Disruption Costs, Learning by Doing, and Technology Adoption*

Carlos J. Pérez† Carlos J. Ponce‡

Abstract

We study technology adoption in a dynamic model of price competition. Adoption involves disruption costs and learning by doing. Because of disruption costs, the adopting firm begins in a market disadvantage, which may persist if its rival captures the customers that the adopting firm needs to learn the technology. The prospect of future rents by the rival results in: (i) A failure to adopt socially efficient technologies; (ii) An equilibrium preference for technologies that are learned faster but have lower social value; (iii) More technologies being adopted if more firms enter the market.

JEL: L1, O3, D4.

Keywords: Technology adoption; adoption breakdowns; dynamic Bertrand competition; Bertrand sum; discounted Bertrand sum; endogenous impatience.

1. Introduction

The adoption of new technologies is the leading force behind productivity growth in many industries. Still, when a new technology is adopted things often go wrong in the beginning. Adopting firms face mayor adaptation problems and become temporarily less productive than non-adopters. Such problems, known as switchover disruption costs, may be overcome through learning by doing as firms accumulate experience using the new technology. Indeed, the idea that productivity growth first falls and later rises after the adoption of a new technology is supported by the micro-evidence according to Huggett and Ospina (2001).

*Pérez acknowledges support from CONICYT under the FONDECYT 2014 Postdoctoral Grant No. 3140360. Ponce acknowledges support from CONICYT under the FONDECYT 2014 Regular Grant No. 1140582.

†Escuela de Administración, Pontificia Universidad Católica de Chile, Avenida Vicuña Mackenna 4860, 7820436 Macul, Chile; cjperez.econ@gmail.com; +56 (2) 2354 4360. (Corresponding author.)

‡Universidad Alberto Hurtado, Erasmo Escala 1835, 8340539 Santiago, Chile; cjponcebruera@gmail.com; +56 (2) 2889 7765.

The relevance of disruption costs in the introduction of products and processes is well-known in the management literature. Christensen (1997) provides ample evidence of disruptive technologies that result in worse product performance in the near-term. Tyre and Hauptman (1992) list as sources of disruption costs the novelty of technical features, the low applicability of previous knowledge, and the incompatibility of current organizational practices with the arriving innovation. Leonard-Barton (1988) shows that the adaptation to a new technology often requires active cooperation between users and developers. In the economics literature, Holmes et al. (2012) provide an excellent discussion of the importance of disruption costs in a number of innovation episodes.

Likewise, the business strategy consulting industry has found extensive empirical evidence of experience curves (Henderson, 1968). This evidence has had a profound impact on strategic management research and practice. Indeed, the *Harvard Business Review* listed the experience curve as 'one of the five charts that changed the world (Ovans, 2011).' Economists have also studied the empirical evidence on, and the theoretical implications of, learning by doing since the seminal works of Wright (1936) and Arrow (1962a). A review of both strands of the literature may be found in Thompson (2010).

The interplay between disruption costs and learning by doing in strategic settings has been mostly overlooked in the literature. This is unfortunate because many industrial innovations take place in oligopolistic environments where strategic issues play a leading role. In this article we develop a simple model to study technology adoption in a strategic setting with disruption costs and learning by doing.

Our main observation is that non-adopting firms have incentives to undercut prices to prevent the learning of the new technology because this makes the adopting firm a weaker future competitor. The prospect of future rents by non-adopting firms places a pecuniary cost on the adopting firm that, in some cases, renders the adoption of Pareto superior technologies unprofitable. In other words, as 'stealing' current customers from the adopting firm creates future rents without adding any social value, the know-how needed to learn the new technology becomes an artificially overpriced 'asset' in the market. This overpricing renders adoption unprofitable.

We study these issues in a dynamic duopoly model of Bertrand competition in which the adopting firm has a limited amount of time to learn the new technology. This time limit may come from the threat of imitation, the expiration of a patent, etc. In the model, firms offer potentially differentiated products to a sequence of short-lived customers with unit demand. The main advantage of this setting with respect to others, e.g., a Cournot model of competition, is that it isolates the dynamics of adoption by assuming away static equilibrium distortions.

Within this framework, we obtain three main results. First, we confirm that, in some cases, the adopting firm prefers to stick to an old technology rather than to switch to a better one. Second, we show that, for the cases of interest, between two technologies with the same social value, the adopting firm prefers the technology whose flow payoffs are received earlier. This equilibrium bias towards technologies with larger early rewards is called the impatience property. As a corollary, we prove that the bias embedded in the impatience property favors the adoption of technologies that are learned faster but have lower social value. Third, we show that adoption is made easier if more firms enter the market. More precisely, we prove that adding non-adopting firms to our model enlarges the set of (efficient) technologies that are adopted in equilibrium. Taken together, our results should warn regulators of keeping an eye on industries either with few competitors or where technological improvements take longer to settle. In our view, these are the industries in which disruption costs and learning by doing raise a strategic barrier to efficient adoption and productivity growth.

Holmes et al. (2012) also study adoption in the presence of switchover disruptions. Using an Arrow-type model, they show that a more competitive environment favors adoption as the cost of adopting a technology is the forgone profits during the disruption period.[1] Our insight is different as we stress that disruption costs open a future profit opportunity to competing firms. In line with a large body of evidence (see Holmes and Schmitz, 2010), we also show that additional competition may promote adoption. But while in Holmes et al.'s article competition is beneficial because it reduces the forgone profits of the adopting firm, in our case it is so because it limits the spurious rents of non-adopters. Our mechanism thus offers a novel channel through which extra competition facilitates the adoption of a new technology.

In the industrial organization literature, dynamic price competition and learning by doing have been explored by Cabral and Riordan (1994) and, more recently, by Besanko et al. (2010) and Besanko et al. (2014). The goal of these articles is to understand how learning by doing, jointly with organizational forgetting in Besanko et al. (2010), determines pricing and market dominance in a duopolistic setting. Schivardi and Schneider (2008) examine a dynamic investment game with learning and disruptive adoption. Their analysis, however, resembles a multi-stage patent race in which the adopting firm learns the potential of a new technology in a Bayesian fashion.

Our work is also related to a list of macro articles in which learning and disruption costs are at the center of the stage. In perfectly competitive environments, Chari and Hopenhayn

[1] Arrow (1962b) was the first to compare adoption incentives under perfect competition and monopoly. However, in Arrow's article and in the literature that follows, for example Gilbert and Newbery (1982), there is neither learning by doing or disruption costs.

(1991) and Parente (1994) examine adoption when the implementation of a technology entails losing previously acquired knowledge. Jovanovic and Nyarko (1996) add to this literature by studying the full dynamics of technology adoption in a one-agent Bayesian model of learning by doing. Klenow (1998) examines a firm's decision of when to update a process technology. In contrast to these articles, we study adoption in a strategic setting and show that market structure is crucial for the adoption of new technologies. This is a key distinctive feature of our work.

The remainder of this article is organized as follows. Section 2 presents the model. Section 3 provides a stripped-down, illustrative example of an adoption breakdown. Sections 4 and 5 introduce some useful concepts and preliminary findings. Section 6 presents our main results. Section 7 concludes. Proofs are collected in the Appendix.

2. The Model

We present a simple, canonical model of learning by doing and technology adoption in the spirit of Cabral and Riordan, 1994, Besanko et al., 2010, and Besanko et al., 2014.

The industry. Consider an industry with two firms denoted by $i \in \{1, 2\}$ and a finite number of $T + 1$ customers with unit demand. Sales take place over time: at each period only one customer is available to buy from the firms. Time is denoted by $t \in \mathrm{T} := \{0, ..., T\}$ and, without loss of generality, firms do not discount the future. Firms start with a baseline technology that allows them to produce at a cost c_0 a unit of a product that customers value at v_0.[2] With $s_0 := v_0 - c_0$ we denote the constant, positive flow (per-period) surplus that is created every time a firm sells to a customer using the baseline technology.

Technology adoption. To ease the exposition, we assume that only firm 1 can adopt a new technology. The new technology may bring either product quality improvements or cost savings, and it is described by the flow surplus it creates at each sale by firm 1. This flow surplus, in turn, depends on the state of the technology via the formula

$$s(x_t) = v(x_t) - c(x_t),$$

where x_t is the stock of know-how (cumulative experience) in using the technology at the beginning of period t, and v and c are the new technology's value and cost. By making a sale, firm 1 adds to its stock of know-how. Hence, the evolution of firm 1's stock of know-how

[2] Our results also hold true for asymmetric values and costs.

is controlled by the law of motion

$$x_{t+1} = x_t + y_t, \quad x_0 \geq 0 \quad \text{given},$$

where $y_t \in \{0, 1\}$ indicates whether firm 1 sells at period t and x_0 is the stock of know-how at the moment of adoption. When firm 1 makes a sale, i.e., when $y_t = 1$, it gets an additional unit of know-how through *learning by doing*. To ease notation, we set the initial stock of know-how x_0 to zero. Hence, the stock of know-how up to the beginning of period t equals *cumulative sales* with the new technology up to the beginning of that period:

$$x_t = \sum_{k=0}^{t-1} y_k,$$

where $x_t \leq t$ (owing to the fact that there cannot be more sales than periods). Henceforth, we write the flow surplus that firm 1 creates if it sells at period t simply as $s(x_t)$. Because the flow surplus changes in time only if sales do, hereafter we abuse notation and denote cumulative sales with the new technology by x; that is, $x := x_t$. The new technology fulfills:

$$s(0) \leq s_0, \tag{A1}$$

$$s(x+1) \geq s(x). \tag{A2}$$

The first assumption captures the idea of *switchover disruption costs*: the first sale made with the new technology creates a weakly smaller flow surplus than the one created by the baseline technology (see Tyre and Hauptman, 1992; Leonard-Barton, 1988; Holmes et al., 2012; Benkard, 2000). We say there are disruption costs if inequality A1 is strict, and their magnitude is given by $s_0 - s(0)$. The second inequality says that selling an additional unit increases the flow surplus that the new technology creates trough learning by doing—either by raising the value of the product or by saving production costs. We further assume that adoption is socially efficient. Formally, there is a minimum number of cumulative sales $q \leq T + 1$ such that the social value of the new technology is non-negative:

$$\sum_{x=0}^{q-1} s(x) \geq q s_0 + \varepsilon, \tag{A3}$$

where $\varepsilon > 0$ is a sunk cost incurred at adoption. The left-hand side of A3 is the surplus from selling q units with the new technology. Similarly, the right-hand side, which measures the opportunity cost of adopting the technology, is the surplus from selling q units with the baseline technology plus the sunk cost. The sunk cost must be distinguished from disruption costs: the former is a fixed cost while the latter are the initial losses in the efficiency of firm 1 (due, for instance, to higher production costs). We aim to understand adoption decisions for a whole class of technologies: the set S of all technologies s satisfying Assumptions A1–A3.[3] Note that, for social efficiency, every technology in set S should be adopted from the outset and firm 1 should sell at each period using the adopted technology.

Actions, payoffs, and equilibrium. Adoption is decided within the framework of the following extensive-form game: At the beginning of each period, firm 1 chooses whether or not to replace the baseline technology s_0 with a given technology s in S.[4] The choice is irreversible. Then, firms simultaneously announce their flow surplus offers to the current customer: $b := v - p$, where p is the price. (We work with flow surplus offers instead of prices to simplify notation and presentation.) The customer, in turn, decides whether or not to buy from one of the firms. The game continues this way until the last period and firms have complete and perfect information.

Payoffs are as follows: At each period the current customer obtains either a zero surplus if he does not buy or the surplus b offered by his trading partner. Likewise, each firm receives nothing if it does not sell or a flow payoff equal to $s - b = p - c$ if it does. The payoff of each firm equals the sum of its flow payoffs.

Our solution concept is pure-strategy Markov perfect equilibrium (MPE, for short) with the pair (x, t) as state variable. A MPE is a subgame perfect equilibrium in which the surplus-offer (pricing) strategies of the firms depend only on the state variable.

Vocabulary. We refer to continuation payoffs simply as *payoffs*. Likewise, we call per-period payoffs *flow payoffs*. We use the same wording for surpluses and social values. We also distinguish between *pre-adoption payoffs* that correspond to the no-adoption decision and *post-adoption payoffs*. Lastly, though our model encompasses quality improvements, much of our discussion is in terms of cost-saving technologies because some may interpret higher surpluses more easily as lower costs.

3. An Example

We begin our study by providing a stripped-down example where an adoption breakdown occurs in equilibrium.

Example 1. Consider a two-period model in which customers value at 1 the product of either technology. The cost of producing with the baseline technology is $c_0 = 0.5$, whereas

[3] From a formal viewpoint, set S is thus parametrized by the triple (T, s_0, ε).

[4] We assume that firm 1 adopts s if it is indifferent.

the new technology allows firm 1 to produce a second unit at a cost 0.1 after producing the first unit at a cost 0.75. Assumptions A1 and A2 are satisfied because

$$s(0) = 0.25 < 0.5 = s_0 < 0.9 = s(1).$$

Assumption A3 is also met for any $\varepsilon \leq 0.15$ as

$$s(0) + s(1) = 1.15 \geq 1 + \varepsilon = 2s_0 + \varepsilon.$$

Suppose that firm 1 adopts the new technology and sells to the first costumer. Then, firm 1 would also sell to the second customer and get a second-period equilibrium payoff equal to $c_0 - c(1) = 0.5 - 0.1 = 0.4$.[5] Therefore, its payoff in the dynamic competition game would be

$$\underbrace{p - 0.75}_{\substack{\text{Current}\\ \text{payoff}}} + \underbrace{[0.4]}_{\substack{\text{Future}\\ \text{payoff}}}, \tag{1}$$

where p is the price at which firm 1 sells in the first period. This price must satisfy

$$1 - p \geq b_2,$$

where $1 - p$ is the utility that the first costumer gets if he buys from firm 1 and b_2 is the maximum utility that firm 2 would be willing to offer to the first customer.

We now obtain b_2. Firm 2 can offer the first customer a utility equal to s_0. But, in addition, if firm 2 sells to the first costumer, it would prevent the learning of the technology and keep firm 1's cost at 0.75 in the second period. Firm 2 could in this case sell also to the second customer and get a second-period equilibrium payoff equal to the disruption cost $s_0 - s(0) = 0.5 - 0.25 = 0.25$.[6] In short:

$$b_2 = \underbrace{s_0}_{\substack{\text{Flow}\\ \text{surplus}}} + \underbrace{[s_0 - s(0)]}_{\substack{\text{Future}\\ \text{rent}}} = 0.75.$$

As a result, firm 1 may charge the first customer a price (at most) equal to

$$\overline{p} = c_0 - [s_0 - s(0)] = 0.25.$$

Plugging this (maximum) price into (1) we see that firm 1 would incur in a loss equal to

[5] Hereafter we consider only equilibria in which the customer buys from the most efficient firm if they ask for the same price. This tie-breaking rule is standard (see Deneckere and Kovenock, 1996).

[6] Or, equivalently, $c(0) - c_0 = 0.75 - 0.5 = 0.25$.

0.1 in order to sell both units. Firm 1 would thus refrain from selling—and get a zero post-adoption payoff. Consequently, we conclude that there would not be adoption in equilibrium for any sunk cost $\varepsilon > 0$.

4. Definitions

The definitions and discussion of this section are key to follow the results below. With $\pi(x) := s(x) - s_0$ we denote the flow social value at state (x,t) of the new technology with respect to the old one. Note that, because of disruption costs, there is a maximum number of cumulative sales $\hat{x}$ such that $\pi(x) < 0$ for $0 \leq x < \hat{x}$. In words, the new technology is less productive than the baseline technology until the former accrues $\hat{x}$ sales. Also, observe that the positive (negative) part of the flow social value function $\pi^+(x)$ ($\pi^-(x)$) equals the equilibrium payoff that firm 1 (firm 2) would get in a *static* Bertrand game (the positive and negative parts of a function f are defined as $f^+(a) := \max\{f(a), 0\}$ and $f^-(a) := -\min\{f(a), 0\}$ for any element a in the domain of f). More concretely, let (x,t) be a state at which cumulative sales x are below threshold $\hat{x}$. If a static Bertrand game were played at this state, firm 2 would sell to the available customer and obtain a payoff equal to $-\pi(x)$. A parallel observation applies to any state (x,t) with cumulative sales x at or above threshold $\hat{x}$. Lastly, recall that, due to learning by doing, the flow social value $\pi(x)$ is non-decreasing in cumulative sales x. We display flow social values at all feasible states in the *triangular array* A:

$$
\text{Sales} \quad
\begin{array}{ccccc}
\pi(0) & \pi(0) & \dots & \pi(0) & \pi(0) \\
 & \pi(1) & \pi(1) & \dots & \pi(1) \\
 & & \ddots & \vdots & \vdots \\
 & & & \pi(T-1) & \pi(T-1) \\
 & & & & \pi(T),
\end{array}
$$

$$\text{Time}$$

where rows count cumulative sales and columns count periods. For any state (x,t), let $A_{x,t}$ be the subarray with first entry in row $x+1$ and column $t+1$, and last entry in row $T+1+x-t$ and column $T+1$. This subarray stands for the subgame that begins at (x,t).

Example 2. Consider a new technology with associated triangular array A:

$$
\begin{array}{ccc}
-0.75 & -0.75 & -0.75 \\
 & 0 & 0 \\
 & & 2.
\end{array}
\tag{2}
$$

Note that $\pi(x) < 0$ for $x = 0$ and $\pi(x) \geq 0$ for $x \geq 1$, so $\hat{x} = 1$. The reader may think of $\pi(0)$ as the initial cost disadvantage of the new technology with respect to the old one (i.e., as $c_0 - c(0)$). The flow social value of this new technology becomes positive after two consecutive sales as a result of accumulated experience. If in the first period firm 2 sells one unit, the game then shifts to the subgame that corresponds to the subarray $A_{0,1}$:

$$\begin{array}{cc} -0.75 & -0.75 \\ & 0, \end{array}$$

which begins at row 1, column 2 and ends at row 2, column 3 of (2). In this subarray $\pi(0) = -0.75$ indicates that if a static Bertrand game were played at any state of the form $(0, t)$ firm 2 would sell to the customer and obtain a payoff equal to $\pi^{-}(0) = 0.75$.

The advantage of arranging our data in a triangular array will become apparent below. With that in mind, let us introduce the following concepts. First, from the definition of the flow social value $\pi(\cdot)$ it ensues that the social value of using the technology from state (x, t) onward is

$$d(x,t) := \sum_{k=0}^{T-t} \pi(x+k), \tag{3}$$

which obtains as the summation over the main diagonal of subarray $A_{x,t}$. This social value is the payoff that firm 1 would get if firm 2 played its static Bertrand best-response from state (x, t) onward. Accordingly, we refer to (3) as firm 1's *Bertrand sum*. Second, let cumulative sales x be below threshold $\hat{x}$, i.e., let $\pi(x) < 0$. Then, the payoff that firm 2 would get if firm 1 played its static Bertrand best-response from state (x, t) onward is

$$r(x,t) := -(T+1-t)\pi(x), \tag{4}$$

which obtains as the negative of the summation over the first row of $A_{x,t}$. We refer to (4) as firm 2's *Bertrand sum*. Third, the summation of flow social values at *all* states from state (x, t) onward is

$$z(x,t) := \sum_{k=0}^{T-t} \big[T+1-(t+k)\big]\pi(x+k), \tag{5}$$

which obtains as the summation of *all* the entries of $A_{x,t}$. A nice interpretation of (5) comes out of looking at it as a discounted combination of static Bertrand payoffs:[7]

$$z(x,t) = \frac{1}{\delta_{T-t,t}} \sum_{k=0}^{T-t} \delta_{k,t}\, \pi(x+k), \tag{6}$$

where $\delta_{k,t} := 1-(T+1-t)^{-1}k$ may be read as a discount factor.[8] Correspondingly, we refer to it as the discounted Bertrand sum. This sum may contain positive as well as negative terms. More concretely, static Bertrand payoffs of firm 1 enter (6) as nonnegative numbers, whereas those of firm 2 enter as nonpositive numbers. The discounted Bertrand sum is a rough measure of the strategic position of firm 1 against firm 2.

Example 3. For the triangular array (2) of Example 2, it is easy to see that $d(0,0) = 1.25$, $r(0,0) = 2.25$, and $z(0,0) = -0.25$. Also, $d(0,1) = -0.75$, $r(0,1) = 1.5$, and $z(0,1) = -1.5$.

5. Preliminary Results

We begin this section studying the dynamic competition subgame that follows once a new technology has been adopted. The following theorem states existence, uniqueness, and characterizes MPE payoffs.[9] (Recall the definitions of f^+ and f^- for a function f.)

Theorem 1 (Equilibrium payoffs)**.** *There is a unique MPE of the subgame that follows adoption. In this equilibrium the payoffs of firms 1 and 2 are:*

$$\pi_1(x,t) = \min\{d^+(x,t), z^+(x,t)\}, \tag{7}$$

$$\pi_2(x,t) = \min\{r^+(x,t), z^-(x,t)\}. \tag{8}$$

Theorem 1 shows that it suffices to compute the discounted Bertrand sum to single out the selling firm at each state of the game. For instance, the array (2) of Example 2 has $z(0,0) = -0.25$. Hence, $z^+(0,0) = 0$ and $z^-(0,0) = 0.25$, meaning that firm 2 sells at the first period. Then, because $r^+(0,0) = 2.25 > z^-(0,0)$, firm 2 gets the payoff $\pi_2(0,0) = 0.25$ in equilibrium. Theorem 1 also constrains payoffs as follows: First, it says that payoffs are nonnegative. This is intuitive because firms can always ask for prices that customers would not accept. Second, it says that the payoff of a selling firm is the minimum between its Bertrand sum and the discounted Bertrand sum. That is, it says that the most a firm can make at any state is what it would make if its rival played its static Bertrand best-response from that state onward.

[7] Note that (6) results from dividing and multiplying (5) by $(T+1-t)$.

[8] For a given period t, $\delta_{k,t}$ is a function of k. Thus $\delta_{T-t,t}$ corresponds to $k = T-t$.

[9] We focus on MPE in weakly dominant strategies.

Corollary 1 (Monotonicity)**.** *In the unique MPE of the subgame that follows adoption, if a firm sells at period t, then the same firm sells at all subsequent periods.*

This result says that market leadership persists over time. The intuition is simple. Because of learning by doing, the flow surplus of firm 1 increases each time it makes a sale and remains constant each time it misses a sale. Consequently, if firm 1 sells (does not sell) in the current period it will also (neither) sell in the next period because its strategic position will then be stronger (weaker). Corollary 1 and Theorem 1 together yield the next lemma.

Lemma 1 (No delay)**.** *In the unique MPE, adoption takes place only at the initial period.*

Lemma 1 says that adoption is never delayed. Put differently, if firm 1 finds adoption profitable at some period t, it will also find it profitable at an earlier period. The reason is that every time firm 1 delays adoption one period it gives away its most profitable selling opportunity. For instance, by delaying adoption from the first to the second period, firm 1 misses the sale that it could make in the last period at the end of its learning path.

Corollary 1 and Lemma 1 allow us to conclude that adopted technologies are fully learned in equilibrium. As a consequence, the equilibrium outcome is efficient if, and only if, the set S^* of technologies that are adopted in equilibrium coincides with the set S of socially efficient technologies. Because Corollary 1, Lemma 1, and Theorem 1 imply that $\pi_1(0,0)$ is the equilibrium post-adoption payoff of firm 1 and its pre-adoption payoff is zero (since before adoption firms compete a la Bertrand, sell homogeneous products, and have identical costs), the equilibrium outcome is efficient if, and only if, $\pi_1(0,0) \geq \varepsilon$ for *all* s in S. Our next result identifies a sufficient condition for efficiency.

Proposition 1 (Efficient adoption)**.** *With zero switchover disruption costs, the equilibrium outcome is efficient.*

For simplicity, we confine the discussion of Proposition 1 to a cost-saving technology. Without disruption costs, firm 1 is at least as efficient as firm 2 from the outset. Further, because of learning by doing, firm 2's cost will always be higher than firm 1's cost. That is, while firm 2 could sell by pricing below cost, it cannot hope to recoup that loss by pricing above cost in the future. Firm 2 will then play its static Bertrand best response at each period—equating price to cost. As a result, firm 1's adoption payoff equals its Bertrand sum minus the sunk cost—the social value of the technology—and efficiency ensues.

The formal, general argument is as follows. Without disruption costs, the flow social value $\pi(\cdot)$ is nonnegative at *all* states. So, the discounted Bertrand sum is weakly larger than firm 1's Bertrand sum $z(0,0) \geq d(0,0)$. Efficiency ensues because Theorem 1 implies that $\pi_1(0,0) = d(0,0)$ and Assumption (A3) guarantees $\varepsilon \leq d(0,0)$ for all s in S.

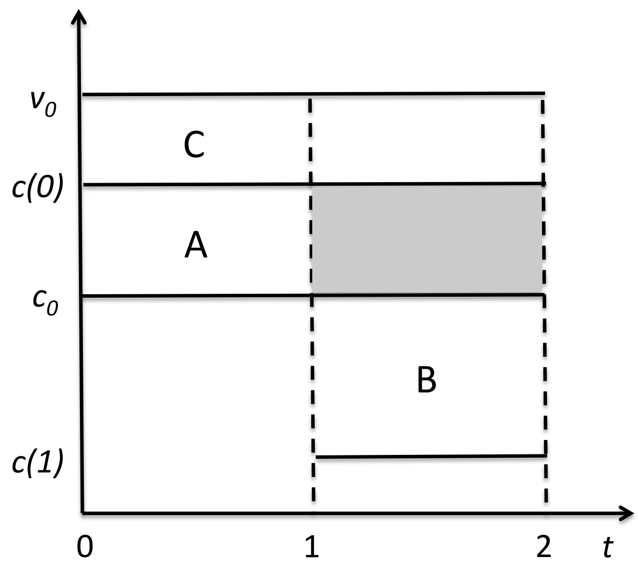

Figure 1: Two-period model with a cost-saving technology.

6. Main Results

Within this section we assume that disruption costs are positive. Our key findings come in three groups: First, we show that some socially efficient technologies are not adopted. Second, we describe the most distinctive features of these technologies. Third, we study the effect of adding more firms on technology adoption.

6.A. Adoption Breakdowns

To ease the exposition, we define the set N of technologies

$$\mathrm{N} := \{s \in \mathrm{S} : z(0,0) < \varepsilon\}.$$

This set contains any technology in S whose discounted Bertrand sum at $t = 0$ is smaller than the sunk cost. Set N, which is clearly non-empty, is a subset of S, the set of socially efficient technologies. Our first key result is Proposition 2.

Proposition 2 (Adoption breakdowns)**.** *In the presence of disruption costs, the equilibrium outcome is inefficient:*

$$\mathrm{S}^* = \mathrm{S} - \mathrm{N}.$$

In words, Proposition 2 says that any technology in set N will not be adopted in equilibrium. The economic intuition behind this proposition is most easily seen within a two-period model with a cost-saving technology like the one displayed in Figure 1. The total surplus that can be created with the baseline technology is twice the sum of the areas A and C in the figure—one for each period. On the other hand, the new technology can create a flow surplus equal to C in the first period and to C plus A plus B in the second period. Consequently, the new technology is weakly preferred to the baseline technology from a social viewpoint if, and only if,

$$B \geq A + \varepsilon. \tag{9}$$

That is, if assumption (A3) holds.

As we know from Corollary 1, and it is shown in Figure 1, if a firm sells to the first customer then a cost advantage will let it sell to the second customer as well. This means that each firm is willing to offer the first costumer not only its first-period flow surplus; C for firm 1 and C plus A for firm 2, but also its potential second-period payoff. The potential second-period payoff of firm 1 is B, so it is willing to offer the first customer C plus B. Likewise, the potential second-period payoff of firm 2 is given by the shaded area of size A in Figure 1. Moving backwards, this means that firm 2 is willing to offer the first customer C plus two times A. As a result, the new technology will be adopted if, and only if,

$$B \geq 2A + \varepsilon.$$

Conversely, adoption will *not* take place if, and only if, the discounted Bertrand sum at $t = 0$ fulfills

$$z(0,0) = 2\underbrace{[C - (A + C)]}_{\pi(0)} + \underbrace{(C + A + B) - (A + C)}_{\pi(1)} = -2A + B < \varepsilon.$$

That is, if, and only if, the new technology s is in set N, as we aimed to show. Because this inequality may hold even if efficiency condition (9) is satisfied, Proposition 2 states that the equilibrium outcome is inefficient.

The cause of adoption breakdowns is, therefore, that firm 2 may appropriate the shaded area in Figure 1 by precluding the learning of the new technology. The key economic intuition behind Proposition 2 is that equilibrium adoption does not merely require the social value of the technology to be nonnegative, it also requires this social value to be sufficiently high to compensate for the potential second-period payoff of firm 2 represented by the shaded area in Figure 1.

The interpretation of Proposition 2 in the general model, although similar to that of the two-period, cost-saving model, is more involved and takes a recursive formulation. We have seen that our game is not just a sequence of independent one-shot Bertrand games because each sale made by firm 1 improves his strategic position against firm 2 and vice versa. We also have seen that, for this reason, both firms are willing to offer more than their flow surplus to the current customer. How much they offer will depend, of course, on the return they expect from the improved strategic position. The next proposition shows that, in order to sell, each firm must offer each customer its rival's potential payoff from an extra sale.

Proposition 3 (Recursive payoffs)**.** *In the unique MPE, the post-adoption payoffs of firms 1 and 2 can be written recursively as:*

$$\pi_1(x,t) = \max\left\{ d(x,t) - \sum_{k=0}^{T-(t+1)} \pi_2(x+k, t+1+k), 0 \right\},$$
$$\pi_2(x,t) = \max\left\{ r(x,t) - \sum_{k=0}^{T-(t+1)} \pi_1(x+1, t+1+k), 0 \right\},$$

where $\pi_i(\cdot, T) = \max\{(-1)^{i-1}\pi(\cdot), 0\}$ *for* $i \in \{1,2\}$.

To understand this proposition, suppose that firm 1 adopts the new technology and sells at every period. If firm 2 always played its static Bertrand best response, firm 1 would get its Bertrand sum $d(t,t)$ at each $t \in T$ as put forward in section 4. However, firm 2 may offer more than its flow surplus s_0 to some customers. Actually, it would be willing to offer

$$u_2(t,t) = s_0 + \pi_2(t, t+1). \tag{10}$$

That is, it would be willing to offer not only its flow surplus, but also the continuation payoff that it would earn if it sold to the current customer. Proposition 3 says that firm 1 must transfer each customer a flow surplus equal to $u_2(t,t)$ to move forward along its learning path. As a result, the payoff of firm 1 is given by its Bertrand sum minus the sum of transfers it must grant to the customers at each succeeding period. If the total amount to be transferred plus the sunk cost of adoption is greater than firm 1's Betrand sum at $t = 0$, i.e. if $\pi_1(0,0) < \varepsilon$, then the new technology will not be adopted. That is, anticipating the pricing strategy of firm 2, firm 1 will not adopt the new technology in response to a credible *threat* rather than to an actual price undercutting.

A 'dual' interpretation of the model and results so far is as follows. As 'stealing' customers from firm 1 slows down the learning of the new technology and improves the future strategic position of firm 2, experience becomes an artificially overpriced 'asset' in the market. In the absence of disruption costs, know-how 'prices' are undistorted and the equilibrium outcome is efficient. However, in the presence of disruption costs, know-how is overpriced because of the spurious rents that firm 2 may appropriate. It is ultimately this overpricing what leads firm 1 to desist from adopting in the first place.

How can we tell adopted from non-adopted technologies apart? One may think, as our previous result also suggests, that new technologies are adopted or not according to their social value. Although this is not true in general, we show below that technologies with a sufficiently high social value are indeed adopted. To state this result formally, let us define the set G of technologies as

$$\mathrm{G} := \{s \in \mathrm{S} : d(0,0) \geq \varepsilon + K\},$$

where $K := \frac{1}{2}T(T+1)|\pi(0)|$. We have the following proposition.

Proposition 4. *In the unique MPE, every technology in set* G *is adopted.*

Proposition 4 just says that firm 2 cannot prevent the learning of extremely productive technologies. For technologies outside set G, however, the social value rule is insufficient to decide whether adoption will take place in equilibrium. In particular, the inter-temporal distribution of the social value of these technologies becomes a key factor as those technologies which are learned faster deliver higher adoption payoffs. The next example illustrates this point.

Example 4. Consider two technologies, s and s', with triangular arrays:

$$\begin{array}{ccc} -0.75 & -0.75 & -0.75 \\ & 0.75 & 0.75 \\ & & 1.25, \end{array} \qquad \begin{array}{ccc} -0.75 & -0.75 & -0.75 \\ & 0 & 0 \\ & & 2, \end{array}$$

and identical sunk costs $\varepsilon = \varepsilon' < 0.5$. Even though both technologies have the same social value $d(0,0) - \varepsilon = d'(0,0) - \varepsilon' = 1.25 - \varepsilon$ and disruption costs $-\pi(0) = -\pi'(0) = 0.75$, it follows from Theorem 1 that $\pi_1(0,0) = 0.5$, whereas $\pi'_1(0,0) = 0$.[10] Hence, technology s is adopted but technology s' is not.

To state this finding formally we need a definition that captures the idea of a technology delivering larger early flow social values than another. Because the desired payoff comparisons can only be made between technologies with equal social values and disruption costs, we say that technologies s and s' are *equivalent*, and denote this relation by $s \sim s'$, if, and only if,

$$s, s' \in \mathrm{S}, \; d(0,0) = d'(0,0), \text{ and } \pi(0) = \pi'(0).$$

[10] Recall that the size of the disruption cost is $s_0 - s(0) = -\pi(0)$.

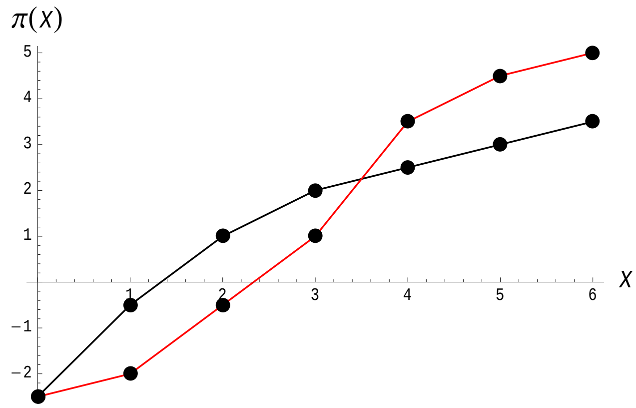

Figure 2: Two equivalent technologies: technology s' (red) yields a zero adoption payoff; technology s (black) is learned earlier and yields a positive adoption payoff.

Definition 1. Let $s \sim s'$. Then, technology s is *learned earlier* than technology s' if both $\pi(1) \geq \pi'(1), ..., \pi(k) \geq \pi'(k)$ and $\pi(k+1) \leq \pi'(k+1), ..., \pi(T) \leq \pi'(T)$ hold with at least one strict inequality.

The idea behind Definition 1 is very simple. First, we compare only equivalent technologies. Then, we say that technology s is learned earlier than technology s' if during the first k sales the former delivers higher flow social values than the latter. As both technologies have the same social value, technology s' must deliver higher flow social values during the last $T-k$ sales. Geometrically, technology s is learned earlier than technology s' if s' crosses s at most once from below (see Figure 2).[11]

Proposition 5 (The impatience property)**.** *Let technology s be learned earlier than technology s'. Then, firm 1's adoption payoff with technology s is weakly higher than its adoption payoff with technology s'.*

The result shows that firm 2 is less willing to price below cost against new technologies that are learned earlier (alternatively, know-how is less overpriced for these technologies). The proposition generalizes the following intuitive argument. Consider a technology and perturb it by shifting a unit of social value from the last period to the first period. (Clearly, the perturbed tecnology is learned earlier than the original.) The cost of this perturbation is to reduce the last-period social value in one unit—a *static* cost of one unit. The benefit is, however, twofold: First, it increases the first-period flow social value in one unit. This *static* benefit exactly compensates the static cost, leaving the social value of the technology unchanged. Second, and key to the result, it increases the price at which firm 1 can sell to the first costumer. This *dynamic* benefit ensues because, as firm 1 becomes more efficient

[11]Ignoring the point $(0, \pi(0))$.

at the initial period, the continuation payoff that firm 2 would earn from selling to the first customer diminishes.

There is indeed a more general illustration of the impatience embedded in equilibrium payoffs. The idea suggests itself from the discounted Bertrand sum. For this purpose, we write

$$z(0,0) = \frac{1}{\delta_{T,0}} \sum_{k=0}^{T} \delta_{k,0}\, \pi(k), \tag{11}$$

where $\delta_{k,0} = 1 - (T+1)^{-1}k$ is read as an *endogenously determined* discount factor. The rule says that a technology is adopted if, and only if, the discounted value in (11) is weakly larger than ε. However, efficiency is characterized by the (non-discounted) Bertrand sum

$$d(0,0) = \sum_{k=0}^{T} \pi(k)$$

being weakly larger than ε (see (3)). As we have seen, this endogenous discounting has two consequences: First, a positive social value is not enough for adoption; second, technologies whose flow surpluses are delivered earlier have more chances of being adopted. A third consequence can be glimpsed from Proposition 5.

Proposition 6 (Inefficient choice)**.** *For any technology s in* N *with positive social value, there is another technology s' in* S^* *with a smaller social value.*

The proposition says that, if firm 1 could choose between technologies, the present bias embedded in the impatience property would favor the adoption of technologies with smaller but 'better' inter-temporally distributed social value. There is indeed a trade-off between the social value of a technology and its temporal distribution. That is, firm 1 would be willing to sacrifice efficiency in exchange for higher early flow surpluses.

6.C. The Value of Increasing Competition

Starting with Schumpeter (1942) there is a long-standing debate in economics as to whether competition is beneficial or detrimental to technology adoption. Moreover, to the extent that the degree of competition may be influenced by policy measures, it is important to understand the mechanism by which its variation affects adoption decisions. These are very important issues in industrial organization to which our model brings some novel insights.

We address these matters by adding a third, non-adopting firm with constant, positive flow surplus s_3 to our model, but the reader will be convinced that adding two or more firms would not affect our conclusions. In order to make the appropriate comparisons across models, we need every technology s in S to remain Pareto efficient in the three-firm model (in particular, inequality A3 must hold for every s in S if s_0 is replaced by s_3). Consequently, we require $s_3 \leq s_0$. This implies that firm 1 will continue to face competitive pressure *only* from firm 2, its most efficient rival.

Our main finding is that the addition of a third firm with $s_3 \leq s_0$ promotes adoption. The economic intuition behind this result is simple: The spurious rents that firm 2 may appropriate by preventing the learning of the new technology are now limited by the competitive pressure of firm 3. This, in turn, leads to a higher adoption payoff as it reduces the surplus that firm 1 must concede to customers along its learning path. The next example paves the way for the formal statement of the result in Proposition 7.

Example 5. Consider the two-period model of Example 1 and add a third, non-adopting firm with unitary cost $c_3 \in [0.5, 0.75]$. Let $\varepsilon = 0.1$. Because $v_3 = v_0 = 1$ and $c_0 = 0.5$, we have $s_0 = 0.5$ and $s_3 \in [0.25, 0.5]$. Recall that the new technology lets firm 1 produce a second unit at a cost 0.1 after producing the first unit at a cost 0.75. Therefore, Assumption A2 is satisfied just as in Example 1, and the three-firm versions of Assumptions A1 and A3 are also met:

$$\begin{aligned} s(0) &\leq \max\{s_0, s_3\}, \\ s(0) + s(1) - 2\max\{s_0, s_3\} &\geq 2s_0 + \varepsilon. \end{aligned}$$

Suppose that firm 1 adopts the new technology and sells to the first customer. Then, firm 1 will also sell to the second customer and get a second-period payoff equal to $0.5 - 0.1 = 0.4$ because the presence of firm 3 does not alter the equilibrium price in this state. Firm 1's adoption payoff is thus

$$\underbrace{p^\dagger - 0.75}_{\substack{\text{Current}\\ \text{payoff}}} + \underbrace{[0.4]}_{\substack{\text{Future}\\ \text{payoff}}}, \tag{12}$$

where $p^\dagger$ is the equilibrium price at period 0 if firm 3 is present. (Figures that are specific to the three-firm model are denoted with †.) This price must satisfy

$$1 - p^\dagger \geq b_2^\dagger,$$

where $1 - p^\dagger$ is the surplus (utility) that the first customer gets if he buys from firm 1 and $b_2^\dagger$ is the maximum surplus that firm 2 is willing to offer to the first customer.

We know that firm 2 is willing to concede at least its flow surplus to the first customer, that is, $b_2^\dagger \geq s_0 = 0.5$. Besides, by selling to the first customer, firm 2 could prevent the learning of the new technology and keep the cost of firm 1 at 0.75 in the second period. But now, and this is the key observation, firm 2 would compete for the second customer against firm 3, which is *more efficient than firm 1* in this state. So, firm 2 would get a second-period payoff of $s_0 - s_3$.[12] In short:

$$b_2^\dagger = \underbrace{s_0}_{\substack{\text{Flow}\\ \text{surplus}}} + \underbrace{[s_0 - s_3]}_{\substack{\text{Future}\\ \text{rent}}} = 1 - s_3.$$

Consequently, firm 1 can charge the first customer a price (at most) equal to:

$$\overline{p}^\dagger = c_0 - [s_0 - s_3] = 0.5 - [0.5 - s_3].$$

Plugging this (maximum) price into (12) we see that it would result in a non-negative adoption payoff for any $s_3 \geq 0.35$ or, equivalently, for any $c_3 \leq 0.65$. Because the pre-adoption payoff of firm 1 is zero and the sunk cost is 0.1, the presence of a third firm results in firm 1 adopting the new technology for any $c_3 < 0.55$. Moreover, the second-period, spurious rent of firm 2 goes down, whereas the payoff of firm 1 goes up, with s_3.

Incidentally, adding a third firm has the *only* effect of reducing the size of the shaded area in Figure 1 in the two-period, cost-saving model. Note that the potential second-period payoff of firm 1 is still given by B, whereas its first-period flow surplus remains to be C. Likewise, the first-period flow surplus of firm 2 continues to be C plus A. Therefore, the presence of firm 3 only reduces the potential second-period profits of firm 2 that cause adoption breakdowns.

It is worth mentioning the link between extra competition and equilibrium prices in the two-period model. Combining (12) and (1) we see that the difference between adoption payoffs with two and three firms is

$$p^\dagger - p = \max\{s_3 - s(0), 0\}. \tag{13}$$

This equation captures what we call the 'protective' effect of competition: By decreasing the socially spurious rents of firm 2, extra competition permits firm 1 to charge higher prices and thus appropriate a higher share of the social value of the new technology. (Of course, there is no protective effect if $s_3 \leq s(0)$.)

We proceed now to generalize these results. Let $\mathrm{S}^\dagger$ be the set of adopted technologies with three firms. We say that the value of competition is (weakly) *positive* if, and only if, $\mathrm{S}^\dagger \supseteq \mathrm{S}^*$.[13] Likewise, we say that the value of competition is (weakly) *increasing* in s_3 if, and

[12] Or, equivalently, $c_3 - c_0$.

[13] Recall that S^* is the set of adopted technologies with two firms.

only if, set $S^\dagger$ is non-decreasing in s_3; i.e., $\hat{S}^\dagger \supseteq \tilde{S}^\dagger$ if $\hat{s}_3 > \tilde{s}_3$.

Proposition 7 (The value of increasing competition)**.** *Let a third firm with $s_3 \leq s_0$ be added to our original model. Then, the value of competition is positive and increasing. Furthermore, if $s_3 = s_0$ the MPE outcome is efficient.*

Arrow (1962b) showed that adoption incentives are stronger under competition than under monopoly. For a non-drastic innovation, the cost savings due to adoption are higher in a competitive industry because output is larger than under monopoly. For a drastic innovation the idea is slightly different as post-adoption output under competition and monopoly coincide. What discourages innovation in this case is that, by adopting a technology, the monopolist loses pre-adoption profits that are zero under perfect competition. More competition is also beneficial in Holmes et al. (2012). In their model, the adopting firm reduces output during the disruption period and, hence, loses pre-adoption profits. Because such loses increase with the market power of adopting firms, those with less market power have stronger adoption incentives.

The Arrow effect and the Holmes et al. effect work trough the output restrictions that firms with market power undertake relative to competitive firms. In both cases, lost pre-adoption profits are key to show that competition is beneficial. On the contrary, in our model demand is inelastic and pre-adoption profits are zero. Hence none of these forces is present. Firms compete each period for a single customer with unit demand and, therefore, the competitive pressure of additional firms works (only) by limiting the future prices that the second firm may charge once it has impeded the learning of the technology. Our mechanism thus offers a novel channel through which extra competition may facilitate the appropriation of the value of a new technology.

7. Concluding Comments

We have presented a dynamic model of technology adoption that captures the idea that adoption creates socially spurious rents for non-adopters. Within this framework, we have shown that adoption breakdowns may come as a consequence of disruption costs and learning by doing. We have been able to characterize the technologies most prone to experience adoption failures, i.e., technologies with slow learning curves. As a corollary, we have shown that firms may prefer adopting inferior technologies if these can be learned faster. We have also assessed the impact of adding more firms obtaining that increasing competition spurs adoption. Summing up, our results should warn regulators of keeping an eye on industries either with few competitors or where technological improvements take longer to settle. In our view, these are the industries in which adoption failures seem most likely to happen.

Our results generalize straightforwardly in a number of directions. Firms could discount the future. We may consider exogenous technological change by simply letting function s depend on time as well as on accumulated sales. Likewise, old technologies could be subject to exogenous progress.

Other generalizations require significant departures from our setup. Among these, introducing randomness is perhaps the most natural. We learned in Corollary 1 that adopted technologies never fail in equilibrium. This feature of our model arises because we consider only deterministic technologies. A stochastic model could account for the failure of adopted technologies.

References

Arrow, K. J. (1962a). The economic implications of learning by doing. *Review of Economic Studies 29*(3), 155–173.

Arrow, K. J. (1962b). Economic Welfare and the Allocation of Resources for Inventions. In R. Nelson (Ed.), *The Rate and Direction of Inventive Activity.* Princeton University Press.

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

Besanko, D., U. Doraszelski, and Y. Kryukov (2014). The economics of predation: What drives pricing when there is learning-by-doing. *American Economic Review 104*(3), 868–897.

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

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

Chari, V. V. and H. Hopenhayn (1991). Vintage human capital, growth, and the diffusion of new technology. *Journal of Political Economy 99*(6), 1142–1165.

Christensen, C. M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.* Harvard Business Review Press.

Deneckere, R. J. and D. Kovenock (1996). Bertrand-edgeworth duopoly with unit cost asymmetry. *Economic Theory 8*(1), 1–25.

Gilbert, R. J. and D. M. G. Newbery (1982). Preemptive patenting and the persistence of monopoly. *American Economic Review 72*(3), 514–526.

Henderson, B. (1968). The experience curve. *BCG Perspectives 87*.

Holmes, T. J., D. K. Levine, and J. A. Schmitz (2012). Monopoly and the incentive to innovate when adoption involves switchover disruptions. *American Economic Journal: Microeconomics 4*(3), 1–33.

Holmes, T. J. and J. A. Schmitz (2010). Competition and Productivity: A Review of Evidence. Research Department Staff Report 439, Federal Reserve Bank of Minneapolis.

Huggett, M. and S. Ospina (2001). Does productivity growth fall after the adoption of new technology? *Journal of Monetary Economics 48*(1), 173–195.

Jovanovic, B. and Y. Nyarko (1996). Learning by doing and the choice of technology. *Econometrica 64*(6), 1299–1310.

Klenow, P. J. (1998). Learning curves and the cyclical behavior of manufacturing industries. *Review of Economic Dynamics 1*(2), 531–550.

Leonard-Barton, D. (1988). Implementation as mutual adaptation of technology and organization. *Research Policy 17*(5), 251–267.

Ovans, A. (2011). The charts that changed the world. *Harvard Business Review*. December.

Parente, S. L. (1994). Technology adoption, learning-by-doing, and economic growth. *Journal of Economic Theory 63*(2), 346–369.

Schivardi, F. and M. Schneider (2008). Strategic experimentation and disruptive technological change. *Review of Economic Dynamics 11*(2), 386–412.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy.* Harper & Brothers.

Thompson, P. (2010). Learning by doing. In B. H. Hall and N. Rosenberg (Eds.), *Handbook of the Economics of Innovation*, Volume 1, Chapter 10, pp. 429–476. North-Holland.

Tyre, M. J. and O. Hauptman (1992). Effectiveness of organizational responses to technological change in the production process. *Organization Science 3*(3), 301–320.

Wright, T. P. (1936). Factors affecting the costs of airplanes. *Journal of the Aeronautical Sciences 3*(4), 122–128.

A. Proofs

Auxiliary Results

The following intermediate lemmas are useful to prove Theorem 1. Recall that:

$$\hat{x} \equiv \min\{x \in \mathrm{X} : s(x) \geq \ s_0\}.$$

Also, note that:

$$z(x,t) = \sum_{k=0}^{T-t} d(x, t+k), \tag{14}$$

$$= -\sum_{k=0}^{T-t} r(x+k, t+k). \tag{15}$$

Lemma A. If $z(x,t) \geq 0$, then: $(a1)$ $d(x,t) \geq 0$, $(b1)$ $z(x, t{+}1) \leq z(x,t)$, and: $(c1)$ $z(x{+}1, t{+}1) \geq 0$. If $z(x,t) \leq 0$, then: $(a2)$ $r(x,t) \geq 0$, $(b2)$ $z(x+1,t+1) \geq z(x,t)$, and: $(c2)$ $z(x,t+1) \leq 0$.

Proof. $(a1)$: If $z(x,t) \geq 0$, then $d(x,t)$ is the largest summand in (14) because $d(x, t+k)$ decreases with k. $(b1)$: Since, from (14), $z(x,t+1) = z(x,t) - d(x,t)$, $(a1)$ implies $(b1)$. $(c1)$: From (15), $z(x+1,t+1) = z(x,t) + r(x,t)$. If $x \leq \hat{x}$, then, as $\pi(x) \leq 0$, $r(x,t) \geq 0$ which gives the result. If $x > \hat{x}$, then $z(x+1,t+1) > 0$ since it is equal to a negative sum of negative values of r. $(a2)$: If $z(x,t) \leq 0$, then $r(x,t)$ is the largest summand in (15) because $r(x,t)$ is decreasing in t and non-increasing in x. $(b2)$ As in $(c1)$, from (15), $z(x+1,t+1) = z(x,t) + r(x,t)$ and thus $(a2)$ implies $(b2)$. $(c2)$ It follows from (14), as $d(x,t+k)$ decreases with k. □

Lemma B. Functions d, r, and z fulfill:

d. If $z(x+1,t+1) \leq 0$, then $0 \leq r(x,t+1) \leq -z(x,t+1)$ and $0 \leq r(x,t) \leq -z(x,t)$.

e. If $z(x,t+1) \geq 0$, then $0 \leq d(x+1,t+1) \leq z(x+1,t+1)$ and $0 \leq d(x,t) \leq z(x,t)$.

f. If $z(x,t+1) \leq 0$ and $z(x+1,t+1) \geq 0$, then either:

  f1. $z(x+1,t+2) \geq 0$ and $-z(x,t+1) \leq r(x,t+1)$, or:

  f2. $z(x+1,t+2) \leq 0$ and $z(x+1,t+1) \leq d(x+1,t+1)$.

Proof. The proof repeatedly uses the results from Lemma A.

(d): If $z(x{+}1,t{+}1) \leq 0$ then, $z(x{+}1,t{+}2) \leq 0$ by $(c2)$. Hence, by $(c1)$, $z(x,t{+}1) \leq 0$. This, in turn, implies that $r(x,t{+}1) \geq 0$ by $(a2)$. Since, from (15), $z(x,t{+}1) = z(x{+}1,t{+}2) - r(x,t{+}1)$, we have the first part of d. For the second part, note that, by $(c1)$, $z(x,t) \leq 0$, which, by $(a2)$, implies that $r(x,t) \geq 0$. Since by (15), $z(x,t) = z(x+1,t+1) - r(x,t)$, we have the second part of d.

(e): If $z(x, t+1) \geq 0$ then, $d(x, t+1) \geq 0$ by ($a1$) and $z(x+1, t+2) \geq 0$ by ($c1$). Also, as $s(\cdot)$ is non-decreasing in x, $d(x+1, t+1) \geq 0$. Since, from (14), $z(x+1, t+2) = z(x+1, t+1) - d(x+1, t+1)$, we already have the first part of e. For the second part, note that, by ($c2$), $z(x, t) \geq 0$, which, by ($a1$), implies that $d(x, t) \geq 0$. Since, by (14), $z(x, t+1) = z(x, t) - d(x, t)$, we have the second part of e.

($f1$): $z(x, t+1) \leq 0$ implies, by ($a2$), that $r(x, t+1) \geq 0$. Since, by (15), $z(x, t+1) = z(x+1, t+2) - r(x, t+1)$, we have $f1$.

($f2$): $z(x+1, t+1) \geq 0$ implies, by ($a1$), that $d(x+, t+1) \geq 0$. Since, by (14), $z(x+1, t+1) = d(x+1, t+1) + z(x+1, t+2)$, we have the result. □

We use throughout the following concepts. The value functions of the firms are:

$$\pi_1(x, t) = \max\{s(x) - \bar{b}_2(x, t) + \pi_1(x+1, t+1), \pi_1(x, t+1)\},$$
$$\pi_2(x, t) = \max\{s_2 - \bar{b}_1(x, t) + \pi_2(x, t+1), \pi_2(x+1, t+1)\}.$$

With $\bar{b}_i(x, t)$, $i = 1, 2.$, we denote the maximum bidding function, i.e., the surplus firm i is willing to transfer to the customer at state (x, t):

$$\bar{b}_1(x, t) = s(x) + \pi_1(x+1, t+1) - \pi_1(x, t+1), \tag{16}$$
$$\bar{b}_2(x, t) = s_2 + \pi_2(x, t+1) - \pi_2(x+1, t+1). \tag{17}$$

(firm 1 sells at (x, t) if $\bar{b}_1(x, t) = \bar{b}_2(x, t)$.)

Proof of Theorem 1

The proof is by backwards induction. Let $t = T$. The result is then obvious for the $T+1$ triangular subarrays $A_{x,t}$ for $x \in \mathrm{X}$, i.e., terminal states of the form $(\cdot, t)$ for which $d(\cdot, t) = -r(\cdot, t) = z(\cdot, t) = s(\cdot) - s_0$. Using the maximum bidding functions, payoffs, at any non-terminal state (x, t), in a MPE are:

$$\pi_1(x, t) = \max\{s(x) - s_0 + \pi_1(x+1, t+1) + \pi_2(x+1, t+1) - \pi_2(x, t+1), \pi_1(x, t+1)\}, \tag{18}$$

$$\pi_2(x, t) = \max\{s_0 - s(x) + \pi_2(x, t+1) + \pi_1(x, t+1) - \pi_1(x+1, t+1), \pi_2(x+1, t+1)\}. \tag{19}$$

Let us now consider a generic time period t. We prove that the result is true for the $t+1$ triangular subarrays $A_{x,t}$ for $x \in \{0, ..., t\}$ if it is true for the $t+2$ triangular subarrays $A_{x,t+1}$, the induction hypothesis.

(a): If $z(x+1,t+1) \leq 0$, we know from (d) in Lemma B, equations (7) and (8) that $\pi_1(x+1,t+1) = \pi_1(x,t+1) = 0$ and that $\pi_2(x,t+1) = r(x,t+1)$. On the other hand, we have, by ($a2$) in Lemma A, that $r(x+1,t+1) \geq 0$. This, in turn, implies, by definition of function $r(\cdot,\cdot)$, that $\pi(x+1) \leq 0$. As $\pi(x)$ is non-decreasing in x, we have that $\pi_2(x+1,t+1) \leq r(x+1,t+1) \leq r(x,t+1)$. Plugging these into (18) and (19) gives $\pi_1(x,t) = 0$ and $\pi_2(x,t) = r(x,t)$.

(b): If $z(x,t+1) \geq 0$, we know from (e) in Lemma B, equations (7) and (8) that $\pi_2(x,t+1) = \pi_2(x+1,t+1) = 0$ and that $\pi_1(x+1,t+1) = d(x+1,t+1)$. On the other hand, we have, by ($a1$) in Lemma A, that $d(x,t+1) \geq 0$. As $d(x,\cdot)$ is non-increasing in t, we have that $\pi_1(x,t+1) \leq d(x,t+1) \leq d(x,t)$. Plugging these into (18) and (19) gives $\pi_1(x,t) = d(x,t)$ and $\pi_2(x,t) = 0$.

(c): If $z(x,t+1) \leq 0$ and $z(x+1,t+1) \geq 0$, equations (7) and (8) say that $\pi_2(x+1,t+1) = \pi_1(x,t+1) = 0$. Then either $\pi_1(x+1,t+1) = d(x+1,t+1)$, or $\pi_1(x+1,t+1) = z(x+1,t+1)$. Let us first regard the case in which $\pi_1(x+1,t+1) = d(x+1,t+1)$. It follows then from equation (7) that $z(x+1,t+1) \geq d(x+1,t+1) \geq 0$ and thus that $z(x+1,t+2) \geq 0$, since $z(x+1,t+2) = z(x+1,t+1) - d(x+1,t+1)$. Hence, from ($f1$) in Lemma B, we know that $\pi_2(x,t+1) = z(x,t+1)$. Plugging these into (18) and (19) gives $\pi_1(x,t) = \max\{d(x,t) + z(x,t+1), 0\} = \max\{z(x,t), 0\}$ and $\pi_2(x,t) = \max\{-z(x,t), 0\} = \max\{r(x,t) - z(x+1,t+1), 0\}$. A parallel argument shows that the same result holds when $\pi_1(x+1,t+1) = z(x+1,t+1)$.

The previous paragraph is not valid for $t = T-1$, because state $(x+1,t+2)$ is not feasible. [It is easy to see that this problem appears if, and only if, we are at state $(\hat{x}-1,T-1)$.] We have that $\pi_1(\hat{x}-1,T) = \pi_2(\hat{x},T) = 0$, $\pi_1(\hat{x},T) = s(\hat{x}) - s_0$, and $\pi_2(\hat{x}-1,T) = s_0 - s(\hat{x}-1)$. Plugging these into (18) and (19) completes the proof.

Proof of Corollary 1

From Theorem 1, the result is obvious for the case in which $t = T-1$ and for the $T+1$ triangular subarrays $A_{x,t}$ for $x \in X$, i.e., terminal states of the form $(\cdot,t)$. Consider now a generic non-terminal state (x,t).

(a): If $z(x+1,t+1) \leq 0$, it follows from (a) in Theorem 1, (16) and (17) that $\bar{b}_1(x,t) = s(x)$ and $\bar{b}_2(x,t) = s_0 + [r(x,t+1) - \pi_2(x+1,t+1)]$. Then, $\bar{b}_1(x,t) < \bar{b}_2(x,t)$ since $s_0 \geq s(x)$ and, from (a) in Theorem 1, $\pi_2(x+1,t+1) \leq r(x+1,t+1) \leq r(x,t+1)$.[14] Thus, firm 2 will be the trading firm at date t. Since the state moves to $(x,t+1)$ and by ($c2$) in Lemma A, $z(x+1,t+2) \leq 0$, we have the result.

(b): If $z(x,t+1) \geq 0$, it follows from (b) in Theorem 1, (16) and (17) that $\bar{b}_2(x,t) = s_0$ and $\bar{b}_1(x,t) = s(x) + [d(x+1,t+1) - \pi_1(x,t+1)]$. Then, $\bar{b}_1(x,t) \geq \bar{b}_2(x,t)$ since $\bar{b}_1(x,t) - \bar{b}_2(x,t) = d(x,t) - \pi_1(x,t+1)$ and, from (b) in Theorem 1, $\pi_1(x,t+1) \leq d(x,t+1) \leq d(x,t)$. Thus, seller 1 will be the trading seller at date t. Since the state moves to $(x+1,t+1)$ and by ($c1$) in Lemma A, $z(x+1,t+2) \geq 0$, we have the result.

[14] That $s_0 \geq s(x)$ follows from $\pi(x) \leq 0$.

(c): If $z(x,t+1) \leq 0$ and $z(x+1,t+1) \geq 0$, we know from (c) in Theorem 1 that $\pi_2(x+1,t+1) = \pi_1(x,t+1) = 0$. Then either $\pi_1(x+1,t+1) = d(x+1,t+1)$, or $\pi_1(x+1,t+1) = z(x+1,t+1)$. Let us first regard the case in which $\pi_1(x+1,t+1) = d(x+1,t+1)$. It follows then from ($c$) in Theorem 1, (16) and (17) that $\bar{b}_1(x,t) = s(x) + d(x+1,t+1)$ and $\bar{b}_2(x,t) = s_0 + z_2(x,t+1)$. Thus $\bar{b}_1(x,t) - \bar{b}_2(x,t) = d(x,t) - z(x,t+1) = z(x,t)$. When $z(x,t)$ is positive, seller 1 will be the trading seller at date t. Since the state moves to $(x+1,t+1)$ and by ($c1$) in Lemma A, $z(x+1,t+1) \geq 0$, we have the result. Clearly, the same result holds when $z(x,t)$ is negative and seller 2 is the trading seller at date t. Finally, a parallel argument shows that the same result is true when $\pi_1(x+1,t+1) = z(x+1,t+1)$.

Proof of Lemma 1

If the set of adopted technologies $\mathrm{S}^* = \emptyset$, the result holds trivially. We assume hereafter that S^* is non-empty. If a technology is adopted at date t, then:

$$\pi_1(0,t) \geq \varepsilon.$$

As S^* is non-empty, there is a date $t^* \in \mathrm{T}$ such that: (i) $\pi_1(0,t^*) \geq \varepsilon$; and that: ($ii$) $\pi_1(0,t^*) \geq \pi_1(0,t)$ $\forall t \in \mathrm{T}$. Also, note that it must be that $\pi_1(0,0) \geq 0$ since $\pi_1(0,0) \geq \pi_1(0,t)$ $\forall t \in \mathrm{T}$, by Theorem 1.

(a): If $\pi_1(0,0) = z(0,0)$, then $\pi_1(0,t) = 0$ $\forall t \geq 1$, by (c) in Theorem 1. Therefore $t^* = 0$.

(b): If $\pi_1(0,0) = d(0,0)$ and if, $\forall t \geq 1$, $\pi_1(0,t)$ is smaller than ε, the result holds trivially. Thus, let $\pi_1(0,t) \geq \varepsilon$ for at least one $t \geq 1$. Then, fromTheorem 1, we have that $\pi_1(0,t) \leq d(0,t)$. And as, $\pi_1(0,0) = d(0,0) > d(0,t)$ for $\forall t \geq 1$, it follows that $t^* = 0$.

Proof of Proposition 1

If switchover disruption costs are zero, i.e., $\pi(0) \geq 0$, $d(x,t) \geq 0$ for every state (x,t) and every $s \in \mathrm{S}$. Thus, $z(x,t) \geq d(x,t)$ for every state (x,t) and every $s \in \mathrm{S}$. This, in turn, implies, from Theorem 1, that $\pi_1(x,t) = d(x,t)$ for every $s \in \mathrm{S}$. Hence, it follows that $\pi_1(0,0) = d(0,0) \geq \varepsilon$ for every $s \in \mathrm{S}$ by Assumption A3.

Proof of Proposition 2

Let s be an element of S^*, then $\pi_1(0,0) \geq \varepsilon$. This, in turn, implies that $z(0,0) \geq \varepsilon$ since, from Theorem 1, $\pi_1(0,0) = \min\{d(0,0), z(0,0)\}$. Therefore, $s \in \mathrm{N}^c$ and $\mathrm{S}^* \subset \mathrm{N}^c$. Conversely, let s be an element of N^c, then $z(0,0) \geq \varepsilon$. If $\pi_1(0,0) = z(0,0)$, then clearly $s \in \mathrm{S}^*$. If $\pi_1(0,0) = d(0,0)$, then $s \in \mathrm{S}^*$ since $d(0,0) \geq \varepsilon$ by Assumption A3. Hence, $s \in \mathrm{S}^*$ and $\mathrm{N}^c \subset \mathrm{S}^*$. This completes the proof.

Proof of Proposition 3

We break the proof in two parts and several steps:
Part One (Seller 2 profits): Recall that $-z(x,t) = r(x,t) - z(x+1,t+1)$.

Step 1: If $x = \hat{x} - 1$, then $\pi_1(x+1, t+1+k) = d(x+1, t+1+k) \geq 0$ for $0 \leq k \leq T - (t+1)$. Since:

$$z(x+1,t+1) = \sum_{k=0}^{T-(t+1)} d(x+1, t+1+k),$$

the proof for $x = \hat{x} - 1$ is complete.

Step 2: If $z(x+1,t+1) < 0$, we know from (d) in Lemma B and (8) that $\pi_2(x,t) = r(x,t)$. By $(c2)$ in Lemma A and (7) we know that $\pi_1(x+1, t+1+k) = 0$ for $0 \leq k \leq T - (t+1)$.

Step 3: This an auxiliary result. Suppose that $z(x,t) \geq 0$ and let:

$$\hat{k} = \max_{0 \leq k \leq T-t} \{k | z(x, t+k) \geq 0\}.$$

Then $z(x,t+k) \geq d(x,t+k) \geq 0$ if $k < \hat{k}$ and $0 \leq z(x,t+k) \leq d(x,t+k)$ if $k = \hat{k}$. Both inequalities come from the definition of $\hat{k}$, the fact that $z(x,t+k) = d(x,t+k) + z(x,t+k+1)$ and $(a1)$ in Lemma A.

Combining these facts with Theorem 1, we get that $\pi_1(x,t+k)$ is equal to $d(x,t+k)$ if $k < \hat{k}$, equal to $z(x,t+\hat{k})$ if $k = \hat{k}$, and zero otherwise.

Step 4: If $z(x+1,t+1) \geq 0$ and $x < \hat{x} - 1$, we write

$$z(x+1,t+1) = \sum_{l=0}^{\hat{k}-1} d(x+1, t+1+l) + z(x+1, t+1+\hat{k}),$$

where $\hat{k}$ is the integer defined in the previous step. Combining $-z(x,t) = r(x,t) - z(x+1,t+1)$, Step 3 and (8), we have the result.

Part Two (Seller 1 profits): Recall that $z(x,t) = d(x,t) + z(x,t+1)$.

Step 1: If $z(x,t+1) \geq 0$, we know from (e) in Lemma B and (7) that $\pi_1(x,t) = d(x,t)$. By $(c1)$ in Lemma A and (8) we know that $\pi_2(x+k, t+1+k) = 0$ for $0 \leq k \leq T - (t+1)$.

Step 2: This an auxiliary result. Suppose that $z(x,t) < 0$ and let:

$$\hat{k} = \max_{0 \leq k \leq T-t} \{k | z(x+k, t+k) < 0\}.$$

Then $0 \leq r(x+k,t+k) \leq -z(x+k,t+k)$ if $k < \hat{k}$ and $0 \leq -z(x+k,t+k) \leq r(x+k,t+k)$ if $k = \hat{k}$. Both inequalities come from the definition of $\hat{k}$, the fact that $-z(x+k,t+k) = r(x+k,t+k) - z(x+k+1,t+k+1)$ and $(a2)$ in Lemma A.

Combining these facts with Theorem 1, we get that $\pi_2(x+k, t+k)$ is equal to $r(x+k, t+k)$ if $k < \hat{k}$, equal to $-z(x+k, t+\hat{k})$ if $k = \hat{k}$, and zero otherwise.

Step 3: If $z(x, t+1) < 0$, we write:

$$-z(x, t+1) = \sum_{l=0}^{\hat{k}-1} r(x+l, t+1+l) - z(x+\hat{k}, t+1+\hat{k}),$$

where $\hat{k}$ is the integer defined in the previous step. Combining $z(x,t) = d(x,t) + z(x,t+1)$, Step 2 and (7), we have the result.

Proof of Proposition 4

For any $s \in \mathrm{S}$:

$$\begin{aligned} z(0,1) &= \sum_{k=0}^{T-1} (T-k)\pi(k). \\ &\geq \sum_{k=0}^{T-1} (T-k)\pi(0) = \frac{1}{2}T(T+1)\pi(0) := -K. \end{aligned}$$

If s is an element of G, then $z(0,0) = d(0,0) + z(0,1) \geq d(0,0) - K \geq \varepsilon$. If $\pi_1(0,0) = z(0,0)$, then $s \in \mathrm{S}^*$. If $\pi_1(0,0) = d(0,0)$, then by definition $s \in \mathrm{S}^*$. This completes the proof.

Proof of Proposition 5

Consider any s and s' in S. If $z'^-(0,1) = z^-(0,1) = 0$, we are done. Thus, assume that $z'^-(0,1)$ and $z^-(0,1)$ are strictly positive and let $\Delta := z'^-(0,1) - z^-(0,1)$. Then, as $s \succeq s'$, there is a $1 \leq k \leq (T-1)$ such that:

$$\Delta = \sum_{x=1}^{k} (T-x)\xi(x) + \sum_{x=k+1}^{T-1} (T-x)\xi(x). \tag{20}$$

$$\xi(1) + \cdots + \xi(k) + \xi(k+1) + \cdots + \xi(T) = 0, \tag{21}$$

where empty sums are taken to be zero, $\xi(x) := (\pi(x) - \pi'(x))$, $\xi(x) \geq 0$ for $x \leq k$, and $\xi(x) \leq 0$ for $x \geq k+1$. Using (21) into (20), we have:

$$\begin{aligned} \Delta &\geq -(T-k) \sum_{x=k+1}^{T} \xi(x) + \sum_{x=k+1}^{T-1} (T-x)\xi(x), \\ &= -\sum_{l=1}^{T-k} l\xi(k+l) > 0, \end{aligned}$$

and the proof is complete.

Proof of Proposition 6

Given that by Proposition 2 there are technologies s in N with a positive social value, it suffices to show that there are technologies in N^c with social value arbitrarily close to zero. Simply consider a technology s_ν ($\nu > 0$) with $\pi_\nu(0) = 0$, $\pi_\nu(1) = \cdots = \pi_\nu(T) = \frac{\nu}{T}$, and sunk cost ε. Assume that s_ν has a social value arbitrarily close to zero, i.e., $\nu - \varepsilon$ arbitrarily close to zero. Then as:

$$z_\nu(0,0) = \frac{\nu}{2}(T+1) > \varepsilon,$$

the proof is complete.

Proof of Proposition 7

We aim to show that more technologies are adopted with three sellers than with two sellers. Since only technologies that give positive payoffs are adopted, this is akin to show that $\pi_1^\dagger(0,0) \geq \pi_1(0,0)$ for every technology $s \in \mathrm{S}$ (figures that refer to the three-seller model are denoted with †).

Because $s_3 \leq s_0$, the equilibrium profits of seller 3 are $\pi_3^\dagger(x,t) = 0$ at every state. If $s_3 \leq s(0)$, seller 3 is irrelevant and the three-seller model is identical to the two-seller model. Let $s_3 > s(0)$, and define $x^\dagger$ as

$$x^\dagger \equiv \min\{x \in \mathrm{X} : s(x) \geq s_3\}.$$

For all states such that $x \geq x^\dagger$, seller 3 is irrelevant and payoffs are those given in Theorem 1. Now we consider states with $x < x^\dagger$. Our proof is by backwards induction. We start with states of the form $(x^\dagger - 1, t)$. Define

$$\tau(x) \equiv \min\{t \in \mathrm{T} : \pi_2(x+1, t+1) > 0\}$$

if the set is not empty, and $\tau(x) \equiv T + 1$ if $\pi_2(x+t, t+1) = 0$ for all t.

For all $t \geq \tau(x^\dagger - 1)$ seller 2 makes the minimum bid in the three-seller model, just as he does in the two-seller model, because he gets nothing out of winning the current customer. On the other hand, because $s(x^\dagger - 1) \leq s_3$, seller 2 now competes with seller 3 instead of with seller 1 and thus bids less with three sellers than with two sellers. As a consequence, the profits of the first seller in the three-seller model remain unchanged at zero whereas the profits of the second seller go down (or remain the same if $s(x^\dagger - 1) = s_3$).

At state $(x^\dagger - 1, \tau(x^\dagger - 1) - 1)$ seller 1 bids the same with three sellers as with two sellers because his profits at (future) adjacent states do not change. That is, he gets nothing if he misses the current sale and moves to $(x^\dagger, \tau(x^\dagger - 1))$ if he sells, were Theorem 1 still applies. However, seller 2 now expects a lower profit from winning the current sale—for the same reason as above; he competes now against seller 3—and gets nothing if he misses it. As a consequence, the profits of seller 1 cannot go down and the profits of seller 2 cannot go up with three sellers.

Two things may happen at this state. If seller 1 sells, then he will also sell at any $(x^\dagger - 1, t)$ with $t < \tau(x^\dagger - 1)$, earning the maximum profit $d(x^\dagger - 1, t)$ at each such state. This occurs because seller 2 always bids the minimum, s_0, if he does not expect to win the next sale. If, on the contrary, seller 2 sells, then we must move one period backwards and see what happens at $(x^\dagger - 1, \tau(x^\dagger - 1) - 2)$. At this state seller 1 bids the same with three sellers as with two sellers, whereas seller 2 bids less—for the reasons we gave in the previous paragraph. Therefore, the profits of seller 1 cannot go down, and the profits of seller 2 cannot go up, at this state. Again, two things may happen at $(x^\dagger - 1, \tau(x^\dagger - 1) - 2)$ if seller 2 sells at $(x^\dagger - 1, \tau(x^\dagger - 1) - 1)$: either seller 1 or seller 2 sells. The same argument above applies over and over again and we conclude that the profits of seller 1 cannot go down, and seller 2's profits cannot go up, for any state of the form $(x^\dagger - 1, t)$.

We now consider states of the form $(x^\dagger - 2, t)$. Define $\tau^\dagger(x)$ as the three-seller analog of $\tau(x)$, that is, the first t at which seller 2 gets the maximum payoff $\pi_2^\dagger(x, t) = (T + 1 - t)(s_0 - s_3)$ in equilibrium. We have shown already that $\tau^\dagger(x^\dagger - 1) \geq \tau(x^\dagger - 1)$. For $t \geq \tau^\dagger(x^\dagger - 2)$, seller 2 clearly makes no more, whereas seller 1 makes the same (zero), with three sellers than with two sellers. At $t = \tau^\dagger(x^\dagger - 2) - 1$, seller 1 bids more with three sellers because, as we have shown above, he always gets more by moving to $(x^\dagger - 1, \tau^\dagger(x^\dagger - 2))$ with three sellers than with two—because his profits at any $(x^\dagger - 1, t)$ are higher with three sellers. On the other hand, seller 2 makes less from winning the current customer for the usual reason, i.e., because $s(x^\dagger - 2) \leq s_3$, while he makes nothing if he misses it.

Again, two things may happen at this state. If seller 1 sells, then he also sells at all previous periods and always earns the maximum profit $\pi_1^\dagger(x^\dagger - 2, t) = d(x^\dagger - 2, t)$. If seller 1 sells, however, we must move one period backwards. At $(x^\dagger - 2, \tau^\dagger(x^\dagger - 1) - 2)$, again, seller 1 bids more with three sellers because, as we have shown, his profits for states $(x^\dagger - 1, t)$ are larger with three sellers. On the other hand, seller 2 gets zero if he misses the current sale and, as we have just said in the previous paragraph, he gets less if he moves forward to $(x^\dagger - 2, \tau^\dagger(x^\dagger - 1) - 1)$ with three sellers. The same argument can be applied a finite number of times to show that the profits of seller 1 cannot go down for any state of the form $(x^\dagger - 1, t)$. The same chain of reasoning is repeated backwards for each $x^\dagger - k$ $(k = 3, 4, \dots)$.

Finally, note that, at each step, the change in seller 1's payoffs occurs because the maximum payoff of seller 2, $\pi_2^\dagger(x, t) = (T + 1 - t)(s_0 - s_3)$, goes down with s_3. Therefore, the payoff of seller 1 is non-decreasing in s_3, with a minimum $\pi_1^\dagger(0, 0) = \pi_1(0, 0)$ at $s_3 = s(0)$, and a maximum $\pi_1^\dagger(0, 0) = d(0, 0)$ at $s_3 = s_0$.