. Paradoxically, the best equilibrium can differ even more from the planner solution than the worst equilibrium.

Recall that too low prices cause deadweight loss from overproduction, just as too high prices cause deadweight loss from underproduction. To illustrate that the equilibria involve prices that are too low, we first define $1\left[p_1(\mathbf{e}) < c(e_1) \text{ for some } \mathbf{e} \in \{1,\ldots,M\} \times \{0,\ldots,M\}\right]$ to indicate that a price is below the marginal cost of production in at least one state. Second, we define $1\left[p_1(\mathbf{e}) < p_1^{FB}(\mathbf{e}) \text{ for some } \mathbf{e} \in \{1,\ldots,M\} \times \{0,\ldots,M\}\right]$ to indicate that a price is below the first-best planner solution in some state. Figure 5 shows the distribution of these indicators and Result 4 summarizes:

{Result: MPE p

Result 4 *(1) $p_1(\mathbf{e}) < c(e_1)$ for some $\mathbf{e} \in \{1,\ldots,M\} \times \{0,\ldots,M\}$ in all equilibria in 80% of parameterizations.(2) $p_1(\mathbf{e}) < p_1^{FB}(\mathbf{e})$ for some $\mathbf{e} \in \{1,\ldots,M\} \times \{0,\ldots,M\}$ in all equilibria in 63% of parameterizations.*

We caution that the states with too low prices are not necessarily on the equilibrium path starting from state $(0,0)$.

We next turn from pricing to exit and entry and compare the expected short-run and long-run number of firms between the equilibria and the first-best planner solution. Figure 6 shows the distribution of $N_1 - N_1^{FB}$ as a solid line and Result 5 highlights some findings:

{Result: MPE i

Result 5 *N_1 is larger than N_1^{FB} in 78% of parameterizations and smaller than N_1^{FB} in less than 1% of parameterizations.*

[26] We also find that the deadweight loss under competition is lower than the deadweight loss that arises when firms behave collusively with respect to both pricing and entry/exit decisions. For the baseline parameterization, the deadweight loss as a percentage of value added in the fully collusive solution is 14.32 percent.

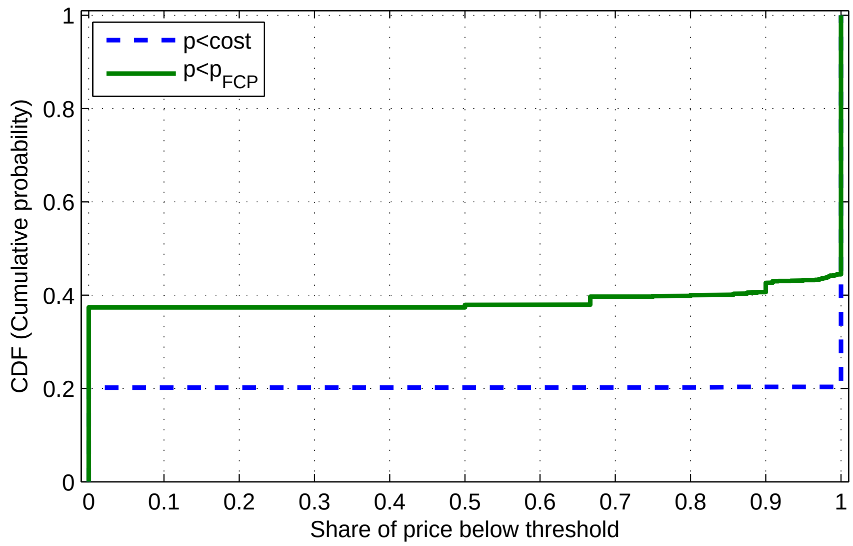

Figure 5: Distribution of $1\left[p_1(\mathbf{e}) < c(e_1) \text{ for some } \mathbf{e} \in \{1, \ldots, M\} \times \{0, \ldots, M\}\right]$ and $1\left[p_1(\mathbf{e}) < p_1^{FB}(\mathbf{e}) \text{ for some } \mathbf{e} \in \{1, \ldots, M\} \times \{0, \ldots, M\}\right]$. Parameterizations and equilibria within parameterizations weighted equally.

{Figure: MPE a

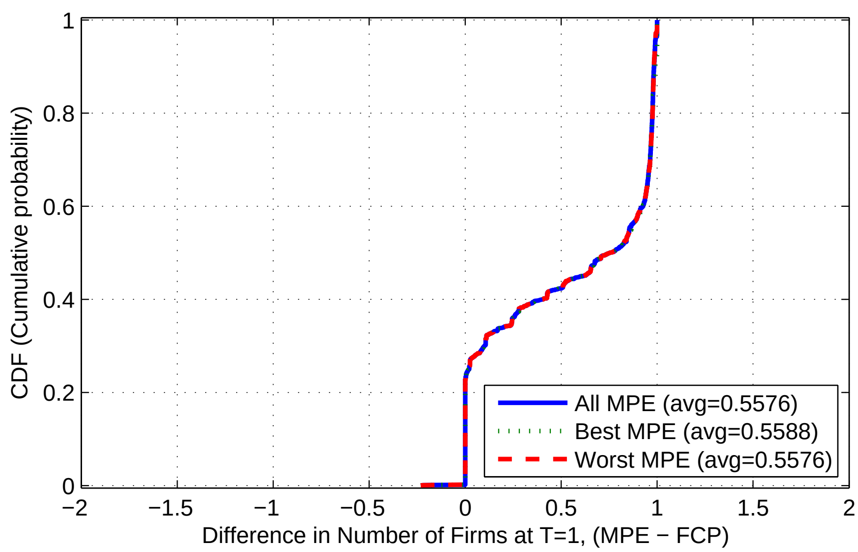

Figure 6: Distribution of $N_1 - N_1^{FB}$. All equilibria (solid line), best equilibrium (dotted line), and worst equilibrium (dashed line). Parameterizations and equilibria within parameterizations weighted equally.

{Figure: distr

Thus, the equilibria typically have too many firms in the short run, consistent with over-entry. They very rarely have too few firms in the short run. Figure 6 also breaks out the best equilibrium as a dotted line and the worst equilibrium as a dashed line. Similar to our examples in Section 5.2, there is no discernible difference between the best and the worst equilibrium.

Figure 7 shows the distribution of $N_\infty - N_\infty^{FB}$ as a solid line and breaks out the best equilibrium as a dotted line and the worst equilibrium as a dashed line. Result 6 summarizes:

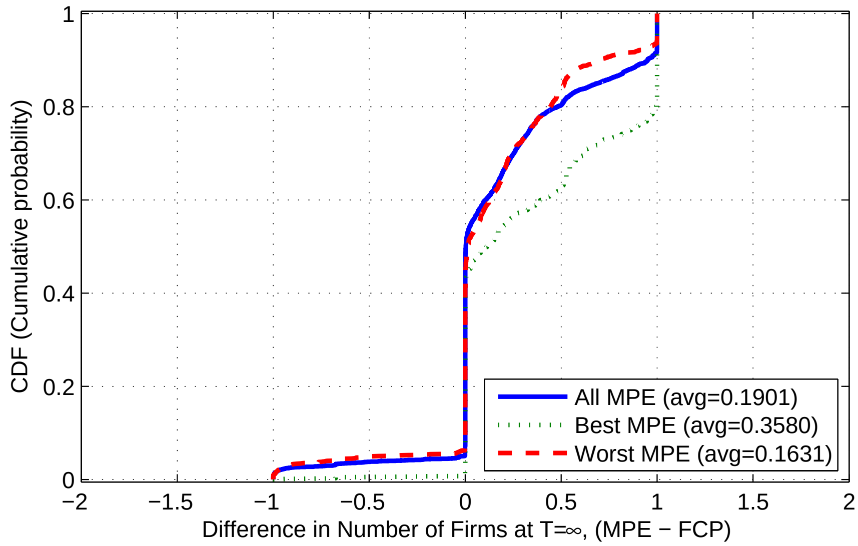

Figure 7: Distribution $N_\infty - N_\infty^{FB}$. All equilibria (solid line), best equilibrium (dotted line), and worst equilibrium (dashed line). Parameterizations and equilibria within parameterizations weighted equally.

{Figure: CDF o

{Result: MPE i

Result 6 *(1) N_∞ is larger than N_∞^{FB} in 60% of parameterizations and smaller than N_∞^{FB} in 5% of parameterizations. (2) For the best equilibrium, N_∞ is larger than N_∞^{FB} in 62% of parameterizations and smaller than N_∞^{FB} in 1% of parameterizations. (3) For the worst equilibrium, N_∞ is larger than N_∞^{FB} in 62% of parameterizations and smaller than N_∞^{FB} in 7% of parameterizations.*

Thus, the equilibria regularly have too many firms in the long run, consistent with under-exit. This tendency is exacerbated in the best equilibrium. The equilibria very rarely have too few firms in the long run.

We finally turn to the speed at which firms move down their learning curves. Recall that the expected time to maturity T^m depends on both the number of incumbent firms and their pricing decisions. Figure 8 shows the distribution of $T^m - T^{m,FB}$ as a solid line and breaks out the best equilibrium as a dotted line and the worst equilibrium as a dashed line. Result 7 summarizes:

{Result:TM}

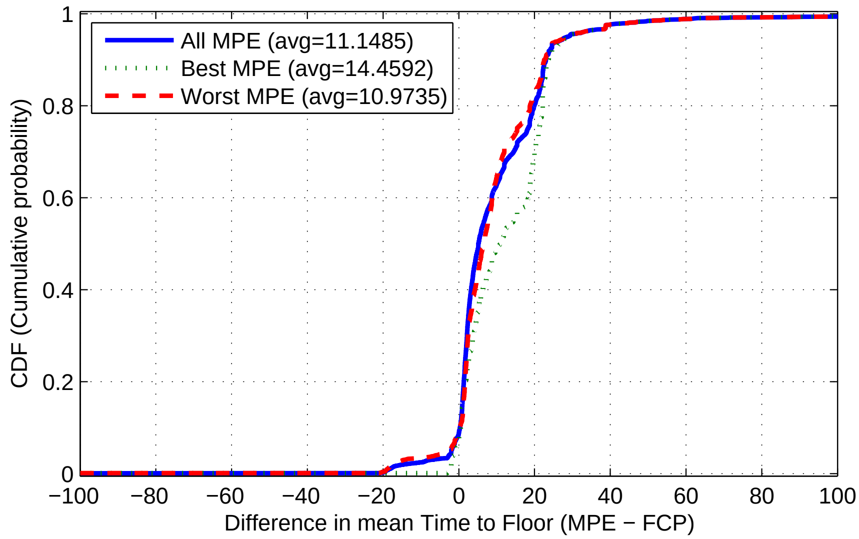

Figure 8: Distribution of $T^m - T^{m,FB}$. All equilibria (solid line), best equilibrium (dotted line), and worst equilibrium (dashed line). Parameterizations and equilibria within parameterizations weighted equally. {fig:TM}

Result 7 *(1) T^m is larger than $T^{m,FB}$ in 91% of parameterizations and smaller than $T^{m,FB}$ in 8% of parameterizations. (2) For the best equilibrium, T^m is larger than $T^{m,FB}$ in 94% of parameterizations and smaller than $T^{m,FB}$ in 7% of parameterization. (3) For the worst equilibrium, T^m is larger than $T^{m,FB}$ in 91% of parameterizations and smaller than $T^{m,FB}$ in 8% of parameterizations.*

The speed of learning in the equilibria is generally too slow. Moreover, the best equilibrium exhausts learning economies even more slowly than the worst equilibrium. This is because pricing is initially less aggressive and more firms split sales in an accommodative equilibrium than in an aggressive equilibrium.

7 Why does dynamic competition lead to low deadweight loss? {Section: Dec

Section 6 leaves us with a puzzle. Dynamic competition leads to low deadweight loss, but this is not because the equilibrium resembles the first-best planner solution. On the contrary, the best equilibrium can differ even more from the planner's solution than the worst equilibrium.

7.1 Decomposition

To better understand why dynamic competition leads to low deadweight loss, we quantify the importance of three factors that together make up deadweight loss: pricing conduct, exit and entry conduct, and market structure. We accordingly decompose the deadweight loss in equation (12) as

$$DWL_\beta = TS_\beta^{FB} - TS_\beta = DWL_\beta^{PR} + DWL_\beta^{EE} + DWL_\beta^{MS},$$

where

$$DWL_\beta^{PR} = \sum_{t=0}^{\infty} \beta^t \sum_{\mathbf{e}} \mu_t(\mathbf{e}) \left[CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e})) \right], \quad (17) \quad \{eq:T1\}$$

$$DWL_\beta^{EE} = \sum_{t=0}^{\infty} \beta^t \sum_{\mathbf{e}} \mu_t(\mathbf{e}) \left[PS^{FB}(\mathbf{e}) - \Pi^{FB}(\mathbf{e}) - (PS(\mathbf{e}) - \Pi(\mathbf{e})) \right], \quad (18) \quad \{eq:T2\}$$

$$DWL_\beta^{MS} = \sum_{t=0}^{\infty} \beta^t \sum_{\mathbf{e}} \left[\mu_t^{FB}(\mathbf{e}) - \mu_t(\mathbf{e}) \right] TS^{FB}(\mathbf{e}), \quad (19) \quad \{eq:T3\}$$

and, recall, $\Pi(\mathbf{e}) = \sum_{n=1}^{2} D_n(\mathbf{p}(\mathbf{e}))(p_n(\mathbf{e}) - c(e_n))$. $\Pi^{FB}(\mathbf{e})$ is defined analogously.

The *pricing distortion* DWL_β^{PR} in equation (17) is the incremental deadweight loss due to state-wise differences in pricing conduct between the equilibrium and the first-best planner solution.[27]

$PS(\mathbf{e}) - \Pi(\mathbf{e})$ is the difference of producer surplus and the static profit of the incumbent firms in state $\mathbf{e}$ and thus the part of producer surplus that accounts for scrap values and setup costs. The *exit and entry distortion* DWL_β^{EE} in equation (18) is therefore the incremental deadweight loss due state-wise differences in exit and entry conduct between the equilibrium and the first-best planner solution. Expected inflows from scrap values in state $\mathbf{e}$ contribute positively to $PS(\mathbf{e}) - \Pi(\mathbf{e})$ and expected outflows from setup costs negatively. A positive value of DWL_β^{EE} thus reflects a tendency for over-entry or under-exit relative to the first-best planner solution while a negative value reflects a tendency for under-entry or over-exit.

The *market structure distortion* DWL_β^{MS} in equation (19) is the incremental deadweight loss due to differences in the evolution of the industry over time between the equilibrium and the first-best planner solution. Recall that the state $\mathbf{e}$ completely describes the number of incumbent firms—and therefore the extent of product variety—along with their cost positions. A negative value of DWL_β^{MS} therefore indicates that the equilibrium puts more weight on more favorable market structures with higher values of $TS^{FB}(\mathbf{e})$ than the planner solution; a positive value indicates the reverse. Note that $TS^{FB}(\mathbf{e})$ is high if firms' cost positions in state $\mathbf{e}$ in relation to the price of the outside good yield large gains from trade.

[27] In our model, there is no distinction between price and quantity distortions. Indeed, consumer surplus can be written as a function of quantities rather than prices as

$$CS(\mathbf{Q}) = v - \sigma \sum_{n=0}^{2} \ln Q_n - \sum_{n=0}^{2} P_n(\mathbf{Q}) Q_n,$$

where $\mathbf{Q} \equiv (Q_0, Q_1, Q_2)$ and $(P_0(\mathbf{Q}),\ P_1(\mathbf{Q}), P_2(\mathbf{Q}))$ solve

$$\begin{aligned} P_0(\mathbf{Q}) &= p_0 \\ Q_1 &= D_1(p_0, P_1(\mathbf{Q}), P_2(\mathbf{Q})) \\ Q_2 &= D_2(p_0, P_1(\mathbf{Q}), P_2(\mathbf{Q})). \end{aligned}$$

$TS^{FB}(\mathbf{e})$ is also high if a large number of firms fosters product variety. Perhaps less obviously, $TS^{FB}(\mathbf{e})$ is high if in state $\mathbf{e}$ there are too many firms from the perspective of the first-best planner, thus allowing the planner to receive scrap values by ceasing to operate excess firms. Of course, these last two effects of the number of firms are mutually exclusive. The factors contributing to a negative value of DWL_β^{MS} are therefore over-entry and under-exit as well as fast exploitation of learning economies. Factors contributing to a positive value of DWL_β^{MS} are under-entry, over-exit, slow exploitation of learning economies, and cost inefficient exit. This last phenomenon—which we saw could arise in the two-step version of our model—contributes to a positive value of DWL_β^{MS} in the same way that slow exploitation of learning economies does: it makes it more likely that the industry stays in higher cost states than it would have under the planner's solution. Table 4 summarizes the distortions that the signs of DWL_β^{EE} and DWL_β^{MS} indicate.

term	positive	negative
DWL_β^{EE}	Over-entry, under-exit	Under-entry, over-exit
DWL_β^{MS}	Under-entry, over-exit, slow exploitation of learning economies, cost inefficient exit	Over-entry, under-exit, fast exploitation of learning economies

Table 4: Decomposition terms and contributing distortions. {Table: summar

DWL_β^{EE} and DWL_β^{MS} can offset each other as they depend on over-entry and under-exit in opposite ways. We therefore define the *non-pricing distortion* as $DWL_\beta^{NPR} = DWL_\beta^{EE} + DWL_\beta^{MS}$. It reflects (1) the *net* social loss from a suboptimal number of firms (setup costs net of scrap values net of social benefits of product variety) and (2) the *gross* social loss from a suboptimal exploitation of learning economies and (3) the *gross* social loss from cost-inefficient exit.

Examples. Table 5 illustrates the decomposition for the aggressive and accommodative equilibria at the baseline parameterization. The pricing distortion $DWL_\beta^{PR} = 10.78$ is the

	DWL_β	DWL_β^{PR}	DWL_β^{EE}	DWL_β^{MS}	DWL_β^{NPR}
aggr. eqbm.	14.43	10.78	4.67	-1.01	3.66
accom. eqbm.	5.01	2.35	7.32	-4.67	2.66

Table 5: Decomposition. Aggressive and accommodative equilibrium. Baseline parameterization. {tbl:DecompBase

largest part of deadweight loss $DWL_\beta = 14.43$ in the aggressive equilibrium. It is mainly driven by the high expected long-run average price $\overline{p}_\infty = 8.28$ (see Table 3) that results as the industry most likely evolves into an mature monopoly. In the accommodative equilibrium, the pricing distortion $DWL_\beta^{PR} = 2.35$ is a smaller part of deadweight loss $DWL_\beta = 5.01$ because the industry evolves into a mature duopoly. Interestingly, DWL_β^{PR} is small even though the expected long-run average price $\overline{p}_\infty = 5.24$ is almost two-thirds larger than in the first-best planner solution ($\overline{p}_{\infty}^{FB} = 3.25$).[28]

In the accommodative equilibrium, the largest part of deadweight loss $DWL_{\beta} = 5.01$ is the exit and entry distortion $DWL_{\beta}^{EE} = 7.32$. This reflects both wasteful duplication of setup costs due to over-entry and scrap values that firms forgo in equilibrium due to under-exit. The latter is analogous to the wasteful duplication of per-period, avoidable fixed costs that arises in the standard war of attrition model in Tirole (1988).[29] In contrast, in the aggressive equilibrium $DWL_{\beta}^{EE} = 4.67$ is a smaller part of $DWL_{\beta} = 14.43$ because the follower is likely to eventually exit the industry and receive the scrap value.

In both equilibria, the market structure distortion DWL_{β}^{MS} is negative, indicating that the industry spends more time in favorable market structures than in the first-best planner solution. This is driven by over-entry and under-exit as opposed to fast exploitation of learning economies. Indeed, in both equilibria learning economies are exhausted more slowly ($T^m = 19.10$ and $T^m = 37.50$, see again Table 3) than in the planner solution ($T^{m,FB} = 15.02$).

As DWL_{β}^{MS} partially offsets DWL_{β}^{EE}, the non-pricing distortion $DWL_{\beta}^{NPR} = 3.66$ is much smaller than the pricing distortion $DWL_{\beta}^{PR} = 10.78$ in the aggressive equilibrium. In the accommodative equilibrium the non-pricing distortion $DWL_{\beta}^{NPR} = 2.66$ is slightly larger than the pricing distortion $DWL_{\beta}^{PR} = 2.35$.

General results. Figure 9 shows the distribution of $\frac{DWL_{\beta}^{PR}}{DWL_{\beta}}$, $\frac{DWL_{\beta}^{EE}}{DWL_{\beta}}$, $\frac{DWL_{\beta}^{MS}}{DWL_{\beta}}$, and $\frac{DWL_{\beta}^{NPR}}{DWL_{\beta}}$. We scale each term of the decomposition by DWL_{β} to better gauge its size. Result 8 highlights some findings:

{Result: Sign

Result 8 *(1) The pricing distortion DWL_{β}^{PR} is positive in 98% of parameterizations. (2) The exit and entry distortion DWL_{β}^{EE} is positive in 81% of parameterizations. (3) The market structure distortion DWL_{β}^{MS} is negative in 70% of parameterizations. (4) The non-pricing distortion DWL_{β}^{NPR} is positive in 92% of parameterizations.*

DWL_{β}^{EE} is typically positive by part (2) of Result 8 because of over-entry and under-exit. Also because of over-entry and under-exit, DWL_{β}^{MS} is typically negative by part (3), as the speed of learning in the equilibria is generally too slow (see Figure 8 and Result 7). Thus, as highlighted in Result 9, DWL_{β}^{MS} typically offsets DWL_{β}^{EE} so that DWL_{β}^{NPR} is very often smaller than its largest component (in absolute value):

{Result: Offse

Result 9 *$|DWL_{\beta}^{NPR}|$ is less than* $\max\left\{\left|DWL_{\beta}^{EE}\right|, \left|DWL_{\beta}^{MS}\right|\right\}$ *in 88% of parameterizations.*

Parts (1) and (4) of Result 8 show that generally both pricing and non-pricing distortions contribute to deadweight loss. As Result 10 shows, the pricing distortion is often larger than the non-pricing distortion:

[28] Dynamic first-best prices $\left(p_1^{FB}(\mathbf{e}), p_2^{FB}(\mathbf{e})\right)$, in turn, coincide with static first-best prices $(c(e_1), c(e_2))$ if $e_1 \geq m$ and $e_2 \geq m$.

[29] In the Online Appendix, we will establish that our model with scrap values is formally equivalent to a model with per-period, avoidable fixed costs but without scrap values.

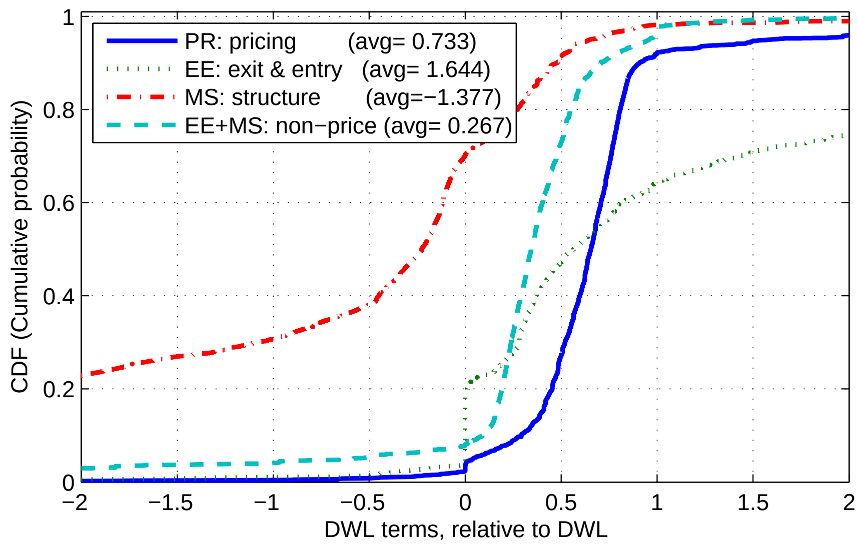

Figure 9: Distribution of $\frac{DWL^{PR}_{\beta}}{DWL_{\beta}}$, $\frac{DWL^{EE}_{\beta}}{DWL_{\beta}}$, $\frac{DWL^{MS}_{\beta}}{DWL_{\beta}}$, and $\frac{DWL^{NPR}_{\beta}}{DWL_{\beta}}$. Parameterizations and equilibria within parameterizations weighted equally. {fig:DecompMPE}

{Result: PR di

Result 10 *The pricing distortion DWL^{PR}_{β} is larger than the non-pricing distortion DWL^{NPR}_{β} in 73% of parameterizations.*

7.2 Why is the best equilibrium so good?

The deadweight loss in the best equilibrium is small although it often differs greatly from the first-best planner solution. Recall that the best equilibrium is often accommodative.

To see why the deadweight loss in an accommodative equilibrium is small, we show in a first step that the pricing distortion DWL^{PR}_{β} is small. Recall from Result 10 that the pricing distortion is often larger than the non-pricing distortion. In a second step, we argue that the non-pricing distortion DWL^{NPR}_{β} is small because of the social benefit of product variety.

Why is the pricing distortion small? At first blush, there appears to be no reason for the pricing distortion in an accommodative equilibrium to be particularly small. At the baseline parameterization, equilibrium prices $(p_1(\mathbf{e}), p_2(\mathbf{e}))$ substantially exceed static first-best prices $(c(e_1), c(e_2))$ even once the industry becomes a mature duopoly.

Proposition 2, however, bounds the contribution $CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e}))$ to the pricing distortion DWL^{PR}_{β} by the demand for the outside good $D_0(\mathbf{p}(\mathbf{e}))$:

{Proposition: '

Proposition 2 *Consider a symmetric state $\mathbf{e} = (e, e)$, where $e > 0$. If $p_0 \geq \kappa$, $p_1(\mathbf{e}) > c(e)$, and $D_0(\mathbf{p}(\mathbf{e})) < \frac{1}{2}$, then*

$$CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e})) \leq \frac{(p_1(\mathbf{e}) - c(e))^2}{\sigma} D_0(\mathbf{p}(\mathbf{e}))(1 - D_0(\mathbf{p}(\mathbf{e}))). \quad (20)$$ {eq: Taylor ap

Note that the bound on the right-hand side of equation (20) approaches zero as the demand for the outside good approaches zero. This often has bite: as the incumbent firms move down their learning curves, they improve their cost positions relative to the price of the outside good and drive the share of the outside good close to zero. Exceptions occur only if learning economies are weak—ρ is close to 1—or inconsequential because the outside good is highly attractive.

While the bound on the right-hand side of equation (20) relies on logit demand, the intuition is more general. To a first-order approximation, the deadweight loss due to market power decreases as market demand becomes less price elastic (Harberger 1954). With logit demand, a decrease in the price elasticity of the aggregate demand $D_1 + D_2$ for the inside goods is associated with a decreased share D_0 of the outside good. In the Online Appendix, we will show that a similar property holds for linear demand with differentiated products.

Why is the non-pricing distortion small? Recall that the non-pricing distortion $DWL_\beta^{NPR} = DWL_\beta^{EE} + DWL_\beta^{MS}$ reflects the net social loss from a suboptimal number of firms as well as the gross social losses from suboptimal exploitation of learning economies and cost inefficient exit. Recall, too, that, broadly speaking, in particular accommodative equilibria have too many firms, both in the short run and in the long run, consistent with over-entry and under-exit (Results 5, 6, and parts (2) and (3) of 8). Negative values of DWL_β^{MS} thus tend to offset positive values of DWL_β^{EE} (Result 9).

Of course, too many firms give rise to a social loss from the wasteful duplication of setup costs due to over-entry and scrap values that firms forgo in equilibrium due to under-exit. However, there is a "silver lining": too many firms give rise to a social benefit from additional product variety. Because of the social benefit of product variety, the net social loss from too many firms is small. This is especially relevant because accommodative equilibria tend to arise when the degree of product differentiation σ is high. Hence, the social benefit of product variety tends to be large.

Learning economies accentuate the silver lining of additional product variety. Suppose the industry evolves into a mature duopoly in an accommodative equilibrium and into a mature monopoly in the first-best planner solution, and that there is no further exit and entry. Then there exists a period t^* such that for period $t \geq t^*$, the equilibrium transient distribution $\mu_t(\mathbf{e})$ puts all mass on state $\mathbf{e}$, where $e_1 \geq m$ and $e_2 \geq m$, and the first-best transient distribution $\mu_t^{FB}(\mathbf{e})$ puts all mass on state $\mathbf{e}$, where either $e_1 \geq m$ and $e_2 = 0$ or $e_1 = 0$ and $e_2 \geq m$. Because learning economies are exhausted and there is no further exit and entry, we have $TS^{FB}(\mathbf{e}) = TS^{FB}(m, m)$ if $e_1 \geq m$ and $e_2 \geq m$, and $TS^{FB}(\mathbf{e}) = TS^{FB}(m, 0)$ if either $e_1 \geq m$ and $e_2 = 0$ or $e_1 = 0$ and $e_2 \geq m$. Moreover, $TS^{FB}(m, m) = CS^{FB}(m, m)$ and $TS^{FB}(m, 0) = CS^{FB}(m, 0)$ because the planner sets static first-best prices. Hence, the market structure distortion DWL_β^{MS} can be approximated as

$$\sum_{t=0}^{t^*} \beta^t \sum_{\mathbf{e}} \left[\mu_t^{FB}(\mathbf{e}) - \mu_t(\mathbf{e})\right] TS^{FB}(\mathbf{e}) - \frac{\beta^{t^*+1}}{1-\beta} \left[CS^{FB}(m, m) - CS^{FB}(m, 0)\right],$$

where the last term can be thought of as a reduction in DWL_β^{MS}—and thus in the non-pricing distortion DWL_β^{NPR}—due to the social benefit of additional product variety. From equation (10) it is straightforward to establish that

$$CS^{FB}(m,m) - CS^{FB}(m,0)$$
$$= \sigma\left(\ln\left(\exp\left(\frac{-p_0}{\sigma}\right) + 2\exp\left(\frac{-c(m)}{\sigma}\right)\right) - \ln\left(\exp\left(\frac{-p_0}{\sigma}\right) + \exp\left(\frac{-c(m)}{\sigma}\right)\right)\right),$$

where $c(m) = \kappa\rho^{\log_2 m}$ from equation (1), and

$$\frac{d(CS^{FB}(m,m) - CS^{FB}(m,0))}{d\rho} = -\left(2D_1(c(m),c(m)) - D_1(c(m),\infty)\right)\kappa\rho^{\log_2 m - 1}\log_2 m < 0,$$

since $2D_1(c(m),c(m)) = \frac{2}{2+\exp\left(-\left(\frac{p_0-c(m)}{\sigma}\right)\right)} > \frac{1}{1+\exp\left(-\left(\frac{p_0-c(m)}{\sigma}\right)\right)} = D_1(c(m),\infty)$. Hence, as learning economies strengthen, the reduction in DWL_β^{MS} and DWL_β^{NPR} due to the social benefit of additional product variety increases.

7.3 Why is the worst equilibrium not so bad?

Recall that the worst equilibrium is aggressive. To see why the deadweight loss in an aggressive equilibrium is small (albeit not as small as in the best equilibrium), we show in a first step that the pricing distortion DWL_β^{PR} is small. In a second step, we argue that the non-pricing distortion DWL_β^{NPR} is small because of a fairly efficient winnowing out of firms.

We emphasize that the arguments below do not imply that the deadweight loss in an aggressive equilibrium is close to zero. Rather, each one speaks to an economic force in the model that serves as a "headwind" that keeps the deadweight loss from becoming excessively large.

Why is the pricing distortion small? There again appears to be no reason for the pricing distortion in an aggressive equilibrium to be small. At the baseline parameterization, the industry evolves into a mature monopoly. Proposition 2 bounds the contribution $CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e}))$ to the pricing distortion DWL_β^{PR} by the degree of product differentiation σ and the advantage-building motive $U_1\left(\mathbf{e}^{1+}\right) - U_1\left(\mathbf{e}\right)$:

{Proposition:

Proposition 3 *Consider a state* $\mathbf{e} = (e, 0)$, *where* $e > 0$. *Then*

$$CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e}))$$
$$< \begin{cases} \sigma & \text{if} \quad 0 \leq U_1\left(\mathbf{e}^{1+}\right) - U_1\left(\mathbf{e}\right) < \sigma\left(1 + \exp\left(\frac{p_0 - c(e)}{\sigma}\right)\right), \\ \sigma + \left|U_1\left(\mathbf{e}^{1+}\right) - U_1\left(\mathbf{e}\right)\right| & \text{otherwise.} \end{cases} \quad (21)$$

{eq: Monopoly

Note that the bound on the right-hand side of equation (21) approaches σ as the incumbent firm moves down its learning curve and $U_1\left(\mathbf{e}^{1+}\right) - U_1\left(\mathbf{e}\right)$ approaches zero. While Proposition 3 relies on logit demand, the intuition that the threat of substitution to the outside good holds market power in check transcends the logit specification.

Proposition 3 has bite because aggressive equilibria tend to arise when the degree of product differentiation σ is low. This is intuitive: pricing aggressively to marginalize one's rival or altogether force it from the industry is especially attractive if products are close substitutes so that firms on an equal footing would fiercely compete.

Why is the non-pricing distortion small? While an aggressive equilibrium usually involves *delayed* exit (relative to the first-best planner solution), it involves rather brisk *eventual* exit. Thus, the industry usually quickly evolves towards the first-best market structure in an aggressive equilibrium, which tends to keep DWL_{β}^{MS}, and thus DWL_{β}^{NPR}, small. The small non-pricing distortion reflects the resulting relatively low net social loss from a suboptimal number of firms. Put another way, in an aggressive equilibrium, this net social loss is small because competition for the market resolves itself quickly and winnows out firms in a fairly efficient way.

8 Conclusion

{Section: Con

This paper studies a setting in which the traditional allocative and distributive roles of pricing are supplemented by a third role: investment. The investment role, which arises when firms jostle for competitive advantage through the prices they set, leads to dynamic competition.

In the introduction we floated the idea that it should be obvious that dynamic competition along a learning curve would be extremely efficient. In contrast to dynamic competition in rent-seeking models or even conventional wars of attrition where competition among firms squanders resources, competition in our setting creates socially valuable know-how as firms move down their learning curves. And instead of throwing this surplus away on wasteful activities, firms transfer it to consumers through lower prices.

However, our computations illustrate that the conduct and market structure that arise under dynamic competition often differ greatly from the conduct and structure that a first-best planner prefers: dynamic competition usually does not *look like* the socially efficient solution. Moreover, as the special case of our model with a two-step learning curve illustrates, even if price competition transfers *all* the benefits of learning economies from firms to consumers, the presence of a learning curve does not eliminate wasteful coordination failures that can arise in natural monopoly markets. In sum, dynamic competition when pricing has an investment role is not a "magic bullet" that completely eliminates inefficiencies that can arise in dynamic or static interactions in oligopoly markets.

Still, our results suggest the following conclusion: when price has an investment role dynamic competition operates *reasonably efficiently*. Learning economies in our model reduce competition from the outside good so much that it minimizes quantity distortions from oligopoly pricing. The presence of learning economies in our model also tends to bias the entry process toward generating more entry rather than less, making it more likely that there is eventually competition in the market. Of course, excessive entry into a market which a planner would monopolize is inefficient because it gives rise to a duplication of setup costs. But when there is sufficiently strong product differentiation, we tend to get accommodative equilibria in which two firms persist in the industry, creating long-run gains from product variety that partially offset this inefficiency. And when product differentiation is weak, aggressive equilibria arise which, although they usually involve delayed exit (relative to the first-best planner solution), involve fairly quick eventual exit, evolving the industry rapidly to the first-best market structure.

So why is dynamic competition reasonably efficient in our model? Ultimately, the answer boils down to the key fundamental that creates the investment role of price in the first place: the learning curve.

- The learning curve is a powerful force for minimizing pricing distortions in accommodative equilibria evolving into a mature duopoly in the long run.
- The learning curve helps lead to an outcome in which exit and entry distortions err on the side of too many firms than too few firms, and the distortions are partly offset by benefits from product variety (in accommodative equilibria) or eventual exit (in aggressive equilibria).
- Learning economies accentuate the long-run benefits from product variety when there is over-entry and under-exit in accommodative equilibria.

Put simply, when price serves as an investment, dynamic competition is efficient because of the efficiency-enhancing properties of the investment.

From a policy perspective, our results suggest that in settings in which price serves as an investment, the upside from competition policy or regulatory interventions—beyond those aimed at preventing firms from colluding on price and entry/exit decisions—is likely to be fairly limited.[30] Though we show in Besanko et al. (2014) that welfare gains from conduct restrictions on pricing are possible in the setting we analyze here, achieving those gains would require detailed knowledge of demand and cost parameters. Unless a competition authority executes flawlessly based on exact knowledge of market primitives, it may be better not to intervene (except to police collusive behavior) and tolerate the "not so bad" welfare losses that would typically arise under dynamic competition.

[30]Analysis (not shown in the paper) suggests that enforcement of statutes against collusion is likely to be important in our setting. Not only does full collusion lead to higher deadweight losses than dynamic competition does, it results in significantly lower consumer surplus. For example, in the baseline parameterization, the discounted present value of consumer surplus under the fully collusive solution is 32.54, as compared to 93.87 and 103.29 under the aggressive and acommodative equilibrium, respectively. (The latter two numbers come from Table 3

A Omitted expressions {Section: App

Exit decision of incumbent firm. The probability of incumbent firm 1 exiting the industry in state $\mathbf{e}'$ is {ap:ExitExpress

$$
\begin{aligned}
\phi_1(\mathbf{e}') &= 1 - F_X(\widehat{X}_1(\mathbf{e}')) \\
&= \begin{cases} 1 & \text{if} \quad \widehat{X}_1(\mathbf{e}') < \overline{X} - \Delta_X, \\ \frac{1}{2} - \frac{\left[\widehat{X}_1(\mathbf{e}') - \overline{X}\right]}{2\Delta_X} & \text{if} \quad \widehat{X}_1(\mathbf{e}') \in [\overline{X} - \Delta_X, \overline{X} + \Delta_X], \\ 0 & \text{if} \quad \widehat{X}_1(\mathbf{e}') > \overline{X} + \Delta_X \end{cases}
\end{aligned}
$$

and the expectation of the scrap value conditional on exiting the industry is

$$
\begin{aligned}
E_X\left[X_1 | X_1 \geq \widehat{X}_1(\mathbf{e}')\right] &= \frac{\int_{F_X^{-1}(1-\phi_1(\mathbf{e}'))}^{\overline{X}+\Delta_X} X_1 dF_X(X_1)}{\phi_1(\mathbf{e}')} \\
&= \frac{1}{\phi_1(\mathbf{e}')} \left[Z_X(0) - Z_X\left(1 - \phi_1(\mathbf{e}')\right)\right],
\end{aligned}
$$

where

$$
Z_X(1-\phi) = \frac{1}{\Delta_X^2} \begin{cases} -\frac{1}{6}\left(\overline{X} - \Delta_X\right)^3 & \text{if} \quad 1-\phi \leq 0, \\ \frac{1}{2}\left(\Delta_X - \overline{X}\right)\left(F_X^{-1}(1-\phi)\right)^2 + \frac{1}{3}\left(F_X^{-1}(1-\phi)\right)^3 & \text{if} \quad 1-\phi \in \left[0, \frac{1}{2}\right], \\ \frac{1}{2}\left(\Delta_X + \overline{X}\right)\left(F_X^{-1}(1-\phi)\right)^2 - \frac{1}{3}\left(F_X^{-1}(1-\phi)\right)^3 - \frac{1}{3}\overline{X}^3 & \text{if} \quad 1-\phi \in \left[\frac{1}{2}, 1\right], \\ \frac{1}{6}\left(\overline{X} + \Delta_X\right)^3 - \frac{1}{3}\overline{X}^3 & \text{if} \quad 1-\phi \geq 1 \end{cases}
$$

and

$$
F_X^{-1}(1-\phi) = \overline{X} + \Delta_X \begin{cases} -1 & \text{if} \quad 1-\phi \leq 0, \\ -1 + \sqrt{2(1-\phi)} & \text{if} \quad 1-\phi \in \left[0, \frac{1}{2}\right], \\ 1 - \sqrt{2\phi} & \text{if} \quad 1-\phi \in \left[\frac{1}{2}, 1\right], \\ 1 & \text{if} \quad 1-\phi \geq 1. \end{cases}
$$

Entry decision of potential entrant. The probability of potential entrant 1 not entering the industry in state $\mathbf{e}'$ is {ap:EntryExpres

$$
\begin{aligned}
\phi_1(\mathbf{e}') &= 1 - F_S(\widehat{S}_1(\mathbf{e}')) \\
&= \begin{cases} 1 & \text{if} \quad \widehat{S}_1(\mathbf{e}') < \overline{S} - \Delta_S, \\ \frac{1}{2} - \frac{\left[\widehat{S}_1(\mathbf{e}') - \overline{S}\right]}{2\Delta_S} & \text{if} \quad \widehat{S}_1(\mathbf{e}') \in [\overline{S} - \Delta_S, \overline{S} + \Delta_S], \\ 0 & \text{if} \quad \widehat{S}_1(\mathbf{e}') > \overline{S} + \Delta_S \end{cases}
\end{aligned}
$$

and the expectation of the setup cost conditional on entering the industry is

$$
\begin{aligned}
E_S\left[S_1 | S_1 \leq \widehat{S}_1(\mathbf{e}')\right] &= \frac{\int_{\overline{S}-\Delta_S}^{F_S^{-1}(1-\phi_1(\mathbf{e}'))} S_1 dF_S(S_1)}{(1 - \phi_1(\mathbf{e}'))} \\
&= \frac{1}{\phi_1(\mathbf{e}')} \left[Z_S\left(1 - \phi_1(\mathbf{e}')\right) - Z_S(1)\right],
\end{aligned}
$$

where

$$Z_S(1-\phi) = \frac{1}{\Delta_S^2}\begin{cases} -\frac{1}{6}\left(\overline{S}-\Delta_S\right)^3 & \text{if} \quad 1-\phi \leq 0, \\ \frac{1}{2}\left(\Delta_S-\overline{S}\right)\left(F_S^{-1}(1-\phi)\right)^2+\frac{1}{3}\left(F_S^{-1}(1-\phi)\right)^3 & \text{if} \quad 1-\phi \in \left[0,\frac{1}{2}\right], \\ \frac{1}{2}\left(\Delta_S+\overline{S}\right)\left(F_S^{-1}(1-\phi)\right)^2-\frac{1}{3}\left(F_S^{-1}(1-\phi)\right)^3-\frac{1}{3}\overline{S}^3 & \text{if} \quad 1-\phi \in \left[\frac{1}{2},1\right], \\ \frac{1}{6}\left(\overline{S}+\Delta_S\right)^3-\frac{1}{3}\overline{S}^3 & \text{if} \quad 1-\phi \geq 1 \end{cases}$$

and

$$F_S^{-1}(1-\phi) = \overline{S}+\Delta_S\begin{cases} -1 & \text{if} \quad 1-\phi \leq 0, \\ -1+\sqrt{2(1-\phi)} & \text{if} \quad 1-\phi \in \left[0,\frac{1}{2}\right], \\ 1-\sqrt{2\phi} & \text{if} \quad 1-\phi \in \left[\frac{1}{2},1\right], \\ 1 & \text{if} \quad 1-\phi \geq 1. \end{cases}$$

Producer surplus (decentralized exit and entry). Producer surplus in state $\mathbf{e}$ is

$$PS(\mathbf{e}) = \sum_{n=1}^{2} PS_n(\mathbf{e}),$$

where

$$PS_1(\mathbf{e}) = \Pi_1(\mathbf{e}) + \sum_{n=0}^{2} D_n(\mathbf{p}(\mathbf{e}))\Big\{1\left[e_1 \neq 0\right]\phi_1\left(\mathbf{e}^{n+}\right)E_X\left[X_1|X_1 \geq \widehat{X}_1\left(\mathbf{e}^{n+}\right)\right]$$
$$-1\left[e_1 = 0\right]\left(1-\phi_1\left(\mathbf{e}^{n+}\right)\right)E_S\left[S_1|S_1 \leq \widehat{S}_1\left(\mathbf{e}^{n+}\right)\right]\Big\}$$

is producer surplus of firm 1 in state $\mathbf{e}$ with

$$\Pi_1(\mathbf{e}) = D_1(\mathbf{p}(\mathbf{e}))(p_1(\mathbf{e})-c(e_1)).$$

$PS_2(\mathbf{e})$ and $\Pi_2(\mathbf{e})$ are analogous.

Producer surplus (centralized exit and entry). Producer surplus in state $\mathbf{e}$ is

$$PS^{FB}(\mathbf{e}) = \Pi^{FB}(\mathbf{e}) + \sum_{n=0}^{2} D_n\left(\mathbf{p}^{FB}(\mathbf{e})\right)$$
$$\times\Big\{\psi_{1,1}^{FB}\left(\mathbf{e}^{n+}\right)E_{\mathbf{X},\mathbf{S}}\left[-1\left[e_1=0\right]S_1-1\left[e_2=0\right]S_2|U^{FB}\left(\mathbf{e}^{n+},\mathbf{X},\mathbf{S}\right)=U_{1,1}^{FB}\left(\mathbf{e}^{n+},\mathbf{X},\mathbf{S}\right)\right]$$
$$+\psi_{1,0}^{FB}\left(\mathbf{e}^{n+}\right)E_{\mathbf{X},\mathbf{S}}\left[-1\left[e_1=0\right]S_1+1\left[e_2\neq 0\right]X_2|U^{FB}\left(\mathbf{e}^{n+},\mathbf{X},\mathbf{S}\right)=U_{1,0}^{FB}\left(\mathbf{e}^{n+},\mathbf{X},\mathbf{S}\right)\right]$$
$$+\psi_{0,1}^{FB}\left(\mathbf{e}^{n+}\right)E_{\mathbf{X},\mathbf{S}}\left[1\left[e_1\neq 0\right]X_1-1\left[e_2=0\right]S_2|U^{FB}\left(\mathbf{e}^{n+},\mathbf{X},\mathbf{S}\right)=U_{0,1}^{FB}\left(\mathbf{e}^{n+},\mathbf{X},\mathbf{S}\right)\right]$$
$$+\psi_{0,0}^{FB}\left(\mathbf{e}^{n+}\right)E_{\mathbf{X},\mathbf{S}}\left[1\left[e_1\neq 0\right]X_1+1\left[e_2\neq 0\right]X_2|U^{FB}\left(\mathbf{e}^{n+},\mathbf{X},\mathbf{S}\right)=U_{0,0}^{FB}\left(\mathbf{e}^{n+},\mathbf{X},\mathbf{S}\right)\right]\Big\}$$

with $\Pi^{FB}(\mathbf{e}) = \sum_{n=1}^{2} D_n\left(\mathbf{p}^{FB}(\mathbf{e})\right)\left(p_n^{FB}(\mathbf{e})-c(e_n)\right)$.

Industry structure, conduct, and performance. The expected short-run and long-run number of firms is

$$N_1 = \sum_{\mathbf{e}} \mu_1(\mathbf{e}) N(\mathbf{e}), \quad N_\infty = \sum_{\mathbf{e}} \mu_\infty(\mathbf{e}) N(\mathbf{e}),$$

where number of firms in state $\mathbf{e}$ is

$$N(\mathbf{e}) = \sum_{n=1}^{2} 1[e_n > 0].$$

The expected long-run average price is

$$\overline{p}_\infty = \sum_{\mathbf{e} \geq (0,0)} \frac{\mu_\infty(\mathbf{e})}{1 - \mu_\infty(0,0)} \overline{p}(\mathbf{e}),$$

where (share-weighted) average price in state $\mathbf{e}$ is

$$\overline{p}(\mathbf{e}) = \sum_{n=1}^{2} \frac{D_n(p_1(\mathbf{e}), p_2(\mathbf{e}))}{1 - D_0(p_1(\mathbf{e}), p_2(\mathbf{e}))} p_n(\mathbf{e}).$$

The expected time to maturity is

$$T^m = E[\min\{t \geq 0 | \mathbf{e}_t \in \Omega\} | \mathbf{e}_0 = (0,0)],$$

where $\mathbf{e}_t$ is the state of the industry in period t and

$$\Omega = \{(m,0), \ldots, (M,0), (0,m), \ldots, (0,M), (m,m), \ldots, (M,M)\}$$

is the set of states in which the industry is either a mature monopoly or a mature duopoly. $\min\{t \geq 0 | \mathbf{e}_t \in \Omega\}$ is the so-called first passage time into the set of states Ω. It can be shown that T^m is the solution to a system of linear equations (Kulkarni 1995, equation (4.72)).

The expected NPV of consumer surplus CS_β is defined analogously to the expected NPV of total surplus TS_β in equation (11).

B Proofs

{Section: App

Proof of Proposition 1. The method of proof is to show, first, that the value functions and the policy functions in Table 1 are consistent with each other in every state and them to show that in every state, firms would not make a one-shot deviation from their equilibrium strategies. The details of this analysis will be presented in the Online Appendix. ■

Proof of Proposition 2. Define the sum of consumer surplus and static profit to be

$$\Phi(p) = CS(p,p) + 2D_1(p,p)(p - c(e)).$$

Using this definition, in a symmetric state $\mathbf{e} = (e,e)$, where $e > 0$, we can write

$$CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e})) = \Phi(p_1^{FB}(\mathbf{e})) - \Phi(p_1(\mathbf{e})).$$

We have

$$\Phi'(p) = -\frac{1}{\sigma}(p - c(e))D_0(p,p)(1 - D_0(p,p)), \qquad (22)$$

{eq: Deriv of I

$$\Phi''(p) = -\frac{1}{\sigma^2}\left((p - c(e))(1 - 2D_0(p,p)) + \sigma\right) D_0(p,p)(1 - D_0(p,p)). \qquad (23)$$

{eq: 2nd deriv

Hence, $\Phi(p)$ is strictly quasiconcave in p and attains its maximum at $p = c(e)$. Thus, we obtain

$$CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e})) \leq \Phi(c(e)) - \Phi(p_1(\mathbf{e})). \qquad (24)$$

{eq: prelimina

We bound the right-hand side of equation (24). Let $\widetilde{p}$ be such that $D_0(\widetilde{p},\widetilde{p}) = \frac{1}{2}$, so $1-2D_0(p,p) \geq 0$ for all $p \leq \widetilde{p}$ because $D_0(p,p)$ increases in p. Equation (23) implies that $\Phi(p)$ is strictly concave in p over the interval $[c(e), \widetilde{p}]$. This interval is non-empty: the assumption $p_0 \geq \kappa$ coupled with $\kappa \geq c(e)$ implies $D_0(c(e), c(e)) \leq D_0(\kappa,\kappa) \leq \frac{1}{3} < \frac{1}{2} = D_0(\widetilde{p},\widetilde{p})$. As $D_0(p,p)$ increases in p, it must be that $c(e) < \widetilde{p}$.

By assumption, $p_1(\mathbf{e}) \in [c(e), \widetilde{p}]$. From Theorem 21.2 in Simon & Blume (1994) and equation (22) we therefore have

$$\begin{aligned}\Phi(c(e)) - \Phi(p_1(\mathbf{e})) &\leq \Phi'(p_1(\mathbf{e}))(c(e) - p_1(\mathbf{e})) \\ &= \frac{(p_1(\mathbf{e}) - c(e))^2}{\sigma} D_0(p_1(\mathbf{e}), p_1(\mathbf{e}))(1 - D_0(p_1(\mathbf{e}), p_1(\mathbf{e}))). \qquad (25)\end{aligned}$$

This establishes Proposition 2. ■

Proof of Proposition 3. Define the sum of consumer surplus and static profit to be

$$\Phi(p) = CS(p,\infty) + D_1(p,\infty)(p - c(e)).$$

Using this definition, in a state $\mathbf{e} = (e, 0)$, where $e > 0$, we can write

$$CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e})) = \Phi(p_1^{FB}(\mathbf{e})) - \Phi(p_1(\mathbf{e})).$$

We have

$$\Phi'(p) = -\frac{1}{\sigma}(p - c(e))D_0(p,\infty)(1 - D_0(p,\infty)),$$

$$\Phi''(p) = -\frac{1}{\sigma^2}\left((p - c(e))(1 - 2D_0(p,\infty)) + \sigma\right) D_0(p,\infty)(1 - D_0(p,\infty)).$$

Hence, $\Phi(p)$ is strictly quasiconcave in p and attains its maximum at $p = c(e)$. Thus, we obtain

$$CS^{FB}(\mathbf{e}) + \Pi^{FB}(\mathbf{e}) - (CS(\mathbf{e}) + \Pi(\mathbf{e})) \leq \Phi(c(e)) - \Phi(p_1(\mathbf{e})),$$

where $\Phi(c(e)) = v - c(e) + \sigma \ln\left(1 + \exp\left(\frac{c(e)-p_0}{\sigma}\right)\right)$.

$p_1(\mathbf{e})$ is uniquely determined by the solution to the first-order condition (5); it can be written as

$$p_1(\mathbf{e}) = c(e) - \left[U_1\left(\mathbf{e}^{1+}\right) - U_1\left(\mathbf{e}\right)\right] + \sigma\left(1 + W\left(\exp\left(\frac{p_0 - c(e) + \left[U_1\left(\mathbf{e}^{1+}\right) - U_1\left(\mathbf{e}\right)\right]}{\sigma} - 1\right)\right)\right),$$

where $W(\cdot)$ is the Lambert W function. Defining $x = \frac{p_0 - c(e)}{\sigma}$ and $y = \frac{U_1(\mathbf{e}^{1+}) - U_1(\mathbf{e})}{\sigma}$, this can be further written as

$$p_1(\mathbf{e}) = c(e) + \sigma\left(-y + 1 + W\left(\exp(x + y - 1)\right)\right).$$

Hence,

$$\Phi(p_1(\mathbf{e})) = v - c(e) + \sigma\left(\ln\left(\frac{1 + W\left(\exp(x + y - 1)\right)}{W\left(\exp(x + y - 1)\right)}\right) + \frac{y}{1 + W\left(\exp(x + y - 1)\right)} - 1\right)$$

and

$$\Phi(c(e)) = v - c(e) + \sigma \ln\left(1 + \exp(-x)\right).$$

It follows that

$$\begin{aligned} &\Phi(c(e)) - \Phi(p_1(\mathbf{e})) \\ &= \sigma\left(\ln\left(1 + \exp(-x)\right) - \ln\left(\frac{1 + W\left(\exp(x + y - 1)\right)}{W\left(\exp(x + y - 1)\right)}\right) - \frac{y}{1 + W\left(\exp(x + y - 1)\right)} + 1\right) \quad (26) \quad \{\texttt{form1}\} \\ &= \sigma\left(\ln\left(1 + \exp(x)\right) - \ln\left(1 + W\left(\exp(x + y - 1)\right)\right) - W\left(\exp(x + y - 1)\right) + \frac{yW\left(\exp(x + y - 1)\right)}{1 + W\left(\exp(x + y - 1)\right)}\right). \quad (27) \quad \{\texttt{form2}\} \end{aligned}$$

Some properties of the Lambert W function are that $W(z)$ is increasing in z, $W(0) = 0$, and $W(z\exp(z)) = z$ for all $z \geq 0$.

Case 1: $y < 1 + \exp(x)$. We first show that if $y < 1 + \exp(x)$, then

$$\ln\left(1 + \exp(-x)\right) - \ln\left(\frac{1 + W\left(\exp(x + y - 1)\right)}{W\left(\exp(x + y - 1)\right)}\right) + 1 < 1. \quad (28) \quad \{\texttt{cond2}\}$$

To see this, note that equation (28) is equivalent to

$$\begin{aligned} \ln\left(1 + \exp(-x)\right) &\leq \ln\left(\frac{1 + W\left(\exp(x + y - 1)\right)}{W\left(\exp(x + y - 1)\right)}\right) \\ \Leftrightarrow \exp(-x) &< \frac{1}{W\left(\exp(x + y - 1)\right)} \\ \Leftrightarrow \exp(x) &> W\left(\exp(x + y - 1)\right). \end{aligned}$$

If $y = 1+\exp(x)$, then the right-hand side is $W\left(\exp(x + \exp(x))\right) = W\left(\exp(x)\exp(\exp(x))\right) = \exp(x)$. Moreover, because $W(z)$ is increasing in z, $\exp(x) > W\left(\exp(x + y - 1)\right)$ for all $y < 1 + \exp(x)$.

Consider equation (26). From equation (28) it follows that

$$\Phi(c(e)) - \Phi(p_1(\mathbf{e})) < \sigma\left(1 - y\frac{1}{1 + W\left(\exp(x + y - 1)\right)}\right).$$

Moreover, $0 < \frac{1}{1+W(\exp(x+y-1))} < 1$. Therefore, if $y < 0$, then

$$\Phi(c(e)) - \Phi(p_1(\mathbf{e})) < \sigma\,(1 + |y|) \tag{29}$$ {ineqa}

and, if $y \geq 0$, then

$$\Phi(c(e)) - \Phi(p_1(\mathbf{e})) < \sigma. \tag{30}$$ {ineqb}

Case 2: $y \geq 1 + \exp(x)$. We first show that if $y \geq 1 + \exp(x)$, then

$$\ln(1 + \exp(x)) - \ln(1 + W(\exp(x + y - 1))) - W(\exp(x + y - 1)) < 1. \tag{31}$$ {cond3}

To see this, note that

$$\begin{aligned}
&\ln(1 + \exp(x)) - \ln(1 + W(\exp(x + y - 1))) - W(\exp(x + y - 1)) \\
&\leq \ln(1 + \exp(x)) - \ln(1 + W(\exp(x + \exp(x)))) - W(\exp(x + \exp(x))) \\
&= \ln(1 + \exp(x)) - \ln(1 + \exp(x)) - \exp(x) = -\exp(x) < 1.
\end{aligned}$$

Consider equation (27). From equation (31) it follows that

$$\Phi(c(e)) - \Phi(p_1(\mathbf{e})) < \sigma\left(1 + y\frac{W(\exp(x + y - 1))}{1 + W(\exp(x + y - 1))}\right).$$

Moreover, $0 < \frac{W(\exp(x+y-1))}{1+W(\exp(x+y-1))} < 1$. Because $y \geq 1 + \exp(x) > 0$, we have

$$\Phi(c(e)) - \Phi(p_1(\mathbf{e})) < \sigma\,(1 + y)\,. \tag{32}$$ {ineqc}

Collecting equations (29), (30), and (32) establishes Proposition 3. ■

C Aggressive and accommodative equilibria

{Section: App

We offer formal definitions of aggressive and accommodative equilibria, but note from the outset that any attempt to classify equilibria is fraud with difficulty because the different equilibria lie on a continuum and thus morph into each other in complicated ways as we vary the parameters of the model.

Our definition of an aggressive equilibrium hones in on a trench in the pricing decision, and our definition of an accommodative equilibrium on a lack of exit from a duopolistic industry:

Definition 1 *An equilibrium is* aggressive *if*

$$p_1(\mathbf{e}) < p_1(e_1, e_2 + 1), \quad p_2(\mathbf{e}) < p_2(e_1, e_2 + 1), \quad \phi_2(\mathbf{e}) > \phi_2(e_1, e_2 + 1)$$

for some state $\mathbf{e} > (0, 0)$ *with* $e_1 > 1$ *and* $e_1 > e_2$*,*

	aggressive	accommodative	unclassified
best	46.40%	98.04%	58.02%
worst	98.11%	66.17%	27.67%

Table 6: Percentage of parameterizations at which an aggressive, accommodative, or unclassified equilibrium (if one exists) is best or worst. Unweighted. {TAB:CLASS1}

	unique equilibrium	multiple equilibria	
		best	worst
aggressive	2.83%	43.19%	97.96%
accommodative	83.82%	40.66%	1.36%
unclassified	13.35%	16.15%	0.68%

Table 7: Percentage of parameterizations at which the best or worst equilibrium is aggressive, accommodative, or unclassified. Unweighted. {TAB:CLASS2}

Definition 2 *An equilibrium is* accommodative *if*

$$\phi_1(\mathbf{e}) = \phi_2(\mathbf{e}) = 0$$

for all states $\mathbf{e} > (0,0)$.

These definitions are not exhaustive. The percentage of equilibria classified as aggressive is 96.88%, the percentage of equilibria classified as accommodative is 1.99%, and the percentage of unclassified equilibria is 1.13%. Our computations led always to a unique accommodative equilibrium but often to multiple aggressive equilibria at a given parameterization.

Our definitions of aggressive and accommodative equilibria map into worst, respectively, best equilibria. Table 6 shows that an aggressive equilibrium (if one exists) is the worst equilibrium in 98.1% of parameterizations while an accommodative equilibrium (if one exists) is the best equilibrium in 98.0% of parameterizations. Conversely, Table 7 shows that if there is a unique equilibrium, then it is classified as accommodative in 83.8% of parameterizations. If there are multiple equilibria, then the best equilibrium is classified as accommodative in 40.4% parameterizations and as aggressive in 43.2% parameterizations, However, the worst equilibrium is classified as aggressive in 97.9% of parameterizations. To facilitate the exposition and build intuition, we therefore identify the best equilibrium with an accommodative equilibrium and the worst equilibrium—to the extent that it differs from the best equilibrium—with an aggressive equilibrium. If the equilibrium is unique, then we identify it with an accommodative equilibrium.

References

Anderson, S., de Palma, A. & Thisse, J. (1992), *Discrete choice theory of product differentiation*, MIT Press, Cambridge.

Bergemann, D. & Välimäki, J. (2006), 'Dynamic price competition', *Journal of Economic Theory* **127**, 232–263.

Besanko, D., Doraszelski, U. & Kryukov, Y. (2014), 'Efficiency or predation: What drives pricing when there is learning-by-doing?', *American Economic Review* **104**(3), 868–897.

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

Besanko, D., Perry, M. & Spady, R. (1990), 'The logit model of monopolistic competition: Brand diversity', *Journal of Industrial Economics* **38**(4), 397–415.

Bolton, P. & Farrell, J. (1990), 'Decentralization, duplication, and delay', *Journal of Political Economy* **98**(4), 803–826.

Borkovsky, R., Doraszelski, U. & Kryukov, Y. (2010), 'A user's guide to solving dynamic stochastic games using the homotopy method', *Operations Research* **58**(4), 1116–1132.

Borkovsky, R., Doraszelski, U. & Kryukov, Y. (2012), 'A dynamic quality ladder model with entry and exit: Exploring the equilibrium correspondence using the homotopy method', *Quantitative Marketing and Economics* **10**(2), 197–229.

Bulow, J. & Klemperer, P. (1999), 'The generalized war of attrition', *Journal of Political Economy* **89**(1), 175–189.

Cabral, L. (2011), 'Dynamic price competition with network effects', *Review of Economic Studies* **78**, 83–111.

Cabral, L. & Riordan, M. (1994), 'The learning curve, market dominance, and predatory pricing', *Econometrica* **62**(5), 1115–1140.

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

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), 241–262.

Chen, J. (2011), How do switching costs affect market concentration and prices in network industries?, Working paper, UC Irvine, Irvine.

Chen, J., Doraszelski, U. & Harrington, J. (2009), 'Avoiding market dominance: Product compatibility in markets with network effects', *Rand Journal of Economics* **49**(3), 455–485.

Dasgupta, P. & Stiglitz, J. (1988), 'Learning-by-doing, market structure and industrial and trade policies', *Oxford Economics Papers* **40**(2), 246–268.

Dixit, A. & Stiglitz, J. (1977), 'Monopolistic competition and optimum product diversity', *American Economic Review* **67**(3), 297–308.

Doraszelski, U. & Escobar, J. (2010), 'A theory of regular Markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification', *Theoretical Economics* **5**, 369–402.

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

Dube, J., Hitsch, G. & Chintagunta, P. (2010), 'Tipping and concentration in markets with indirect network effects', *Journal of Marketing Research* **29**(2), 216–249.

Dube, J., Hitsch, G. & Rossi, P. (2009), 'Do switching costs make markets less competitive?', *Journal of Marketing Research* **46**, 435–445.

Farrell, J. & Klemperer, P. (2007), Coordination and lock-in: Competition with switching costs and network effects, *in* M. Armstrong & R. Porter, eds, 'Handbook of Industrial Organization', Vol. 3, North-Holland, Amsterdam, pp. 1967–2072.

Harberger, A. (1954), 'Monopoly and resource allocation', *American Economic Review* **44**(2), 77–87.

Koenker, R. & Perry, M. (1981), 'Product differentiation, monopolistic competition, and public policy', *Bell Journal of Economics* **12**, 217–231.

Kulkarni, V. (1995), *Modeling and analysis of stochastic systems*, Chapman and Hall, London.

Maynard Smith, J. (1974), 'Theory of games and the evolution of animal conflicts', *Journal of Theoretical Biology* **47**, 209–221.

Mermelstein, B., Nocke, V., Satterthwaite, M. & Whinston, M. (2014), Internal versus external growth in industries with scale economies: A computational model of optimal merger policy, Working paper, Northwestern University, Evanston.

Mitchell, M. & Skrzypacz, A. (2006), 'Network effects and long-run market shares', *Economic Theory* **29**(3), 621–648.

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.

Segal, I. & Whinston, M. (2007), 'Antitrust in innovative industries', *American Economic Review* **97**(5), 1703–1730.

Simon, C. & Blume, L. (1994), *Mathematics for economists*, Norton, New York.

Tirole, J. (1988), *The theory of industrial organization*, MIT Press, Cambridge.