Disruption costs, learning by doing, and the choice of technology*

Carlos J. Pérez† Carlos J. Ponce‡

October 2013

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 buyers it needs to learn the technology. The prospect of future rents by the rival results in: (i) a failure to adopt Pareto superior technologies; (ii) an equilibrium preference for the choice of technologies with smaller (discounted) social value but flow payoffs that are received earlier in time; (iii) more technologies being adopted as the adopting firm is exposed to more competition.

1. Introduction

The adoption of new technologies is at the center of productivity growth in many industrial sectors. In most of these industries, however, the nature of the adoption

*Pérez acknowledges support from CONICYT under the FONDECYT 2014 Postdoctoral Grant no. 3140360.

†Pontificia Universidad Católica de Chile; cjperez.econ@gmail.com.

‡Ilades-Georgetown; cjponcebruera@gmail.com.

process affects the competitive position of firms, creating losses as well as gains. This article develops a simple model of technology adoption with two distinctive features. First, adoption involves switchover disruption costs: the firm adopting the technology becomes initially less productive than non-adopters. Second, adoption entails learning by doing: the more the firm uses the technology, the more productive it gets.

The key observation of this article is that, upon adoption, non-adopting firms have incentives to undercut prices to prevent the learning of the new technology—as this makes the adopting firm a weaker competitor. The expectation of future rents by non-adopters 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 buyers from the adopting firm creates future rents without adding any social value, buyers become an artificially overpriced 'commodity' in the market. This overpricing may render adoption unprofitable.

This article studies these issues in a dynamic duopoly model of price 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 buyers with unit demand. The main advantage of this setting with respect to others, i.e. a Cournot model of competition, is that it isolates the dynamics of adoption by assuming away static equilibrium distortions.

Within this framework, we first confirm formally 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 (discounted) social value, the adopting firm prefers the technology whose flow payoffs are received earlier. This equilibrium bias towards technologies with greater present payoffs is called the impatience property. Third, we prove that the bias embedded in the impatience property favors the adoption of technologies with smaller social value but flow payoffs that are delivered earlier in time.

The relevance of disruption costs in the introduction of products and processes is well-known in the management literature. Tyre and Hauptman (1992) list among their main causes the novelty of technical features, the low applicability of previous knowledge, and the incompatibility of current organizational practices with the arriving innovation.[1] Leonard-Barton (1988) shows that the adaptation of a technology often requires active cooperation between users and developers. Our model accommodates some of these features. Disruption costs may come not only from higher production costs but also from a lower valuation of the new good. In this case, trade between the buyers and the adopting firm improves the product, becoming a form of 'cooperation.'

Holmes, Levine, and Schmitz (2012) also study adoption in the presence of switchover disruptions. Their article contains an excellent discussion of the importance of disruption costs in a number of innovation episodes. 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.[2] Our insight is different as we stress that disruption costs open a future profit opportunity to competing firms. In an extension of our model, we also prove that, for some parameter values, adding non-adopting firms promotes 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 diminishes the rents of non-adopters.

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, Doraszelski, Kryukov, and Satterthwaite (2010). The goal of

[1] They also show that firms face significant disruption costs despite efforts in problem-solving prior to the introduction of technologies.

[2] Arrow (1962) 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 or disruption costs.

these articles is to understand how learning by doing, jointly with organizational forgetting in Besanko et al.'s article, 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.

Somewhat related to our work is Bergemann and Välimäki (2006), which studies the efficiency of price competition in a general dynamic framework. Our focus here is narrower and more applied. In particular, we are concerned with the adoption of better technologies in a specific dynamic setting—one in which the adopting firm suffers from disruption costs and the adopted technology gets better through learning by doing.

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 (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 exploit the idea that disruption costs are a source of future rents to non-adopters. This is the key distinctive feature of our work.

The remainder of this article is organized as follows. Section 2 presents the model. Sections 3 and 4 introduce some basic concepts and useful preliminary findings. Section 5 presents our main results. Section 6 concludes. In Appendix A we study a version of our model with an infinite number of buyers. Proofs are collected in Appendix B.

2. The model

We consider a quasi-linear utility economy with two sellers, denoted by $i \in \{1, 2\}$, and $T + 1$ buyers with unit demand. We write v_i for the buyers' valuation of a purchase from seller i and c_i for seller i's unit cost. With $s_i \equiv v_i - c_i > 0$ we denote the bounded flow surplus that is created when a buyer trades with seller i. Trading takes place over time: at each date only one short-lived buyer is available to trade with the sellers. Time is denoted by $t \in \mathrm{T} \equiv \{0, ..., T\}$. Sellers discount the future with a discount factor $\delta = 1$.

Technologies and efficiency. For clarity and to ease the exposition, we assume that only the first seller, seller 1, has the option of adopting a technology. For our purpose, a technology is completely specified by a function that associates cumulative sales up to the beginning of period t, x, with the surplus that can be created at that date if trade occurs. Formally, $s : \mathrm{X} \longrightarrow [\underline{s}, \overline{s}]$, where $\mathrm{X} \equiv \{0, ..., T\}$ and $0 < \underline{s} < \overline{s} < +\infty$. Any technology fulfills:

$$s(0) \leq s_2, \tag{A1}$$

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

The first assumption captures the notion of switchover disruption costs: at least the first sale made with a technology creates a weakly smaller flow surplus than the one created by seller 2. The second inequality represents learning by doing. We also assume that the socially efficient allocation requires the adoption of the technology. Formally, there is a minimum number of cumulative sales q in X such that the social value of the technology up to date q is nonnegative:

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

where $\varepsilon \geq 0$ is a sunk cost incurred at adoption.[3]

Our aim is to understand adoption decisions for a whole class of technologies: the set S of all functions s satisfying Assumptions A1–A3. Note that, for efficiency, every technology in S should be adopted at the initial date and trade should take place at each date using the adopted technology.[4]

Actions and payoffs. Adoption occurs within the framework of the following extensive-form game: At the beginning of each date, seller 1 decides whether or not to replace the old technology s_1 with a given s in S. The choice is irreversible. Then, the sellers simultaneously announce their (flow) surplus offers to the buyer: $b_i \equiv v_i - p_i$, where p_i is seller i's price.[5] The buyer then decides whether or not to buy from one of the sellers. The game continues this way until the last date is reached. The sellers have complete information about the history of the game.

Payoffs are as follows: Each buyer obtains either a zero payoff if trade does not take place or a payoff equal to b_i, the surplus offered by his trading partner. Seller i, at date t, receives a payoff either equal to zero if he does not trade, or equal to $s_i - b_i$ if he trades. For the sellers, total payoffs equal the sum of their flow payoffs.

Example 1. Let $T = 1$, $v_1 = v_2 = 1$, and $c_1 = c_2 = 0.5$. Suppose that the technology lets seller 1 produce a second unit at 0.1 after producing the first at 0.75. If seller 2 trades with the first buyer, the cost of seller 1 would still be 0.75 in the second period. In this case, seller 2 makes a profit of $0.75 - 0.5 = 0.25$ in the second period, which, moving backwards, leaves him willing to sell at any price $p_2 \geq 0.25$ in the first period. On the other hand, seller 1 makes at most

[3] With quasi-linearity and no lower bound in the supply of the money good an allocation is Pareto optimal if, and only if, it maximizes the sum of utilities. Assuming that there is a q in X such that the social value of the technology up to date q is nonnegative is the same as saying that the sum of all individual utilities is weakly higher with adoption than without it.

[4] In our economy, with infinitely inelastic demand curves, there is no room for static trade distortions. Hence, even if learning stops after seller 1 achieves q sales, allocative efficiency requires assigning all buyers to him.

[5] For tractability, we focus directly on surplus/money offers rather than on prices.

$0.5 - 0.1 = 0.4$ in the second period if he trades with the first buyer. This means he has to charge $p_1 \geq 0.35$ for the first unit. We conclude that seller 2 sells to both buyers upon adoption and, consequently, there is no adoption for any $\varepsilon > 0$.

Strategies and equilibrium. We describe here the strategies and the equilibrium concept for the sub-game that follows once a technology is adopted. Extending these to the entire game is straightforward but tedious. Besides, it would bring information of little use. As solution concept we use pure-strategy Markov perfect equilibrium (MPE for short) with x as the state variable. A state x in X is feasible if, for any t in T, $0 \leq x \leq t$. The history of the game at the beginning of date t, h_t, is the sequence of actions chosen by buyers and sellers before date t. Let $\mathrm{H_t}$ be the set of all possible histories at the beginning of date t and $\mathrm{H} \equiv \cup_{t\in\mathrm{T}}\mathrm{H_t}$ the set of all possible histories of the game with typical element h. A pure strategy for seller i is a real-valued function $b_i(\cdot) : \mathrm{H} \to \mathbb{R}$. A strategy $b_i(\cdot)$ is Markov if for each feasible state x and histories h, $h' \in \mathrm{H}$ of the same length, then $b_i(h, x) = b_i(h', x)$. In words, two different histories leading to the same x may only influence the behavior of the sellers through the date, t, at which x is attained. A Markov strategy for seller i is written as $b_i(x, t)$ and, for short, we call the pair (x, t), with x feasible, a state. A Markov perfect equilibrium is a sub-game perfect equilibrium in which the sellers use Markov strategies (see Maskin and Tirole, 2001).

3. Definitions

The functions we introduce here play a key role in our equilibrium analysis. We begin by displaying the incremental *flow* surplus, at each state (x, t), for a generic technology s adopted at $t' \leq t$. As the incremental flow surplus that s yields at (x, t) is independent of t, we denote it by $\pi(x) \equiv s(x) - s_2$. Note that $|\pi(x)|$ equals the Nash equilibrium payoff that seller 1 (seller 2) would obtain in a one-shot Bertrand game if $\pi(x) > 0$ ($\pi(x) < 0$).[6] Observe also that Assumption A2 implies that $\pi(x)$ is non-decreasing in x. We arrange these quantities in the *triangular array* A:

$$
\begin{array}{cccccc}
 & \pi(T) & & & & \\
 & \pi(T-1) & \pi(T-1) & & & \\
\text{Sales} & \vdots & \vdots & \ddots & & \\
 & \pi(1) & \pi(1) & \dots & \pi(1) & \\
 & \pi(0) & \pi(0) & \dots & \pi(0) & \pi(0). \\
 & & \text{Time} & & &
\end{array}
$$

where rows count cumulative sales and columns count periods (both running backwards). For any state (x, t), $A_{x,t}$ is the triangular sub-array whose last entry is in row $T + 1 - x$ and column $T + 1 - t$.

Definition 1 (d, r, and z)**.** Define functions d, r, and z on $\{(x, t) \in \mathrm{X} \times \mathrm{T} : x \leq t\}$ as follows: For any state (x, t), let $d(x, t)$ be the summation over the (outer) diagonal of $A_{x,t}$, let $r(x, t)$ be the negative of the summation over the last row of $A_{x,t}$, and let $z(x, t)$ be the sum of all entries of $A_{x,t}$.

The incremental surplus $d(x, t)$ can be written as:

$$d(x,t) = \sum_{k=0}^{T-t} [s(x+k) - \max_i \{s_i\}] + K(t),$$

where $K(t) \equiv (T + 1 - t)(\max_i\{s_i\} - s_2)$ is the equilibrium (continuation) payoff of the first seller without adoption.[7] (In what follows, let K be seller 1's reservation payoff at $t = 0$, $K(0)$, to lighten the notation.)

Two observations are in order. First, Assumption A3 implies $d(0, 0) - \varepsilon \geq K$, yet $d(x, t)$ may be either smaller than $K(t)$ or even negative if the remaining

[6] In this equilibrium sellers do not use weakly dominated strategies.

[7] The payoff in the Nash equilibrium in which both sellers offer $\min\{s_1, s_2\}$ at each date.

buyers cannot accommodate q sales. As a result, function z may take positive as well as negative values. Second, function z displays a sign preserving property: if $z(x,t)$ is positive then $z(x+1,t+1)$ is also positive. Similarly, if $z(x,t)$ is negative so is $z(x,t+1)$.

The triangular array of Example 1 is

$$\begin{array}{cc} 0.4 & \\ -0.25 & -0.25, \end{array} \tag{1}$$

which has $d(0,0) = 0.15$, $r(0,0) = 0.5$, and $z(0,0) = -0.1$.

4. Preliminary results

We begin this section by studying the dynamic competition sub-game that follows once a technology is adopted. We proceed then to characterize the adoption decision.

Lemma 1. *In a MPE,* $b_1(x,t) = b_2(x,t) \geq 0$.

Lemma 1 says two things. First, that in a MPE trade takes place at each date. Trade would not happen only if the maximum surplus offered to the corresponding buyer were negative. But then, as the flow surplus of each seller is bounded away from zero, a seller could deviate profitably and sell. Second, that in a MPE the buyer must be indifferent between the two offers; otherwise, the active seller could decrease the surplus he concedes to the buyer and increase his payoff.

In what follows, we focus attention on cautious MPE. In a cautious MPE (to ease notation, still MPE), the non-trading seller must be, at each state, indifferent between selling or not to the current buyer (see Bergemann and Välimäki, 2006, for a definition). Under this equilibrium concept seller 2 trades at $p_2 = 0.35$ with the first buyer in Example 1. Recalling that for a real-valued function f, $f^+(a) \equiv \max\{f(a), 0\}$ and $f^-(a) \equiv -\min\{f(a), 0\}$, we give:

Theorem 1 (Equilibrium payoffs)**.** *There is a unique MPE. In this equilibrium the payoffs of the sellers are:*

$$\pi_1(x,t) = \min\{d^+(x,t), z^+(x,t)\}, \tag{2}$$

$$\pi_2(x,t) = \min\{r^+(x,t), z^-(x,t)\}. \tag{3}$$

Theorem 1 provides an algorithm that resolves the intricacies of the dynamic competition game in a simple way. In particular, it shows that, for any state (x,t), it suffices to sum all flow surpluses in the corresponding triangular sub-array, i.e. to compute $z(x,t)$, to single out the trading seller. For example, from the triangular array (1) of Example 1 we get $\pi_1(0,0) = 0$ and $\pi_2(0,0) = 0.1$, meaning that seller 2 sells both units upon adoption (see Corollary 1 below).

The following observations throw light on the result. First, each seller obtains a non-negative payoff since he can always offer nothing. Thus, when the payoff of a seller is positive, his rival's payoff must be zero. For example, if $z^+(x,t) = z(x,t)$, then $\pi_2(x,t) = z^-(x,t) = 0$. Second, payoffs are sign monotone: if $\pi_1(x,t)$ is positive, then $\pi_1(x+1,t+1)$ is also positive. Similarly, if $\pi_2(x,t)$ is positive so is $\pi_2(x,t+1)$. This property is inherited directly from the sign preserving property of function z. We further interpret and use these implications below.

Corollary 1 (Monotonicity)**.** *In the unique MPE, if a seller sells at date t, then he makes all subsequent sales.*

One way to get intuition on Corollary 1 is to note that for any technology there is a minimal $\hat{x} \leq q$ such that $s(x) \geq s_2$ for all $x \geq \hat{x}$. If $x \geq \hat{x}$, the result is obvious. The cases in which cumulative sales have not yet attained threshold $\hat{x}$ are more involved, but, in essence, a seller will sell at date t only if he obtains a positive payoff. Then, as payoffs are sign monotone, his payoff from selling the next date will also be positive. Corollary 1 and Theorem 1 together yield:

Lemma 2 (No delay)**.** *For the unique MPE, adoption takes place only at the initial date.*

Lemma 2 says that adoption takes place without delay. The intuition is simple. When adoption occurs, the first seller, independently of the selected adoption date, receives full compensation of his reservation payoff. Hence, he does not get any benefit, and indeed loses money, from delaying adoption because the number of profitable trades may only diminish. This is particularly clear in Example 1, where missing the first sale makes the technology inefficient.

Corollary 1 and Lemma 2 have two important implications. First, adopted technologies attain their maximum social value. Second, the maximum adoption payoff is $\pi_1(0,0) - \varepsilon$. Therefore, the MPE is efficient if, and only if, the set S^* of technologies adopted in equilibrium fulfills:

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

Equivalently, the MPE is efficient if, and only if, $\pi_1(0,0) - \varepsilon \geq K$ for every technology in S. Our next result identifies a sufficient condition for efficiency.[8]

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

Without switchover disruption costs, efficiency ensues because the first seller gets, at least, the social value of the technology (i.e. equation (A3) with $q = T$). Formally, all entries in the triangular array A are non-negative without switchover disruption costs and, as a result, $z(0,0) \geq d(0,0) \geq \varepsilon + K$. Theorem 1 implies then that $\pi_1(0,0) = d(0,0)$ for every technology in S.

To fully appropriate the incremental surplus of the technology, the first seller must obtain, at each date t, a flow payoff equal to $s(t) - s_2$. That is, the maximum

[8] To ease the exposition, in what follows, we abuse our terminology and refer directly to the efficiency properties of the MPE.

surplus that seller 2 offers to the buyer must equal his flow surplus at each date. This 'bidding' behavior reflects the fact that, without switchover disruption costs, the continuation payoff of seller 2 is always zero as he cannot gain any market power.

5. Main results

This section contains our key findings. They come in three groups. First, we show that some efficient technologies are not adopted. Second, we describe the most distinctive features of these technologies. Third, we study the effect of adding more sellers.

5.A. Adoption breakdowns

To ease the exposition, we define the subset N of technologies as:

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

Clearly, set N is non-empty and is contained in S (see Figure 1). Our first key result generalizes the outcome of Example 1:

Proposition 2 (Adoption breakdown)**.** *The unique MPE is inefficient:*

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

Proposition 2 says that technologies in set N are not adopted. That is, the maximum adoption payoff $\pi_1(0,0) - \varepsilon$ is smaller than the reservation payoff for every technology s in N. Note that this is true even in the extreme case of a zero sunk cost, i.e. if ε is zero.

Our game is not a mere sequence of independent one-shot Bertrand games because a sale made by seller 1 improves his strategic position against seller 2

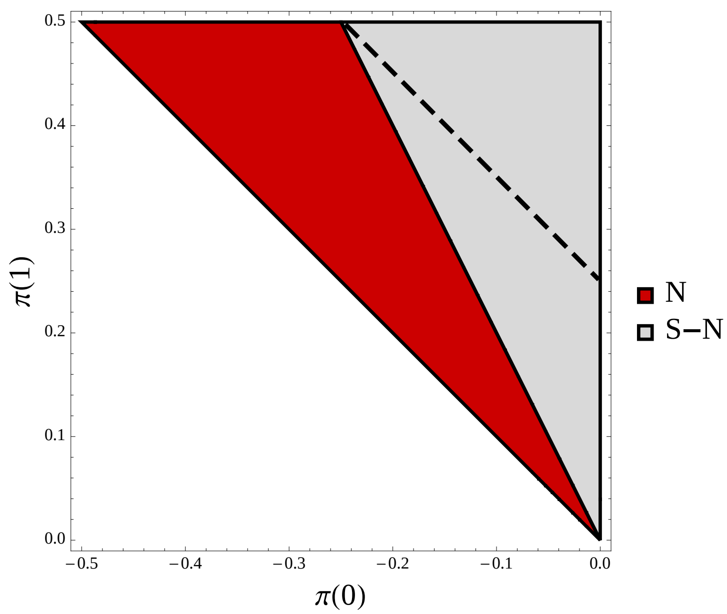

Figure 1: Sets N and $\mathrm{S}-\mathrm{N}$ of technologies with non-negative costs in Example 1, if $\varepsilon = 0$.

and vice versa. Both sellers are, in principle, willing to offer more than the flow surplus to the current buyer. How much do they offer depends, of course, on the return they expect from the improved strategic position. In our cautious equilibrium this means that the trading seller offers precisely his rival's valuation of an extra sale. The following proposition is meant to clarify this point and thus to improve our understanding of Proposition 2.

Proposition 3 (Recursive payoffs)**.** *In the unique MPE, payoffs obtain recursively from:*

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

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

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

Corollary 1 and Lemma 2 tell us we may observe one of two possible paths in equilibrium. Suppose that seller 1 adopts the technology—and sells at each date. Function $d(t,t)$ (from T to $\mathbb{R}$) gives the payoff of seller 1 if seller 2 plays at each date his one-shot Bertrand best response, i.e. if he offers at each date a surplus equal to s_2. Function $d(t,t)$ is thus an upper bound for the equilibrium payoff of seller 1. At each date, however, seller 2 may be tempted to deviate by offering more than s_2 to the current buyer. In fact, he would be willing to offer, at most:

$$u(t,t) \equiv s_2 + \pi_2(t, t+1). \tag{4}$$

That is, he would be willing to offer not only the buyer's intrinsic alternative social value, s_2, but also the continuation payoff he would earn from deviating at date t. Proposition 3 says that seller 1 must transfer to each buyer t in T a surplus equal to $u(t,t)$ to preclude deviations from the equilibrium path.

Therefore, the payoff function of seller 1 is just the incremental surplus of the technology minus the sum of the money payments he must transfer to the buyers to prevent deviations from the equilibrium path. If the total amount to be transferred is greater than the maximum appropriable rents, then the technology is not adopted—and the payoff from the dynamic competition game is zero for the first seller. A parallel argument explains the second equation in Proposition 3, i.e. the payoff function of the second seller. In particular, in Example 1 seller 1 must transfer $u(0,0) = 0.5 + 0.25$ to seller 2 to preclude a deviation in the first period. Still, he can transfer only $s(0) + \pi_1(1,1) = 0.25 + 0.4$. Following Proposition 3, the payoff of seller 2 in the first period may thus be written as $\pi_2(0,0) = \max\{0.5 - 0.4, 0\}$.

Less formally, as 'stealing' buyers from the first seller prevents the learning of the technology and improves the future market position of seller 2, buyers become an artificially overpriced 'commodity' in the market. Indeed, the mechanics of the equilibrium resemble the workings of a second price auction in which, to move forward in his preferred direction, seller 1 must transfer to each buyer a flow surplus equal to the valuation $u(t,t)$ of his rival. It is ultimately this social over-pricing of buyers what renders the adoption of better technologies unprofitable.

This also clarifies a key conceptual point. The root of the adoption failure is not the resistance of potential losers, but the market gains that non-adopters could obtain after adoption takes place. This is further highlighted in the next remark.

Remark 1. The adoption of a new technology may let seller 2 appropriate rents he could not appropriate otherwise. Consider Example 1. Following adoption, the second seller earns a payoff of 0.1, yet he makes zero profits without adoption. With zero sunk costs, this implies a transfer of wealth to seller 2, who benefits from the new situation without adding any social value. Moving one step backwards, we see these new rents constitute a form of resistance to adoption.

5.B. Endogenous impatience

The previous results lead us to question how can we tell adopted from non-adopted technologies apart. One might think that the distinctive feature of adopted technologies is a large social value. And although, as we shall see, this is not true in general, technologies with a sufficiently high social value are adopted. Formally, define the subset G of technologies as:

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

where $M_s \equiv \frac{1}{2}T(T+1)|\pi(0)|$. (Set G is the triangle above the dashed line in Figure 1.) Then, we have the following result:

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

For technologies outside G the social value rule is insufficient to decide whether adoption takes place. In particular, the inter-temporal distribution of the social value becomes crucial because technologies with larger early flow surpluses give higher adoption payoffs.

Example 2. Let $\varepsilon = 0$ and consider two technologies, s and s', with triangular arrays:

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

Although both technologies have the same incremental surplus $d(0,0) = d'(0,0)$ and switchover disruption costs $\pi(0) = \pi'(0)$, it follows from Theorem 1 that $\pi_1(0,0) = 0.5$, whereas $\pi_1'(0,0) = 0$.

To convey this idea formally we use a few more concepts. First, we say that technologies s and s' are surplus equivalent if, and only if, they have equal incremental surpluses. We denote by $[s]$ the equivalence class of technology s, that is, $[s] \equiv \{s' \in \mathrm{S} : d'(0,0) = d(0,0)\}$. Second, consider any technology s' in $[s]$ such that $\pi(0) = \pi'(0)$. We say that s is learned faster than s' during k sales if there is an $x \in X$ such that $\pi(x) \geq \pi'(x)$, $\pi(x+1) \geq \pi'(x+1)$,...,$\pi(x+k) \geq \pi'(x+k)$, with at least one strict inequality. Finally, we say that s is *learned earlier* than s', $s \succeq s'$, if s is learned faster than s' during the first k sales and s' is learned faster than s during the remaining $T-k$ sales. Geometrically, $s \succeq s'$ if technology s' crosses technology s at most once from below (see Figure 2).[9]

Proposition 5 (The impatience property)**.** *Let $s \succeq s'$. Then, seller 1's payoff $\pi_1(0,0)$ with technology s is weakly higher than his payoff $\pi_1'(0,0)$ with technology s'.*

The result shows that the total money payments that seller 1 must transfer to the buyers diminish as the technology is learned earlier. The proposition generalizes the following intuition.

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

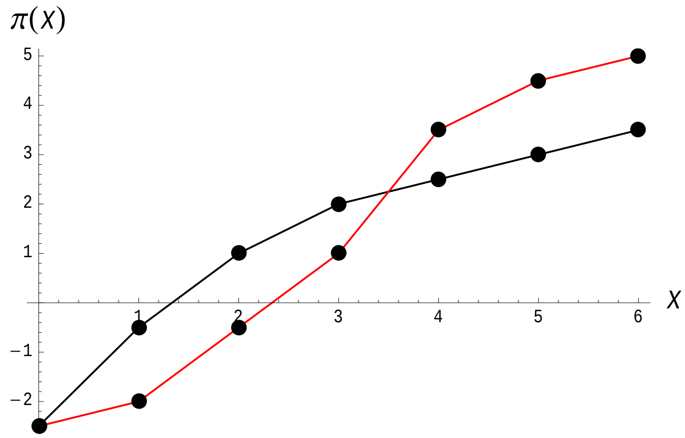

Figure 2: Two surplus equivalent technologies: technology s' (red) yields zero profits; technology s (black) is learned earlier and yields a positive payoff.

Consider a technology and perturb it by shifting a unit of surplus from the last date to date one. The cost of this perturbation is to reduce the last date surplus in a unit—a *static* cost of one unit. The benefit is, however, twofold. First, it increases the date one surplus in a unit. This *static* benefit just compensates the static cost, leaving the social value of the perturbed technology unchanged. Second, and key to the result, it (weakly) decreases the money payment that seller 1 must transfer to the date-1 buyer. This *dynamic* benefit ensues because, as seller 1 becomes more efficient at date one, the continuation payoff that seller 2 would earn from selling to the date-1 buyer diminishes.

There is indeed a more general illustration of the impatience embedded in equilibrium payoffs, though less clear-cut than the impatience property. The idea comes from understanding function z as a *discounted* sum of flow profits. For this purpose we write:

$$\frac{1}{T+1} z(0,0) = \sum_{t=0}^{T} \delta_t \, \pi(t), \tag{5}$$

where $\delta_t \equiv 1-(T+1)^{-1}t$ is read as an *endogenously* determined discount factor. The rule says that a technology is adopted only if the discounted value in (5) is weakly higher than $K+\varepsilon$. However, efficiency is measured by the non-discounted sum of flow profits (A3). As we have seen, this endogenous discounting has two consequences: first, a positive social value is not enough for adoption; second, technologies whose payoffs are received earlier have more chances of being adopted. These are neatly exemplified in Figure 1, where equivalence classes $[s]$ are parallel straight lines of slope -1. There, all classes below the dashed line can be partitioned into two convex subsets of adopted, and not adopted, technologies. 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 result shows that, if the first seller could choose between technologies, the present bias embedded in the impatience property may favor the adoption of technologies with smaller but 'better' inter-temporally distributed social value.

5.C. Adding more sellers

In this subsection we add a third seller to our model, but the reader shall be convinced that adding two or more sellers is straightforward. To be sure we make legitimate comparisons across models, we keep the set S of technologies fixed. For this purpose, recall that a third seller with flow surplus s_3 fulfills Assumptions A1 and A3 for all s in S if, and only if, $s_2 \leq s_3 \leq \max_i\{s_i\}$. Therefore, adding a third seller that fulfills these two inequalities leaves the reference set S of technologies unchanged.

We distinguish among two possible cases: $s_3 = s_2$, which always occurs if $s_1 \leq s_2$, and $s_2 < s_3 \leq s_1$. In the first case, neither seller 2 or seller 3 make profits in equilibrium because any strategy granting a positive payoff to one of them can be replicated by the other. In particular, note that none of them offers less than his flow surplus for states (x, t) with $\hat{x} - x > T - t$, which eliminates any incentive to offer more than just the flow surplus at earlier states. As a consequence, all technologies in S are adopted if $s_2 = s_3$.

In the second case seller 3 is more efficient than seller 2. This means that the second seller never sells in equilibrium, since seller 3 can bid higher for every buyer. In fact, seller 1 can be seen, without any loss of generality, as competing solely against seller 3 in this case.

There are two opposite effects in action when $s_2 < s_3$. The first effect is that seller 1 now faces a more efficient rival that might, in principle, bid more aggressively for the buyers. This is the *predatory* effect of competition which clearly hinders adoption. The second effect, which eases adoption, is that the amount seller 3 needs to offer to any buyer is now bounded below by s_2 (and $s_2 \geq s(x)$ for $x < \hat{x}$). This is the *protective* effect of competition that appears because, in equilibrium, seller 2 always offers his flow surplus. It turns out that which effect prevails depends on the magnitude of s_3: if s_3 is close to s_2 we approach the case $s_3 = s_2$ in which the protective effect dominates; as s_3 goes up the predatory effect gains importance.

The model with three sellers may be solved as a model with just two sellers, sellers 1 and 3, after updating the value function of seller 3 to:

$$\pi_3(x,t) = \max\{s_3 - \max\{\bar{b}_1(x,t), s_2\} + \pi_3(x,t+1), \pi_3(x+1,t+1)\}.$$

This expression simply takes into account that seller 3 now pays, at least, seller 2's bid—and not just the maximum bid of seller 1, $\bar{b}_1(x,t)$. Unfortunately, this model is not symmetric and there is no neat characterization as Theorem 1's in this case. However, we may make good use of the machinery developed above and consider an auxiliary model with a fictitious technology $\dot{s}(x) \equiv \max\{s(x), s_2\}$ for seller 1. A seller equipped with this technology never bids less than $\max\{\bar{b}_1(x,t), s_2\}$, and so his payoff $\dot{\pi}_1$ is an upper bound for seller 1's payoff. Proposition 7 formalizes our previous discussion.

Proposition 7 (The value of extra sellers)**.** *Fix a pair of flow surpluses fulfilling* $s_2 < s_1$*, a technology* s *in* S*, and add a third seller with flow surplus* s_3 *in* $[s_2, s_1]$*. Then:*

i. If $\pi_1(0,0) < d(0,0)$*, there is a threshold* $s_3^\dagger$ *in* $(s_2, s_1]$ *such that seller* 1*'s payoff goes up for all* $s_3 < s_3^\dagger$*.*

ii. If $s_1 - s_2 > s_2 - s(0)$*, there is a threshold* s_3^* *in* (s_2, s_1) *such that seller* 1*'s payoff goes down for all* $s_3 > s_3^*$*.*

Part ii of the proposition says that the predatory effect prevails if s_3 is large with respect to s_2 but also s_2 is close to $s(0)$. We should recall that adoption is easier the lower is the disruption cost $|\pi(0)| = s_2 - s(0)$ in the first place.

If we add a third seller in Example 1...

6. Conclusion

We have presented a dynamic model of technology adoption based on the idea that adoption creates socially spurious rents to 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 assessed the impact of adding more sellers obtaining mixed results. Nonetheless, we defend that a market in which sellers are closer substitutes of each other is more competitive and, under this view, our results show that competition has a positive value (Makowski and Ostroy, 2001). Summing up, our results should warn regulators of keeping an eye on industries with either little competition 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. In Appendix A we consider a model with infinitely many buyers. The discount factor δ may be any number within the interval $(0, 1)$. 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. Finally, the case in which both sellers may adopt a technology can also be treated.

Other generalizations require significant departures from our setup. Among these, introducing randomness is perhaps the most natural. We leaned in Corollary 1 that adopted technologies never fail. This feature of our model arises because learning curves are known with certainty. A model in which learning fluctuates randomly can easily incorporate the failure of adopted technologies.

References

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

Bergemann, D. and J. Välimäki (2006). Dynamic price competition. *Journal of Economic Theory 127*(1), 232–263.

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

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.

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

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.

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.

Makowski, L. and J. M. Ostroy (2001). Perfect competition and the creativity of the market. *Journal of Economic Literature 39*(2), 479–535.

Maskin, E. and J. Tirole (2001). Markov perfect equilibrium I. Observable actions. *Journal of Economic Theory 100*(2), 191–219.

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

Royden, H. (1988). *Real Analysis* (3rd ed.). Pearson.

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

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.

A. Large numbers of buyers

In this appendix we study markets with a large number of buyers. Our approach is to construct a continuous-time version of the model in which sales take place at the beginning of each of $T + 1$ periods of length $\Delta \equiv (T + 1)^{-1}$. We then let the number of buyers grow without bound and characterize the limit of the corresponding sequence of equilibria.

Define the following surplus function on the unit interval:

$$\pi^T(x) \equiv \sum_{k=0}^{T} \pi(k)\mathbf{1}_{[\Delta k, \Delta(k+1))}(x), \tag{6}$$

for x in $[0, 1)$, and $\pi^T(1) \equiv \pi(T)$. The interpretation is that $\pi^T(\Delta k)$ ($k = 0, \ldots, T$.) is the incremental flow surplus in state $(k, t) \in \mathrm{X} \times \{0, \ldots, \Delta T\}$ in a version of the model in which time runs in discrete steps of size Δ. From here it is straightforward to build a continuous-time version of the model with state-space $\{(x, t) \in [0, 1]^2 : x \leq t\}$ in which a sale of size Δ is made at each date $t_k = \Delta k$ ($k = 0, \ldots, T$.) to a buyer of life span $[t_k, t_{k+1})$. (Nothing happens when $t_k < t < t_{k+1}$.) We shall consider sequences of such continuous-time economies—with an increasing number of buyers.

Definition 2 (d^T, r^T, and z^T)**.** Define functions d^T, r^T, and z^T on $\{(x, t) \in$

$[0,1]^2 : x \leq t\}$ as:

$$d^T(x,t) \equiv \int_0^{1-t} \pi^T(x+y)\,dy,$$
$$r^T(x,t) \equiv (1-t)\pi^T(x),$$
$$z^T(x,t) \equiv \int_0^{1-t} [1-(t+y)]\pi^T(x+y)\,dy.$$

Functions d^T and r^T are continuous-time analogs of functions d and r in Definition 1, while z^T is a scaled-down analog of function z. Let $\{\pi^T\}$, $T = 1, 2, \ldots,$ be a sequence of surplus functions that converges pointwise to some function π^∞ on the unit interval. Define functions d^∞, r^∞, and z^∞ as the (pointwise) limits of functions d^T, r^T, and z^T as T goes to $+\infty$.[10] The next proposition characterizes the limit of the sequence of equilibria that corresponds to the sequence $\{\pi^T\}$, $T = 1, 2 \ldots,$ of surplus functions. Furthermore, it shows that, in the limit, there is no adoption without seller 1 fully appropriating the incremental surplus of the technology.

Proposition 8 (Large numbers payoffs)**.** *Consider the continuous-time model outlined above and let* $\{\pi^T\}$*,* $T = 1, 2, \ldots,$ *be a sequence of surplus functions converging pointwise to* π^∞*. Then, as* T *goes to* $+\infty$*, the payoffs of the sellers converge to:*

$$\pi_1^\infty(x,t) = \begin{cases} d^\infty(x,t) & \textit{if } z^\infty(x,t) > 0 \\ 0 & \textit{otherwise} \end{cases},$$
$$\pi_2^\infty(x,t) = \begin{cases} r^\infty(x,t) & \textit{if } z^\infty(x,t) < 0 \\ 0 & \textit{otherwise} \end{cases}.$$

Both adoption breakdowns and the impatience property are neatly illustrated

[10] Because $|\pi| \leq \bar{s}$, the limits exist by the Bounded Convergence Theorem (Royden, 1988).

in our large-numbers model. Let us write down function z^∞ at the initial state as:

$$z^\infty(0,0) = d^\infty(0,0) - \int_0^1 x\, \pi^\infty(x)\, dx. \tag{7}$$

Since the sign of $z^\infty(0,0)$ decides whether a technology is adopted in equilibrium, we see in (7) that a positive incremental surplus, $d^\infty(0,0) > 0$, is insufficient for adoption to take place. Likewise, suppose we are given two surplus equivalent technologies such that one is learned earlier than the other. Since the future is weighted heavier inside the integral and π^∞ is increasing, the technology which is learned earlier must have a higher $z^\infty(0,0)$.

In the limit, our continuous-time model approximates arbitrarily well an economy with a continuum of buyers distributed uniformly on the unit interval. In such economy, a technology is characterized by a non-decreasing function $\sigma : [0,1] \to [\underline{s}, \overline{s}]$ that fulfills $\sigma(0) \leq s_2$ and $\int_0^q \sigma(x)\, dx \geq q \max_i\{s_i\} + \varepsilon$ for some q in $(0,1]$. Function σ is the counterpart of function s, and gives the instantaneous flow surplus of the technology. The interpretation is that a new buyer shows up at each instant in the unit interval, though a share q of the market is needed for efficiency. The model generalizes in the obvious way to any interval $\mathrm{T} = [0,T]$ (with T in $\mathbb{R}_{++}$). A parametric example should help to fix the idea.

Example 3. Consider Example 1 with $\mathrm{T} = [0,1]$ and $\varepsilon = 0$. Suppose that the technology lets seller 1 produce at a cost $c(x) = \alpha x^{-\beta}$, with α in $(0, 0.5]$ and β in $[0,1)$.[11] (The power rule is the most common specification in empirical research; see Thompson 2010.) Let us be more specific and consider only technologies with total cost $\int_0^1 c(x)\, dx = 0.375$, which is the unitary cost of the technology in Example 1. This implies the parametric relation $\alpha = 0.375(1-\beta)$. A straightforward

[11] Function c requires $\underline{s} = -\infty$. This poses no technical problem because $\beta < 1$ ensures the convergence of the integral.

computation gives:

$$z^{\infty}(0,0)=\int_0^1 (1-x)[c_2-c(x)]\,dx=\frac{0.125-0.25\beta}{2-\beta}.$$

Therefore, technologies with $\beta < 0.5$ are adopted, whereas technologies with $\beta > 0.5$ are not. If $\beta < 0.5$ the first seller makes a profit of 0.125. Equilibrium prices are always constant and equal to 0.5. Note, also, that the equilibrium is the same as if seller 1 had $c(x) = 0.375$ if $\beta < 0.5$, and $c(x) = c_1$ if $\beta \geq 0.5$.

Our main results thus extend to a model with infinitely many buyers. The innovation is that we do no longer have predatory pricing in equilibrium: In the limit each buyer has a perfect substitute—who shows up an instant after him—and, for this reason, sellers do not need to undercut prices to win any particular sale.

B. Proofs

Proof of Lemma 1

The best response of a generic buyer is to buy from the seller who offers the highest non-negative surplus and not to buy otherwise. For notational ease, let us assume that seller 1 is the trading seller at state (x,t). Also, let us assume, momentarily, that trade happens at state (x,t). If the equality were not satisfied, seller 1 could decrease $b_1(x,t)$ by an infinitesimal amount and the buyer would still buy from him. Now we show that trade happens at state (x,t). Suppose not. Then, the highest flow surplus offered by the sellers must be negative. But as $s(\cdot)$ and s_2 are bounded away from zero, seller i can offer a surplus $0 < b_i(x,t) < s_i$ that is accepted by the buyer and gives him a strictly positive flow payoff.

Auxiliary results

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

$$\hat{x} \equiv \min\{x \in \mathrm{X} : s(x) \geq s_2\}.$$

Also, note that:

$$z(x,t) = \sum_{k=0}^{T-t} d(x, t+k), \tag{8}$$

$$= -\sum_{k=0}^{T-t} r(x+k, t+k). \tag{9}$$

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 (8) because $d(x,t+k)$ decreases with k. $(b1)$: Since, from (8), $z(x,t+1) = z(x,t) - d(x,t)$, $(a1)$ implies $(b1)$. $(c1)$: From (9), $z(x+1,t+1) = z(x,t) + r(x,t)$. 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 (9) because $r(x,t)$ is decreasing in t and non-increasing in x. $(b2)$ As in $(c1)$, from (9), $z(x+1,t+1) = z(x,t) + r(x,t)$ and thus $(a2)$ implies $(b2)$. $(c2)$ It follows from (8), as $d(x,t+k)$ decreases with k. $\square$

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 (9), $z(x, t+1) = z(x+1, t+2) - r(x, t+1)$, 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 (9), $z(x, t) = z(x+1, t+1) - r(x, t)$, 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 (8), $z(x+1, t+2) = z(x+1, t+1) - d(x+1, t+1)$, 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 (8), $z(x, t+1) = z(x, t) - d(x, t)$, 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 (9), $z(x, t+1) = z(x+1, t+2) - r(x, t+1)$, we have $f1$.

($f2$): $z(x+1, t+1) \geq 0$ implies, by ($a1$), that $d(x+, t+1) \geq 0$. Since, by (8), $z(x+1, t+1) = d(x+1, t+1) + z(x+1, t+2)$, we have the result. $\square$

We use throughout the following concepts. The value functions of the sellers 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 Seller i is willing to transfer to the buyer at state (x,t):

$$\bar{b}_1(x,t) = s(x) + \pi_1(x+1, t+1) - \pi_1(x, t+1), \tag{10}$$

$$\bar{b}_2(x,t) = s_2 + \pi_2(x, t+1) - \pi_2(x+1, t+1). \tag{11}$$

(Seller 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 sub-arrays $A_{x,t}$ for $x \in \mathrm{X}$, i.e. terminal states of the form $(\cdot, t)$ for which $d(\cdot,t) = -r(\cdot,t) = z(\cdot,t) = s(\cdot) - s_2$. Using the maximum bidding functions, payoffs, at any non-terminal state (x,t), in a MPE are:

$$\pi_1(x,t) = \max\{s(x) - s_2 + \pi_1(x+1,t+1) + \pi_2(x+1,t+1) - \pi_2(x,t+1), \\ \pi_1(x,t+1)\}, \tag{12}$$

$$\pi_2(x,t) = \max\{s_2 - s(x) + \pi_2(x,t+1) + \pi_1(x,t+1) - \pi_1(x+1,t+1), \\ \pi_2(x+1,t+1)\}. \tag{13}$$

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

(a): If $z(x+1,t+1) \leq 0$, we know from (d) in Lemma B, equations (2) and (3) 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 (12) and (13) 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 (2) and (3) 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 (12) and (13) 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 (2) and (3) 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 (2) 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 (12) and (13) 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 $\pi_1(\hat{x}-1,T) = \pi_2(\hat{x},T) = 0$, $\pi_1(\hat{x},T) = s(\hat{x}) - s_2$, and $\pi_2(\hat{x}-1,T) = s_2 - s(\hat{x}-1)$. Plugging these into (12) and (13) 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 sub-arrays $A_{x,t}$ 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, (10) and (11) that $\bar{b}_1(x,t) = s(x)$ and $\bar{b}_2(x,t) = s_2 + [r(x,t+1) - \pi_2(x+1,t+1)]$. Then, $\bar{b}_1(x,t) < \bar{b}_2(x,t)$ since $s_2 \geq s(x)$ and, from (a) in Theorem 1, $\pi_2(x+1,t+1) \leq$

$r(x+1,t+1) \leq r(x,t+1)$.[12] Thus, seller 2 will be the trading seller 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, (10) and (11) that $\bar{b}_2(x,t) = s_2$ 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.

(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, (10) and (11) that $\bar{b}_1(x,t) = s(x) + d(x+1,t+1)$ and $\bar{b}_2(x,t) = s_2 + z_2(x,t+1)$. 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 $\pi_1(x+1,t+1) = z(x+1,t+1)$.

Remark 2. We will use throughout the following notation. Recall that:

$$d(x,t) = w(x,t) + K(t), \tag{14}$$

where:

$$w(x,t) \equiv \sum_{k=0}^{T-t} [s(x+k) - \max_i \{s_i\}], \tag{15}$$

and $K(t) \equiv (T-t+1)\,[\max_i\{s_i\} - s_2]$.

[12] That $s_2 \geq s(x)$ follows from $\pi(x) \leq 0$.

Proof of Lemma 2

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:

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

Let $\Pi(0,t) \equiv \pi_1(0,t) - K(t)$. As S^* is non-empty, there is a date $t^* \in \mathrm{T}$ such that: (i) $\Pi(0,t^*) \geq \varepsilon$; and that: (ii) $\Pi(0,t^*) \geq \Pi(0,t)$ $\forall t \in \mathrm{T}$. 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, $\Pi(0,t) = 0$ $\forall t \geq 1$, and $t^* = 0$.

(b): If $\pi_1(0,0) = d(0,0)$ and if, $\forall t \geq 1$, $\Pi(0,t)$ is smaller than ε, the result holds trivially. Thus, let $\Pi(0,t) \geq \varepsilon$ for at least one $t \geq 1$. Then, by (14) and Theorem 1, we have that $\Pi(0,t) \leq w(0,t)$. And as, $\Pi(0,0) = w(0,0) > w(0,t)$ 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 $s \in \mathrm{S}$. As, from (14), $d(0,0) = w(0,0) + K$, it follows that $\Pi(0,0) = w(0,0) \geq \varepsilon$ for every $s \in \mathrm{S}$ by Assumption A3.

Proof of Proposition 2

Let s be an element of S^*, then $\pi_1(0,0) \geq K + \varepsilon$. This, in turn, implies that $z(0,0) \geq K + \varepsilon$ 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 $z(0,0) \geq K + \varepsilon$. 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 $d(0,0) = w(0,0) + K$ and $w(0,0) \geq \varepsilon$ 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 (3) that $\pi_2(x,t) = r(x,t)$. By $(c2)$ in Lemma A and (2) 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 (3), 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 (2) that $\pi_1(x,t) = d(x,t)$. By $(c1)$ in Lemma A and (3) 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 (2), we have the result.

Proof of Proposition 4

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

$$z(0,1) \equiv \sum_{k=0}^{T-1}(T-k)\pi(k).$$
$$\geq \sum_{k=0}^{T-1}(T-k)\pi(0) \equiv -M_s.$$

If s is an element of G, then $z(0,0) = d(0,0) + z(0,1) \geq d(0,0) - M_s \geq K + \varepsilon$. If $\pi_1(0,0) = z(0,0)$, then $s \in \mathrm{S}^*$. If $\pi_1(0,0) = d(0,0)$, then $s \in \mathrm{S}^*$ since $d(0,0) = w(0,0) + K$ and $w(0,0) \geq \varepsilon$ by Assumption A3. 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 $\Delta \equiv z'^{-}(0,1) - z^{-}(0,1)$. Then, as $s \succeq s'$, there is a $1 \leq k \leq (T-1)$ such that:

$$\Delta = \sum_{x=1}^{k}(T-x)\xi(x) + \sum_{x=k+1}^{T-1}(T-x)\xi(x). \tag{16}$$

$$\xi(1) + \cdots + \xi(k) + \xi(k+1) + \cdots + \xi(T) = 0, \tag{17}$$

where empty sums are taken to be zero, $\xi(x) \equiv (\pi(x) - \pi'(x))$, $\xi(x) \geq 0$ for $x \leq k$, and $\xi(x) \leq 0$ for $x \geq k+1$. Using (17) into (16), we have:

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

and the proof is complete.

Proof of Proposition 6

It suffices to show that there are technologies in N^c with social value arbitrarily close to zero. Simply consider a technology s_ϵ ($\epsilon > 0$) with $\pi_\epsilon(0) = 0$, $\pi_\epsilon(1) = \cdots = \pi_\epsilon(T) = \frac{K+\epsilon}{T}$, and sunk cost $\varepsilon = 0$. It has a social value equal to ϵ and:

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

Proof of Proposition 7

If a fictitious seller 1 has access to technology $\dot{s}$, with $\dot{s}(x) = \max\{s(x), s_2\}$ for $x \in \mathrm{X}$, he always bids more than s_2. If, in addition, $s_2 \leq s_3$, we may conclude that Seller 2 is irrelevant in this three-sellers model. We may then compute the profits of Seller 1, $\dot{\pi}_1$, as we do in the two sellers' model. Also, since $\dot{s}(x) \geq s(x)$ for all $x \in \mathrm{X}$, we have that the profits $\dot{\pi}_1$ of the fictitious Seller 1 are an upper bound for Seller 1's profits—both with three sellers. Therefore, it suffices to show that there is $s_3^* \in (s_2, s_1)$ such that $\dot{\pi}_1 \leq \pi_1$ for all $s_3 > s_3^*$.

From Theorem 1 is obvious that $\dot{\pi}$ *decreases continuously* with s_3. Let $s_3 = s_1$. Since in this case $s_3 - s_2 > s_2 - s(0)$ by assumption, we have that $\dot{\pi}(x) = \max\{s(x), s_2\} - s_3 < s(x) - s_2 = \pi(x)$ for all $x \in \mathrm{X}$. Then, by continuity and monotonicity with respect to s_3, there is $s_3^* \in (s_2, s_1)$ such that $\dot{\pi}_1 \leq \pi_1$ for all $s_3 > s_3^*$, and the proof is complete.

Proof of Proposition 8

Let π^T be a surplus function as defined in (6). Consider a discrete-time model with $\pi(y) = \Delta\pi^T(\Delta y)$, $y = 0, \ldots, T$. Quantities d, r, and z at state $(y, u) \in \mathrm{X} \times \mathrm{T}$

are:

$$
\begin{aligned}
d(y,u) &= \sum_{k=0}^{T-u} \Delta \pi^T(\Delta(y+k)),\\
r(y,u) &= (1-\Delta u)\pi^T(\Delta y),\\
z(y,u) &= \sum_{k=0}^{T-u} [1-\Delta(u+k)]\pi^T(\Delta(y+k)).
\end{aligned}
$$

Clearly, we have that $d(y,u) = d^T(\Delta y, \Delta u)$, $r(y,u) = r^T(\Delta y, \Delta u)$, and $z(y,u) = \Delta^{-1} z^T(\Delta y, \Delta u)$. The equilibrium payoffs of the sellers in the discrete-time model are given in Theorem 1:

$$
\begin{aligned}
\pi_1(y,u) &= \min\left\{\max\{d^T(\Delta y, \Delta u), 0\}, \max\{\Delta^{-1} z^T(\Delta y, \Delta u), 0\}\right\},\\
\pi_2(y,u) &= \min\left\{\min\{r^T(\Delta y, \Delta u), 0\}, -\min\{\Delta^{-1} z^T(\Delta y, \Delta u), 0\}\right\}.
\end{aligned}
$$

Therefore, the payoffs in the continuous-time model are simply:

$$
\pi_i^T(x,t) := \sum_{y=0}^{T} \sum_{u=0}^{T} \pi_i(y,u) \mathbf{1}_{[\Delta y, \Delta(y+1))}(x) \mathbf{1}_{[\Delta u, \Delta(u+1))}(t),
$$

for $i = 1, 2$. Since $\Delta^{-1} z^T(\Delta y, \Delta u)$ diverges to $\pm\infty$ as T goes to $+\infty$, whereas d^T and r^T converge to d^∞ and r^∞, the proof is complete.