Mergers and the Market for Organization Capital

André L. Faria∗†
The University of Chicago

November 17, 2002

Abstract

In this paper, mergers are an equilibrium outcome in which acquirers “marry” targets so as to gain access to their organization capital. Firms that are good at learning about a new technology are not necessarily those that can manage it well. If the organization capital is tradable, there should exist a market for it as there are gains from trade. The model generates a merger wave after deregulation or a shock to technology. I compare the implied wave to the actual merger wave in the telecom sector, which was deregulated just before the merger wave took place.

Keywords: Mergers, organization capital, merger waves, learning, technological change, deregulation

JEL: G34, L16, L96, O33.

∗Department of Economics. The University of Chicago. 1126 E. 59th Street, Chicago, IL 60637. Phone: 1-773-947-9408. E-mail: alfaria@uchicago.edu. URL: http://home.uchicago.edu/~alfaria.
†I thank Boyan Jovanovic for his advice and encouragement, and Fernando Alvarez, Robert E. Lucas Jr. and Luigi Zingales for very helpful comments and suggestions. I also thank Luís Braido, Pedro Carneiro, Jean-Pierre Dubé, François Gourio, João Hrotkó, Esteban Rossi-Hansberg, Mehmet Yorukoglu, and participants in the Federal Reserve Bank of Chicago brown bag seminar, in the VII LACEA annual conference (Madrid) and in informal seminars at the GSB and the Economics department of The University of Chicago for their comments. I acknowledge financial support from Fundação para a Ciência e Tecnologia, John M. Olin Foundation and The University of Chicago. All errors are mine.
1 Introduction

Evidence suggests that a plant’s productivity rises as a by-product of the production process.\textsuperscript{1} Evidence also shows that a discrete technological change is associated with a productivity drop early on, and then a gradual rise in productivity that eventually slows down.\textsuperscript{2} This suggests the existence of learning that is firm specific and tied to the technology as well.\textsuperscript{3} Prescott and Visscher (1980) call the stock of such learning organization capital.

A merger is one way that a firm can acquire the organization capital of another. The merger may arise if there is a gain to trading the organization capital. Gains arise if for some firm the cost of external acquisition of this capital is lower than the cost of internally developing it, while for other the long run benefits of the new technology are smaller than the gains it obtains from the transfer of its organization capital.

In the model, such gains to trade arise because managers differ in their ability. I assume a firm ran by a manager with high ability has larger profits than a firm ran by a low ability manager. The new technology imposes reduction in the profits in the short run, though it leads to higher profits in the long run. Hence the opportunity cost of adoption of the technology at an early stage is higher for the firm with high managerial skills. However, the firm with the best manager is the one that gains more from the new technology when it is “mature”. Therefore, a firm with low managerial skill has a comparative advantage at developing and learning about a new technology, because it loses relatively less. But once the technology is “mature” this firm is not the one that can benefit most of it, so there are gains from trade and room for a market.

As in Holmes and Schmitz (1990), in equilibrium there is a specialization of functions, with some firms innovating and others acquiring their businesses. However, in this paper I disregard issues of entry and exit, and focus on the transition process instead of the stationary equilibrium. As a consequence of a technological or deregulatory shock, mergers

\footnotesize{\textsuperscript{1}Bahk and Gort (1993).}
\footnotesize{\textsuperscript{2}Klenow (1998) briefly surveys the literature that presents this evidence.}
\footnotesize{\textsuperscript{3}Irwin and Klenow (1993) show that a firm learns three times more from its own cumulative production than from other firm’s cumulative output.}
occur clustered in time - merger wave - and they are an equilibrium outcome of a model in which acquirers “marry” targets to have access to their organization capital.

The main contribution of this paper is the dynamic matching model I use to explain merger waves and the attempt to derive endogenously the characteristics of the participants in the merger. I show how a model like this explains some of the empirical findings about mergers: mergers occur in waves, the $q$ (ratio of market value to the replacement cost of capital) of the acquirer tends to be larger than that of the target, the best takeovers in terms of value creation are those in which a high $q$ firm takes over a low $q$ firm, the combined value of the firms rises after merger. I also show that this model delivers time series implications for the ratio of the $q$ of target to the $q$ of the acquirer and for the ratio of productivities of target and acquirer: both ratios must rise during the merger wave.

The rest of the paper is organized as follows. In section 2, I briefly review the related literature. In section 3, I present the model and in section 4 I characterize the equilibrium and I do some comparative statics. In section 5, I study the planner’s problem and I present an equivalence result between the market equilibrium and the planner’s solution. In section 6, I show some examples that help to understand the dynamics of the model. In section 7, I present the empirical implications and in section 8 I make a brief empirical assessment of the model looking at the merger wave of the 90s in the telecommunications industry. In section 9, I conclude.

## 2 Literature review

In this section I briefly review some of the new empirical findings and theories related to the merger activity.

One of the most relevant observations is the time and industry clustering in merger activity (Mitchell and Mulherin 1996; Andrade, Mitchell and Stafford 2001; Andrade and Stafford 2002). The fact that merger activity clusters through time is not new (Nelson 1959; Gort 1969); however, the industrial clustering of merger activity, i.e., that waves come not only over time but by industry, is new and seems to favor the interpretation that merger waves are the result of industry-level shocks.
Mergers create value (Andrade, Mitchell and Stafford 2001; Andrade and Stafford 2002) and the creation of value that results from mergers is higher if the acquirer has a high $q$ (ratio of market value to replacement cost of capital) and the target has a low $q$ (Lang, Stultz and Walkling 1989; Servaes 1991). Microstudies also suggest that newly acquired plants see their productivity improved (Lichtenberg and Siegel 1990; Schoar 2000). Other evidence suggests that the $q$ of acquirers is higher than that of targets in the majority of the cases and acquirers are larger than targets (Andrade and Stafford 1999; Andrade, Mitchell and Stafford 2001).

Interpreting $q$ as a measure of managerial ability, this empirical evidence suggests that mergers are a way of reallocating resources from low skilled managers to high skilled managers (as argued by Manne 1965), and that the value created is larger, the larger are the differences between the skills of the managers of the acquirer and the target.

Jovanovic and Rousseau (2002 a, b) develop a model that accounts for several of these empirical findings. In their model, the only objective of takeovers is to acquire physical capital, i.e., mergers are simply used capital trades. This approach totally disregards the issue of the identity of the firms merging, it overlooks the fact that mergers are like marriages and that a firm is more than just a collection of physical assets.4

An alternative equilibrium explanation of mergers is developed by Toxvaerd (2002). He develops a model in which acquirers (imperfectly) compete over time for scarce targets. This scarcity of targets leads to strategic behavior and in equilibrium all the potential acquirers bid for the targets simultaneously. The fact that mergers are clustered in a moment in time depends crucially on the assumption of the scarcity of targets. Moreover, the fact that the sets of targets and acquirers are not an equilibrium result, but a given of the model, does not help to understand the characteristics of the the participants in a merger.

Matsusaka (2001) develops a dynamic matching model to study conglomerate mergers and divestitures. He shows that diversification can be a value maximizing strategy for a firm with high managerial skills. A firm with high managerial skills experiments until it

---

4Between 1997 and 2001, The Economist published at least 60 articles in which mergers were compared to marriages, some of them with colorful titles as “Hold my hand” or “The runaway bride”.

4
finds the business that best matches its abilities. A good match leads to specialization. Matsusaka focuses his attention on the so called diversification discount and not in the merger wave.

Arrow (1983) suggested that small firms have an advantage in pursuing revolutionary innovations, but they do not have access to the capital markets that would allow them to implement and further develop their innovations once they reach maturity. Large firms, on the other hand, are less able to develop revolutionary innovations but they have the financial means that allow them to implement the mature innovations created by the small firms. Hence, there are gains from trade and mergers must take place. This is the only paper I know that tries to derive endogenously the characteristics of participants in a merger: targets are small are acquirers are large. However, Arrow does not try to explain why mergers tend to occur in waves.

My paper contributes to this literature. The basic idea is similar in spirit to that of Arrow (1983). I develop a model in which a technological or deregulatory shock leads to a merger wave. Mergers occur as an equilibrium outcome as acquirers marry targets to have access to the organization capital the latter firms developed after the shock.

3 The model

I assume the horizon is infinite and time is continuous. The agents of this economy are firms. Firms are fully characterized by their managerial ability $m \in M = [m_l, m_h]$. There is a distribution $\Gamma(m)$ of firms. The profit function of each firm is $\pi(m) = m$.\footnote{If no factors are used at production, profit and output coincide. However, one could see $\pi(m) = m$ as a reduced form of a problem of the type $\pi(m) = \max_I m^\alpha t^{1-\alpha} - w t$, as described by Lucas (1978). The profit function is $\pi(m) = \left( \frac{\alpha}{\alpha} \right)^{\frac{1}{\alpha}} (1 - \alpha) m$ for all $m$. If one drops the common term $\left( \frac{\alpha}{\alpha} \right)^{\frac{1}{\alpha}} (1 - \alpha)$ the profit function is $\pi(m) = m$.}

At some date (date 0), a technological shock occurs. This technological shock translates in the emergence of a new technology. Firms can choose between the old and the new technology. If firm $m$ chooses the new technology, the profits $t$ periods after the
adoption is:

$$\pi(t, m) = A(t) m,$$

(1)

where $A(t)$ is the learning accumulated after $t$ periods. I assume $A(t)$ is a strictly increasing and strictly concave function of $t$, $0 < A(0) < 1$ and learning is bounded \( \lim_{t \to \infty} A(t) = \bar{A} > 1 \). These assumptions imply the new technology leads to higher profits in the long run, but it is not free. To adopt, firms must incur an opportunity cost (foregone profits). The particular form of $\pi(\cdot, \cdot)$ treats learning independent of the managerial ability.\(^7\)

I call $A(t)$ organization capital: it is a capital that is embedded in the firm and cannot exist outside it. Prescott and Visscher (1980) thought of it as the knowledge about the matching between workers or between workers (skills) and tasks. Atkeson and Kehoe (1997, 2002) followed the path of Arrow (1962) and Rosen (1972) and thought of this organization capital as the learning the firm accumulates over time. Bahk and Gort (1993) present empirical evidence that shows that an important component of learning is what they call organization learning. They conclude there is a component of learning-by-doing whose returns are not captured by labor, i.e., there is a component of learning that is firm specific and “enters into what we call the firm’s stock of organization capital”.

The interpretation of $A(t)$ deserves some attention. $A(t)$ is a form of capital accumulated by the firm $t$ periods after the adoption of the new technology. This capital is closely tied to the firm and does not exist without it: it can be knowledge about the best organizational procedure, the marketing and development of the product or the technology itself. The learning is firm specific, i.e., there are no spillover effects in the learning process.

I assume the learning is only transferable through the transfer of ownership of the

\(^6\)Jovanovic and Nyarko (1996) show how that a function of the type of $A(t)$ is a reduced form of learning in a model with Bayesian learning.

\(^7\)Human capital is not unidimensional (for models in which human capital has more than one characteristic see, for example, Heckman and Scheinkman 1986). In this case, it is as if all individuals are equally good at engineering, but they differ in the managerial skills.
One crucial assumption is that $A(t)$ is not a public good. If a firm wants to benefit from $A(t)$, it can only do so either by adopting the technology at its onset or by buying it through the acquisition of an early adopter later on.

Other assumptions concern the way the transfer occurs. First, I assume that to each seller corresponds only one buyer, that is, the organization capital can only be sold to one firm. This assumption is a consequence of the private nature of $A(t)$ and does not look very restrictive as it seems consistent with reality: the target is bought by a single firm.\footnote{I am not the first one to think of this form of transaction of knowledge. Rosen (1972) mentions that specific knowledge “vested in the firm” is transferable by selling the firm. He also provides several examples.} \footnote{The main results of the paper still hold if only $\delta A(t), \ 0 < \delta < 1$, is transferable. I provide the results upon request. In Appendix A I show the case when organization capital is non-tradable ($\delta = 0$).}

Figure 1: Learning
acquirer, not by many. Second, once a firm has \( A(t) \), it does not gain anything by acquiring any other firm with the (same) organization capital. This implies that a buyer acquires only one firm because any additional organization capital is redundant.

In what follows I assume that if the organization capital were non-tradable every firm would be better off adopting the new technology, that is, I add the assumption \( r \int_{0}^{\infty} e^{-rt} A(t) \, dt > 1 \) to the set of assumptions I have made about this learning function. Therefore, a firm has two alternatives, either develops the organization capital on its own or it buys one firm that has developed it.

At the time there is a technological revolution, the firm must decide whether to be a (potential) seller or a buyer. A seller is a firm that develops a technology, learns about it and after it acquires some level of knowledge about it, sells itself to a buyer (or not). A buyer is a firm that only works with “mature” technologies; it does not make major innovations, but applies and slightly develops them after buying the knowledge other firms accumulated about it. All firms behave competitively.

### 3.1 Buyer’s problem

The value of being a buyer \( m \) includes the decision of when to buy the new technology from an early adopter \( m' \):

\[
V_B^B (m, m') = \max_{t^d} \int_{0}^{t^d} e^{-rt} md t + \int_{t^d}^{\infty} e^{-rt} A(t) \, dt - e^{-rt} q (m', t^d).
\]  

(2)

The value of the buyer is the sum of the discounted earnings with the old technology plus the discounted earnings with the new technology minus the discounted price paid by the organization capital. I assume the buyer takes the management of the newly formed firm. Though this is not always the case, it is the most typical.\(^{10}\) Equation (2) is the mathematical expression of Gort’s statement: “The managers of the acquiring firm (usually the larger firm) exercise control over the combined properties after mergers.

\(^{10}\)Martin and McConnell (1991) show the CEO of the target firm is four times more likely to be replaced in the year after the takeover than in the previous years (the hazard rate passes from 11\% to 46\%). In 56\% of the cases, the reason for change in management is “change of control”.

8
generally take the initiative in making a purchase offer, and ordinarily retain the corporate identity of the acquiring firm”. Sellers differ in the price they are willing to accept to sell the organization capital at date $t$: $q(m', t)$ reads as the price seller $m'$ is willing to accept on the organization capital at date $t$. Hence, from all the sellers $m'$ the buyer chooses the one that maximizes its value, that is, the problem of buyer $m$ is then:

$$V^B(m) = \max_{m'} \left\{ \max_{t^d} \int_0^{t^d} e^{-rt} mdt + \int_{t^d}^{\infty} e^{-rt} A(t) mdt - e^{-rt} q(m', t^d) \right\}.$$ (3)

This problem can be re-interpreted as

$$V^B(m) = \max_{t^d} \int_0^{t^d} e^{-rt} mdt + \int_{t^d}^{\infty} e^{-rt} A(t) mdt - e^{-rt} \min_{m'} \left\{ q(m', t^d) \right\}. \quad \text{(4)}$$

The buyer does not care with the managerial ability of the seller $m'$ as this is not used in the newly formed firm. Therefore, a firm that buys organization capital at date $t$ does not care about the identity of the seller, and so it only looks at the least expensive organization capital in the market at that date, at the quoted market price of organization capital at date $t$: $q(t) = \min_{m'} q(m', t)$. That is, the price schedule $q(\cdot)$ describes the market prices of the non-zero transactions. This approach mimics that of Rosen (1974) in his paper on hedonic prices and implicit markets. As Rosen did, I am only concerned with the observed price function and I disregard all the other possible prices for organization capital at date $t$ that do not have transactions associated with them. The problem with which I am concerned is

$$V^B(m) = \max_{t^d} \int_0^{t^d} e^{-rt} mdt + \int_{t^d}^{\infty} e^{-rt} A(t) mdt - e^{-rt} q(t^d). \quad \text{(5)}$$

and I assume $q(\cdot)$ belongs to the class of functions twice continuously differentiable. Buyers and sellers behave competitively in the sense that they take this price schedule $q(\cdot)$ as given, that is, their individual decisions of buying and selling do not affect this function. Buyers decide the date at which they acquire organization capital, taking as given the quoted market price of it. Solving this problem for each buyer I get a unique
relation between the identity of buyers and the date at which they decide to acquire the sellers.

The first order condition for this problem is\(^ {11}\)

\[
(1 - A(t^d)) m - q'(t^d) + rq(t^d) = 0. \tag{6}
\]

This equation implicitly defines \(t^d = \psi(m):\)

\[
(1 - A(\psi(m))) m - q'(\psi(m)) + rq(\psi(m)) = 0. \tag{7}
\]

The inverse function of \(\psi(\cdot)\) is defined because \(\psi'(m) < 0 \, ^{12}\)

\[
m^b = \psi^{-1}(t) = b(t). \tag{8}
\]

The function \(b(\cdot)\) gives the identity of the buyer as a function of the time at which the merger occurs.

### 3.2 Seller’s problem

The value of being a seller includes the decision of the date, \(t^s\), when (if ever) to sell the organization capital and, therefore, the firm:

\[
V^S(m) = \max_{t^s} \int_{0}^{t^s} e^{-rt} A(t) mdt + e^{-rt^s} q(m, t^s). \tag{9}
\]

The value of this firm is the sum of the discounted earnings plus the discounted price of the firm at date \(t\). organization capital.

As I argued above, the buyer only cares with the organization capital he buys at date \(t\), not with the identity of the seller. If the seller \(m\) sells the firm at date \(t\) it means it is the lowest price for organization capital at that date. That is, if the transaction occurs

\(^{11}\)The assumption \(r \int_{0}^{\infty} e^{-rt} A(t) \, dt > 1\) rules out the cases in which the buyer chooses never to buy the new technology, so the solution is always interior.

\(^{12}\)The first order condition defines a maximum only if the second order condition is
\[e^{-r\psi(m)} \left[ (rA(\psi(m)) - A'(\psi(m)) - r)m - r^2q(\psi(m)) + 2rq'(\psi(m)) - q''(\psi(m)) \right] < 0.\] I anticipate the equilibrium result \(A(\psi(m)) > 1\) (later, I show this is indeed the case) and then
\[\psi'(m) = -\frac{1-A(\psi(m))}{(rA(\psi(m)) - A'(\psi(m)) - r)m - r^2q(\psi(m)) + 2rq'(\psi(m)) - q''(\psi(m))} < 0.\]
at date $t$, $q(t) = q(m, t)$, otherwise $q(t) < q(m, t)$ and the seller is not able to sell the organization capital at this date.

For $t^*$ to be the optimal decision, the first order condition is

$$e^{-rt^*} [A(t^*) m + q'(m, t^*) - rq(m, t^*)] \geq 0.$$  \hfill (10)

For the cases in which the solution is interior, $t^* < +\infty$, $q(m, t^*) = q(t^*)$,

$$e^{-rt^*} [A(t^*) m + q'(t^*) - rq(t^*)] = 0.$$  \hfill (11)

This is as if the problem faced the seller was

$$V^S(m) = \max_{t^*} \int_0^{t^*} e^{-rt} A(t) m dt + e^{-rt^*} q(t^*).$$  \hfill (12)

As I said above, buyers and sellers behave competitively in the sense that they take this price schedule $q(\cdot)$ as given, that is, their individual decisions do not affect this function. Sellers decide the date at which they sell the firm, taking as given the observed price of the organization capital. Solving this problem for each seller I get a unique relation between the identity of sellers and the date at which they decide to sell.

The first order condition of this problem, equation (11), implicitly defines $t^* = \phi(m)$:

$$A(\phi(m)) m + q'(\phi(m)) - rq(\phi(m)) = 0.$$  \hfill (13)

As $\phi'(m) > 0$, the inverse function is (locally) defined

$$m^* = \phi^{-1}(t) = s(t).$$  \hfill (14)

The function $s(\cdot)$ gives the identity of the seller as a function of the date at which the transaction occurs.

\footnote{The first order condition only defines a maximum if the second order condition is $e^{-r\phi(m)} [(A'(\phi(m)) - rA(\phi(m))) m + r^2q(\phi(m)) - 2rq'(\phi(m)) + q''(\phi(m))] < 0$. Hence, $\phi'(m) = -\frac{A'(\phi(m)) - rA(\phi(m)) m}{A(\phi(m)) - rA(\phi(m)) m + r^2q(\phi(m)) - 2rq'(\phi(m)) + q''(\phi(m))} > 0$.}
3.3 Firm’s decision

Therefore, the problem for any firm $m$ is

$$V(m) = \max \{ V^S(m), V^B(m) \}. \tag{15}$$

This formulation does not rule out two extreme cases: the case in which a seller decides not to sell the organization capital, and the case in which a firm sticks to the old technology. However, the assumption $r \int_0^\infty e^{-rt} A(t) \, dt > 1$ rules out this last case.

4 Equilibrium

I am looking at the market for organization capital. In the previous subsections I described all the elements necessary for the definition of an equilibrium in this market: the problem faced by sellers and by buyers and how firms select in one of these categories. I only care with the equilibrium price of observed transactions, that is, with the price schedule $q(\cdot)$. Therefore, the relevant definition of equilibrium is as follows:

**Definition 1** An equilibrium is a price function $q(\cdot)$, sets $B = \{ m \in M \mid V^B(m) \geq V^S(m) \}$ and $S = \{ m \in M \mid V^B(m) < V^S(m) \}$ and a pair of functions $t^s(\cdot), t^d(\cdot)$ such that: (i) prices clear the market at every $t$: $Q^s(t) = Q^d(t)$ where $Q^s(t) = \int_{\{m \in S\} \land t^s(m) = t} d\Gamma(m)$ and $Q^d(t) = \int_{\{m \in B\} \land t^d(m) = t} d\Gamma(m)$; (ii) $t^s(m) = \arg \max_{t^s} \int_0^{t^s} e^{-rt} A(t) \, dt + e^{-rt} q(t^s), t^d(m) = \arg \max_{t^d} \int_{t^d}^\infty e^{-rt} A(t) \, dt - e^{-rt} q(t^d)$ and $V(m) = \max \{ V^S(m), V^B(m) \}$.

The firms that lose less from the adoption of the technology at its onset are those with low managerial ability. The ones that gain most from the adoption of the technology at a “mature” stage are those with a high managerial ability. This suggests there is a threshold level of managerial ability that separates sellers from buyers, where the former have managerial ability lower than the threshold and the latter have managerial skills above the threshold.
Lemma 1 There is a unique $\bar{m}$ such that $V^S(\bar{m}) = V^B(\bar{m})$; $V^S(m) > V^B(m)$ for $m < \bar{m}$, and $V^S(m) < V^B(m)$ for $m > \bar{m}$.

This lemma is the formal expression of the intuition I presented in the lines above. It tells that the set of sellers and buyers are defined by a unique threshold level that divides the interval of managerial abilities in sellers and buyers.

Lemma 2 There is an “inaction” region. All firms with $m \in (\bar{m}, \hat{m})$ adopt the new technology at date zero but choose never to sell it.

Lemma 2 tells that not all the firms that adopt the new technology at date zero are willing to sell it. That is, it is only worth transferring the technology if the differences in managerial abilities between the seller and the (potential) buyer are large enough.
Otherwise, there are no gains from trade, as what the gains to the potential buyer do not offset the costs of transfer (foregone profits) the seller would incur. This result is important as casual observation of reality suggests that not all the firms that adopt a technology are taken over by other firms that use a more obsolete technology: some of these early adopters stand alone and grow.

In equilibrium, $\forall t < +\infty$, buyer $b$ and seller $s$ merge at time $T(b, s)$, where $T(b, s)$ uniquely solves

$$\frac{b}{s} = \frac{A(t)}{A(t) - 1}. \quad (16)$$

The function $T(b, s)$ is homogeneous of degree zero in its arguments, so $T(b, s) = T\left(\frac{b}{s}, 1\right)$ and $T(x, 1)$ is a decreasing function of $x$. Therefore, the last transaction $\bar{t} = T\left(\frac{m - m}{m}, 1\right)$ is the one in which $\underline{m}$ is the seller and $\bar{m}$ is the buyer. Analogously, the first mergers occur between $m_l$ and $m_h$ at date $t = T\left(\frac{m_h}{m_l}, 1\right)$.

Hence, for $t \leq t \leq \bar{t}$, this equation together with the market clearing condition

$$\Gamma(s(t)) = 1 - \Gamma(b(t)) \quad (17)$$

delivers

$$\Gamma(s(t)) + \Gamma\left(\frac{A(t)}{A(t) - 1}s(t)\right) = 1. \quad (18)$$

From this last equation I unambiguously get $s(t)$ and so $b(t)$. Plugging $s(t)$ in

$$e^{-rt}A(t)s(t) = -p'(t) \quad (19)$$
and integrating with respect to \( t \) emerges
\[
p(t) = \int_t^\infty e^{-r\tau} A(\tau) s(\tau) d\tau + C. \tag{20}
\]
Imposing the “boundary” condition that the last seller \( m \) is the one that is just indifferent between selling at \( \bar{t} \) and not selling at all\(^{14}\), the current price function is
\[
q(t) = e^{rt} p(t) = \int_t^\infty e^{-r(\tau-t)} A(\tau) s(\tau) d\tau. \tag{21}
\]
for \( t \leq t \leq \bar{t} \), where \( s(\tau) = m \) for \( \tau \geq \bar{t} \). In summary:

**Proposition 1** The price function is \( q(t) = \int_t^\infty e^{-r(\tau-t)} A(\tau) s(\tau) d\tau \), where \( s(\tau) \) is the identity of the seller at date \( \tau \leq \bar{t} \) and \( s(\tau) = m \) for \( \tau > \bar{t} \), where \( \bar{t} = T \left( \frac{\bar{m}}{m}, 1 \right) \).

Prices must be such that in each moment, the market clears, i.e., the mass of firms that sold the organization capital by date \( t \) must equal the mass of firms that bought the organization capital by this same date. Trade stops at date \( \bar{t} \) and the mass of firms \( \Gamma \left( \bar{m} \right) - \Gamma \left( \frac{m}{\bar{m}} \right) \) does not merge as gains from trade only exist if the difference between the managerial ability of the potential buyer and the potential seller is “big enough”.

The result of this model presents a very strong similarity with the hedonic price model of Rosen (1974) with a single characteristic. The characteristic in which agents trade is time. The model is falsely dynamic as all the decisions are made at date zero and the price function is defined as of that date. As in Rosen’s model, the price function is the joint envelope that unites the tangency points between the continuation value of buyers and sellers.\(^{15,16}\) One difference with respect to Rosen’s paper is that he focuses his

\(^{14}\) Otherwise, there will be a seller \( m + \varepsilon, \varepsilon > 0 \), that is still better off by selling at a finite time, but this violates the fact that \( m + \varepsilon \) must belong to the inaction region. Therefore \( \int_0^\bar{t} e^{-r\tau} A(\tau) m d\tau + p \left( \bar{t} \right) = \int_0^\infty e^{-r\tau} A(\tau) m d\tau, \) and so \( C = \int_\bar{t}^\infty e^{-r\tau} A(\tau) m d\tau. \)

\(^{15}\) The tangency condition in my model is analogous to the smooth pasting condition of the optimal stopping time problem in stochastic calculus literature (see Dixit and Pindyck (1994)).

\(^{16}\) In equilibrium, a buyer and a seller are perfectly matched when their respective value and offer functions “kiss” each other, with the common gradient at that point given by the gradient of the market.
attention on the case in which the range of values of the characteristic, and the identities of sellers and buyers are exogenously determined, whereas in this model the range and the identities are endogenously determined.

It only remains to see who merges with whom. From the first order conditions of the problems of sellers and buyers I obtain

\[ \phi(m^*) = t, \]  
\[ m^b = b(t). \]  

Putting the two equations together I obtain an assignment rule that “maps” sellers into buyers:

\[ m^b = b(\phi(m^*)), \]  
\[ m^b = a(m^*), \]

where \( m^* \in [m_L, m] \) and \( a = b \circ \phi \). The derivative of this function with respect to its argument is negative: \( a'(m) = \phi'(m)b'(\psi(m)) < 0 \), because \( \phi(\cdot) > 0 \) and \( b'(\cdot) < 0 \). In equilibrium the market clearing condition holds, and so the assignment rule \( a(\cdot) \) must satisfy

\[ \Gamma(m) + \Gamma(a(m)) = 1, \forall m \in [m_l, m], \]  

with \( m_h = a(m_l) \) and \( \bar{m} = a\left( \frac{m}{m} \right) \). Summarizing:

**Proposition 2** The assignment rule that “maps” sellers to buyers is \( b = a(s) \), such that 
\[ \Gamma(m) + \Gamma(a(m)) = 1, \forall m \in [m_l, m], m_h = a(m_l) \text{ and } \bar{m} = a\left( \frac{m}{m} \right). \]

\[ clearing \text{ implicit price function } p(z). \text{ Therefore, observations } p(z) \text{ represent a joint envelope of a family of value functions and another family of offer functions” } Rosen (1974). \]
Market clears at each moment

Lemmas 1 and 2 say that low $m$ firms are the early adopters of the new technology and that high $m$ firms are the buyers. This proposition says how they match. The matching starts by the extremes, with the lowest ability firms $m_l$ selling the organization capital
to the highest ability ones $m_h$. This is the intuition behind this result: once the new technology is “mature” low $m$ firms are the ones who gain less from it, while high $m$ firms are the ones that gain most. The urgency to buy the new technology as soon as it arrives to its “maturity” is higher for those firms with a high managerial ability; the willingness to sell it at early dates immediately after the “maturity” is higher for the adopters that can gain less from this technology, the ones with low managerial ability.

4.1 Changes in the distribution function

In this subsection I study the sensitivity of the equilibrium to the distribution function of managerial ability. I look at a very intuitive case in which the new distribution function, $\hat{\Gamma}(\cdot)$, first-order stochastically dominates$^{17}$ the original distribution function, $\Gamma(\cdot)$:

$$\Gamma(x) \geq \hat{\Gamma}(x), \forall x.$$  \hfill (27)

In equilibrium, for $\forall t < +\infty$ it must be that

$$\frac{b(t)}{s(t)} = \frac{A(t)}{A(t) - 1},$$

independently of the distribution function of managerial ability. However, the functions $s(\cdot)$ and $b(\cdot)$ are not independent from it, as $s(\cdot)$ must satisfy the market clearing condition for all $\underline{t} \leq t \leq \bar{t}$:

$$\Gamma(s(t)) + \Gamma\left(\frac{A(t)}{A(t) - 1}s(t)\right) = 1.$$  \hfill (28)

For the new distribution function this market clearing condition becomes

$$\hat{\Gamma}(\hat{s}(t)) + \hat{\Gamma}\left(\frac{A(t)}{A(t) - 1}\hat{s}(t)\right) = 1.$$  \hfill (29)

$^{17}$The distribution function $\hat{\Gamma}(\cdot)$ first order stochastically dominates $\Gamma(\cdot)$ if $\forall$ increasing function $v(\cdot)$, $\int v(x) d\Gamma(x) \geq \int v(x) d\Gamma(x)$. The expression in the text is not the definition itself but the result of a proposition. However, I use it as if it was the definition because it states more clearly the properties of the distribution functions.
If there are two economies identical in everything but the distribution function, the ratio of managerial abilities of acquirer and target in each moment in time are equal, but in the economy with relatively small supply of low quality managers (i.e., the one with distribution function $\hat{\Gamma}(\cdot)$), the abilities of acquirer and target in each moment in time are at least as large as in the other economy. This happens because $\Gamma(x) \geq \hat{\Gamma}(x)$, $\forall x$, and so $\hat{s}(t) \geq s(t)$, for all $t - \leq t \leq \bar{t}$.

The consequence on prices is fairly intuitive. Recall from Proposition 1 that $q(t) = \int_t^\infty e^{-r(\tau-t)}A(\tau)s(\tau) d\tau$. Then, for the economy with relatively more supply of high quality managers it must be that $\hat{q}(t) = \int_t^\infty e^{-r(\tau-t)}A(\tau)\hat{s}(\tau) d\tau$. This lemma is a consequence of the discussion above:

**Lemma 3** If $\hat{\Gamma}(\cdot)$ first-order stochastically dominates $\Gamma(\cdot)$, then $\hat{q}(t) \geq q(t)$, for $\forall t - \leq t \leq \bar{t}$.

Lemma 3 says the price function of the economy with $\hat{\Gamma}(\cdot)$ is bounded below by the price function of the economy with $\Gamma(\cdot)$. Why is this? To say that $\hat{\Gamma}(\cdot)$ first-order stochastically dominates $\Gamma(\cdot)$ means that in the economy with $\hat{\Gamma}(\cdot)$ there are relatively more high quality managers than in the original economy, so the relative demand for organization capital is higher and consequently the equilibrium price of it must rise. Or, in a different way, the firms that supply organization capital in the $\hat{\Gamma}(\cdot)$ economy have higher managerial ability than in the original economy, in any moment in time, and so the price of organization capital rises because the opportunity cost of these firms is higher.

One can think of distribution functions that first order stochastically dominate $\Gamma(\cdot)$, but for which no changes occur in the price function. This happens if the changes in the distribution function occur in the interval $\left(\underline{m}, \underline{m}\right)$, that is, there are many distributions that generate the same price function $q(t)$, as long as the mass in the tails is the same. Suppose the distribution function $\hat{\Gamma}(\cdot)$ satisfies the following condition:

**Condition 1** Either $\exists y$ such that $\hat{\Gamma}(y) < \Gamma(y)$, $y \in [m_l, m_l]$, or $\exists z$ such that $\hat{\Gamma}(z) < \Gamma(z)$, $z \in [\bar{m}, m_h]$. 

19
Then, the proposition follows from the discussion above:

**Proposition 3** If \( \hat{\Gamma} (\cdot) \) first-order stochastically dominates \( \Gamma (\cdot) \) and satisfies Condition 1, then \( \exists t \) such that \( \hat{q} (t) > q (t) \).

### 4.2 Changes in managers’s outside options

Up to now I assumed the targets’ managers have no outside options or the value of it is null. In this subsection I study how the equilibrium outcome changes once I consider the existence of a positive outside option to the target’s manager.

I now assume that the manager can always find an alternative occupation that pays him a wage \( w (m) = wm, \) \( 0 < w \leq 1 \). Then the seller’s problem becomes

\[
\hat{V}^S (m) = \max_{t^s} \int_0^{t^s} e^{-rt} A (t) m dt + e^{-rt^s} \left( \hat{q} (t^s) + \frac{wm}{r} \right).
\]  

(30)

How does this affect the equilibrium? This has two effects in the equilibrium, one in the “intensive” margin and the other in “extensive” margin: it affects the way firms do the matching and it also changes the universe of firms that were matching. Naturally, this reflects in the price function itself.

I start looking at the “intensive” margin, that is, I look at the way buyers and sellers match, assuming that the universe of buyers and sellers is the same as in the original economy. The assignment function for \( m \in [m_i, m] \) remains unchanged (no changes in the distribution function) but the matching timings \( \hat{T} (b, s) = \hat{T} (\frac{b}{s}, 1) \) are now shorter than before, \( \hat{T} (\frac{b}{s}, 1) < T (\frac{b}{s}, 1) \):

\[
\frac{b}{s} = \frac{A \left( \hat{T} \left( \frac{b}{s}, 1 \right) \right) - w}{A \left( \hat{T} \left( \frac{s}{s}, 1 \right) \right) - 1}
\]

(31)

This happens because the opportunity cost of the target’s manager is now lower, so he is willing to transfer the property of the organization capital sooner than before.

The changes in the “extensive” margin are consequence of this reduction in the opportunity cost of the target’s manager. This enables managers with closer managerial
abilities to be in opposite sides of trade, that is, the “inaction” region is now smaller than in the original economy. Therefore, the assignment function is “extended” to \( \hat{m} > m \),

\[
\Gamma (m) + \Gamma (a(m)) = 1, \quad \forall m \in \left[ m_l, \hat{m} \right],
\]

(32)

where \( \hat{m} \) is computed in an analogous way as for the original economy.

The price function remains unaffected in the common range of timings of mergers, i.e.,

\[ q(t) = \hat{q}(t) \]

in that range.\(^\text{18}\)

One interesting case is when the targets’s managers have access to the old technology \( w = 1 \). In this case, the timing of matching is the same for all mergers, independently of the ratio of managerial abilities between acquirer and target, i.e., \( A(T(x, 1)) = 1, \forall x \). Therefore, the price function is a point in the space, i.e., it is only defined for \( t = t_1 \) where \( A(t_1) = 1 \). However, an “inaction” region still exists.\(^\text{19}\)

5 Planner’s problem

In the previous section I looked at this problem from the market perspective. Now I study the model using the planner’s approach. This helps to understand the mechanics of the model and at the same type it unveils the welfare properties of the market solution. The question I try to answer in this section is the following: if there were a planner maximizing the aggregate discounted profits of the industry would he choose the same pattern of mergers the market delivers?

To better understand this problem, I divide the planner’s problem into two stages. At the first stage, the planner divides the firms in two categories: they are either sellers (firms that adopt the technology at an early stage and might decide to sell it to some other firm) or buyers (firms that decide to do not immediately adopt the technology and prefer to acquire it via acquisition of some other firm). The second stage is an assignment

\(^\text{18}\)The “neutrality” result comes from the fact that \( \hat{s}(\tau) = s(\tau) \frac{A(\tau)}{A(\tau) - w} \).

\(^\text{19}\)Because \( \frac{\int_0^\infty e^{-\tau(A(\tau)-1)}d\tau}{\int_0^\infty e^{-\tau(1-A(\tau))}d\tau} > 1 \).
or matching problem in which the planner chooses who “marries” whom and when in order to maximize the social gain. I solve this problem beginning with the second stage.

5.1 The second stage of the planner’s problem

In this stage, the planner maximizes the gains of matching \( b \in B \) with \( s \in S \) taking as given the set of buyers \( B \) and the set of sellers \( S \) (defined at the first stage). The function

\[
G = \max_t G(t)
\]

with \( G(t) = \int_0^t e^{-rt} (1 - A(t)) bd\tau - \int_t^\infty e^{-rt} A(\tau) sd\tau \), describes the gains of matching \( b \) and \( s \). The function \( G(t) \) is the gain of matching \( b \) and \( s \) at an arbitrary date \( t \), it is the difference between the discounted profits of firms \( b \) and \( s \) if they merge at date \( t \) and the discounted profits of these same firms if both were to adopt the new technology at its onset and therefore they would not merge.

If \( A(t) \) is strictly concave and \( b > s \), then \( G(\cdot) \) is single peaked with an interior maximum. The intuition is that the marginal gain to merge decreases with time since learning decreases with time.

The first order condition of problem (33) is

\[
G'(t) = e^{-rt} \{(1 - A(t)) b + A(t) s\}.
\]

I show that for each \( b > s \) there is a unique \( T^*(b, s) \) that solves \( G'(t) = 0 \),

\[
\frac{b}{s} = \frac{A(t)}{A(t) - 1}
\]

since \( \frac{A(t)}{A(t) - 1} \) is decreasing in \( t \), if \( A(\cdot) \) is strictly concave. Moreover, \( T \) is homogeneous of degree zero in \( b \) and \( s \). So that

\[
T^*(b, s) = T^*\left(\frac{b}{s}, 1\right)
\]

and \( T(x, 1) \) is a decreasing function of \( x \). Finally, \( T > t_1 \), where \( t_1 \) solves \( A(t_1) = 1 \), since

\[
\frac{b}{s} = \frac{A(T^*)}{A(T^*) - 1} < \frac{A(t_1)}{A(t_1) - 1} = \infty.
\]
Now I argue that $G(\cdot)$ is single peaked. For that it suffices to show that at $t = T^* (b, s)$, $G'' (T^*) < 0$. This follows since

$$G''(t) = -re^{-r t} \{(1 - A(t)) b + A(t) s\} - e^{-r t} A'(t) (b - s)$$  \hspace{1cm} (38)$$

which at $t = T^*$ is $G''(t) = -e^{rT} (b - s) A'(T^*) < 0$.

Therefore,

$$G(T^* (b, s)) = \int_0^{T^* (b, s)} e^{-r \tau} (1 - A(\tau)) b d\tau - \int_{T^* (b, s)}^{\infty} e^{-r \tau} A(\tau) s d\tau.$$  \hspace{1cm} (39)$$

or

$$sG(T^* (x, 1)) = s \left\{ \int_0^{T^* (x, 1)} e^{-r \tau} (1 - A(\tau)) x d\tau - \int_{T^* (x, 1)}^{\infty} e^{-r \tau} A(\tau) s d\tau \right\}$$  \hspace{1cm} (40)$$

The planner decides to match $b$ with $s$ as long as the gains from matching are positive, i.e., if $G(T^* (x, 1)) \geq 0$.

**Lemma 4** There exists a unique $x_0$ such that $G(T^* (x_0, 1)) = 0$, for $x > x_0$, $G(T^* (x, 1)) > 0$, and for $x < x_0$, $G(T^* (x, 1)) < 0$.

For values of $b$ and $s$ such that there are gains from matching, i.e., $\frac{b}{s} \geq x_0$, the function $G(T^* (b, s))$ has negative cross derivative $\frac{\partial^2 G(T^* (b, s))}{\partial b \partial s} < 0$, so the optimal assignment between buyers and sellers must be what Becker (1973) called negative assortive mating. This monotone matching pattern is independent of the distribution function though the identity of those who merge it is not: $b = f(s)$, with $f'(s) < 0$.\footnote{Legros and Newman (forthcoming) cover matching issues extensively. My model belongs to the class of models for which the utility is fully transferable, without market imperfections and where the “segregation payoff” is zero, just like in Becker’s model.} The function $f(\cdot)$ is (implicitly) defined by:

$$\Gamma(s) = 1 - \Gamma(f(s)) , \forall s \in S ,$$  \hspace{1cm} (41)$$

with

$$f'(s) = - \frac{d\Gamma(s)}{d\Gamma(f(s))} < 0.$$  \hspace{1cm} (42)$$
5.2 The first stage of the planner’s problem

The planner now decides the sets $B$ and $S$. In the second stage, the objective is to maximize the gains from the match, taking as given the identities of buyer and seller. In the first stage the planner chooses in which side of the match a firm will be. As I described before, the gain of matching $b$ with $s$ is:

$$G(T^*(b, s)) = \int_0^{T^*(b, s)} e^{-r\tau} (1 - A(\tau)) b d\tau - \int_{T^*(b, s)}^\infty e^{-r\tau} A(\tau) s d\tau. \quad (43)$$

The matching between buyer and seller must respect the “resources constraint”, that is $b = f(s)$. Plugging $b = f(s)$ the gains of matching are a function of $s$ alone:

$$G(T^*(f(s), s)) = \int_0^{T^*(f(s), 1)} e^{-r\tau} (1 - A(\tau)) f(s) d\tau - \int_{T^*(f(s), 1)}^\infty e^{-r\tau} A(\tau) s d\tau. \quad (44)$$

The matching is worth doing as long as $G(T^*(f(s), s))$ is positive. The threshold $m^*$ such that $G(T^*(f(m^*), m^*)) = 0$ defines the universe of buyers and sellers: $S = [m_L, f(m^*)]$ and $B = (f(m^*), m_H]$, and all those $m \in [m^*, f(m^*)]$ are early adopters that never merge with other firms.

Without surprise, there is an equivalence between the market equilibrium and the planner solution, that is, $T(x, 1) = T^*(x, 1)$, $\forall x$, $m = m^*$ and $a(m) = f(m^*)$ for $\forall m \in [m_l, m]$. This equivalence result summarizes in the following proposition:

**Proposition 4** The market equilibrium solution and the planner solution are equivalent.

Hence, in this model there are not as “too many mergers” or wrong timings to merge. The market delivers exactly the same outcome as would a planner whose objective function was the sum of the discounted profits of the firms.

---

21 It must be clear that $\frac{f(m^*)}{m^*} = x_0$. This value is independent of the distribution function. The particular values of $m^*$ and $f(m^*)$ are not independent from it.

---
6 Examples

In this section I illustrate with examples some of the properties of the model. One of the difficulties of this model concerns the fact that its results are derived from a characteristic that is unobservable to the researcher: the managerial ability. Though there are some proxies for this characteristic (profit, size, $q$, productivity, etc.), it is difficult to establish reasonable values for the managerial ability, as some of the proxies may vary non-linearly relation with the true ability. This makes the calibration of the model hard and difficult to defend. The simulations that follow help to understand the dynamics of the model. For the sake of illustration, I assume the managerial ability ranges between $m_l = 1$ and $m_h = 10$.

Following Parente (1994), I assume a specific functional form for the learning curve

$$A(t) = \tilde{A} - \exp(-kt).$$

I assume that the new technology is slightly better than the old one $\tilde{A} = 1.501$, $k = 0.07$ (taking the case in which organization capital is non-tradable as a reference), and that $r = 0.07$. In what follows, I work with the benchmark case presented in which the targets’s managers do not have outside options.

6.1 Example 1 - Mass of $m_l$ managers

To solve this model and to get some more explicit results I need to be willing to do some explicit assumptions about the distribution of managerial abilities. I start with the easiest (and also more uninteresting) case, the one in which there is an infinitely elastic supply of firms with low managerial ability, i.e., there are as many of $m_l = 1$ firms as “needed”. By assuming a mass of firms with the lowest managerial ability, I am transforming this general equilibrium model in a partial equilibrium model.

Technically, this assumption defines clearly the set of targets, $m = 1$, the set of firms that decides to adopt and not to sell, $m \in (1, \bar{m})$, and the set of acquirers $m \in [\bar{m}, m_H]$.

--

$^{22}$Parente (1994) assumes the same type of learning curve but in his model $\tilde{A} = 1$. 
The price function is as defined in section 4: \( q(t) = \int_t^\infty e^{-r(\tau-t)} A(\tau) d\tau \forall t. \)

As before, this model can be interpreted as a hedonic price model, but now with an infinitely elastic supply. In this case the price function is equal to the cost function (so the price schedule is non-linear), and the buyers choose the amount of characteristic (in this case, time) according to a tangency condition. Though all sellers are identical, mergers occur at different timings because of heterogeneity on the buyers side.

For this set of parameters, it takes approximately 13 years and a half after the introduction of the new technology until the merger wave takes place and this lasts approximately 13 years. The firms with managerial ability between 1 and 3.9279 adopt the technology immediately. The firms with the highest managerial ability (10) start acquiring the ones with the lowest managerial skill (1) approximately 13 years and a half after the adoption and when the new technology is already becoming “mature”. The others with managerial ability lower than 10 but “high enough” follow them.

\[ \text{It is easy to check that this price function effectively delivers the result I claim above, with } \bar{m} = \frac{\int_{\sigma(m)}^{\infty} e^{-r\sigma} A(t) dt}{\int_0^{\sigma(m)} e^{-r(1-A(t))} dt}. \]
6.2 Example 2 - Uniform distribution

In this example I keep assuming that \( m_l = 1 \) and \( m_h = 10 \) but now I assume the managerial ability is uniformly distributed within this range.

As I described above, the timing of mergers is a function of the ratio of managerial abilities of acquirer and target only. The timing of mergers is therefore the same as in the above example. The difference now is who matches with whom. In the example above, the sellers were the lowest managerial ability firms and they were indifferent about the time of selling. Now, the sellers are the firms with managerial ability in the range \([1, 2.322]\) and the buyers are the ones with managerial abilities in the interval \([8.7678, 10]\). The ratio of \( \bar{m} \) to \( m \) is the same in the two examples, because this does not depend on the distribution functions but only on the learning curve and on the interest rate \( \frac{8.7678}{2.322} = 3.79279 = \frac{3.9279}{1} \). The distribution functions matter to determine the values of \( \bar{m} \) and \( m \).

The particular shape of the merger wave is a consequence of the assumptions on the distribution function and on the functional form of the learning: even though the distribution function of managerial abilities is uniform, the merger wave is not uniform over time because the relation between time and identity of the sellers is non-linear.
Figure 6: Timing of mergers

Figure 7: Merger wave
6.3 Example 3 - Uniform distribution with faster learning

In this example I keep assuming that the distribution of managerial ability is uniform but now the learning about the new technology is faster, i.e., $k = 0.1 > 0.07$.

![Faster learning](image8.png)

**Figure 8: Faster learning**

![Comparing merger waves](image9.png)

**Figure 9: Comparing merger waves**

There are two effects. First, the value of the new technology relatively to the old one
is higher because the foregone profits are now smaller; this makes some of the high ability firms less likely to be buyers because the cost of internally developing the technology is now smaller.

Second, the learning curve has a steeper slope that makes early acquisition more attractive, so the merger wave happens earlier and more concentrated than before.

7  Empirical implications of the model

In this section I derive some empirical implications from this model. I divide them in two categories: cross section and time series. In the last category I look at the implications for the dynamic behavior of some of the variables, while in the first one I do not consider the dynamic element.

7.1  Cross section

The literature documented some stylized facts about mergers:

1. The $q$ of the acquirer is larger than that of the target (Andrade, Mitchell and Stafford 2001; Andrade and Stafford 2002).


3. The best takeovers in terms of value creation are those in which a high $q$ firm takes over a low $q$ firm (Lang et al. 1989; Servaes 1991).

The literature has interpreted the market-to-book ratio (or Tobin’s $q$) in very different ways. Some thought of it as a measure of monopoly rents (Lindenberg and Ross 1981); others interpreted it as a measure of intangible capital, either R&D and patents (e.g., Griliches 1981, Megna and Klock 1993, Hall 1993) or managerial ability (Lang, Stultz, Walkling 1989). My model does not have physical capital. Therefore, the value of the firm coincides with the value of its intangible capital, that consists in the managerial ability and the organization capital. Hence, my model gives theoretical content to the ratio of
market value to the replacement cost of physical capital. Defining \( V(m, t) \) as the current value at date \( t \) of a firm with managerial ability \( m \), the \( q \) of it is \( q(m, t) = V(m, t) \).

Under what conditions does this model deliver the stylized fact 1? The \( q \) of the target is smaller than that of the acquirer if \( b(t) > 2 \int_t^\infty \omega(\tau) s(\tau) d\tau \), that is, if the managerial ability of the acquirer in any given moment is at least twice as large as the weighted average of the managerial ability of the targets from that moment on.24 Though I cannot get a general proposition of when this occurs, I am able to identify some special cases. If \( \frac{m_h}{m_l} > 2 \) then the \( q \) of the acquirer is larger than that of the target when the merger wave starts. If \( \frac{m}{m} > 2 \) then the \( q \) of the acquirer is always larger than that of the acquirer.25

Andrade, Mitchell and Stafford (2001) document that in 66% of the mergers, the \( q \) of the acquirer is higher than that of the target. This has clear implications on the technology and the distribution of the managerial abilities.26

In this economy there are no surprises. This is a perfect foresight economy, so all the matches and their value are perfectly known by the moment the new technology arrives. I compare the joint value of target and acquirer in two different economies: one in which firms can merge and other in which firms would like to merge but they are not allowed to do so. In this model, the larger the difference in the managerial abilities of the acquirer and the target, the greater is the value creation of the merger.27 This is consistent both with stylized facts 2 and 3 (there are no mergers if mergers do not create value) and, in a way, with the micro evidence that shows that the plants acquired see their productivity increased (Lichtenberg and Siegel 1990, Schoar 2000).

**Proposition 5** The best takeovers in terms of value creation are those with larger differences in managerial ability.

24 The weight \( \omega(\tau) \) is \( \frac{e^{-\tau}A(\tau)}{\int_0^\infty e^{-\tau}A(\tau) d\tau} \) and so \( \int_0^\infty \omega(\tau) d\tau = 1 \).

25 The value of the ratio \( \frac{m_h}{m_l} \) depends on the distribution and the learning functions together. This is why it is hard to get a general statement about the ratio of \( q \)'s. In the examples I cover in the previous section, the \( q \) of the target is always smaller than the \( q \) of the acquirer.

26 It implies that \( \int_0^{m_h} \frac{dV(m)}{dV(m)} = \frac{2}{\hat{t}} \), where \( \hat{t} \) is such that \( b(\hat{t}) = 2 \int_0^\infty \omega(\tau) s(\tau) d\tau \).

27 This follows from \( \frac{\partial}{\partial m} \left\{ m \int_0^\infty e^{-\tau}A(\tau) d\tau - \int_{\psi(m)}^\infty e^{-\tau}A(\tau)(m + s(\psi(m))) d\tau \right\} > 0 \).
7.2 Time series

This model has the ability to explain the transitional dynamics that occur as a consequence of the technological (or deregulatory) shock. The most known stylized fact is that mergers occur clustered in time or in waves. This model reproduces this fact: the merger wave starts \( t \) periods after the introduction of the new technology. The particular shape of the wave is related with the distribution \( \Gamma(\cdot) \) and with the function \( A(\cdot) \).

As I described above, the firms that first takeover are those with higher managerial ability and the targets are those firms with lower managerial skills. Over time, the managerial ability of the acquirer decreases while that of the target increases. At the same time, the level of organization capital traded increases. This translates in the following proposition (re-interpretation of proposition 2):

**Proposition 6** The ratio \( \frac{m^a(t)A(t)}{m^b(t)} \) rises during the merger wave.

If one is willing to interpret \( m^a(t)A(t) \) as the productivity of the target and \( m^b(t) \) as the productivity of the acquirer, this proposition says that the relative productivity of target to acquirer must rise during the merger wave. In the empirical assessment I discuss this in more detail.

The model predicts a similar pattern for the market-to-book ratio:

**Proposition 7** The ratio of the \( q \) of the target to the \( q \) of the acquirer increases over time during the transition.

8 Empirical assessment

In this section I try to see up to what point this model can explain some of the merger waves that took place in the 90s. The model at use is a model in which technological shocks lead to merger waves. This makes difficult any empirical assessment as some of

\[\text{\footnotesize Jovanovic and Rousseau} \ (2002 \ b) \text{ also derive this same empirical implication from their model. However, their set-up imposes that this ratio converges to one, whereas mine imposes no restrictions about it.}\]
the technological innovations are not well defined in time or even well defined themselves. Therefore, as Andrade et al. (2001) I look at deregulation as the main industry-specific shock and as the possible starting point of the merger wave. I argue that deregulation is a valid “starting point” because:

1. It is well defined in time.29

2. In some industries, it mimics a technological shock, enabling the adoption of technologies that before were not allowed.

3. Sometimes it happens as a reaction to technological innovations.30

In the 90s, four big deregulations took place: utilities (1992), banking (1994), telecommunications and broadcasting (1996). From all these industries, the one that seems to fit more in the arguments I gave is the telecommunications industry in what deregulation seems to have enabled firms to adopt technologies that previously were out of reach for regulatory reasons.31 Moreover, it was amongst the four most active industries in merger activity during the 90s.32

8.1 The Telecommunications Act of 1996 33

In February 1996, the Telecommunications Act became law (Public Law 104-104; 1995 Bill Tracking S. 652). This was the most important reform in the telecommunications sector since the Telecommunications Act of 1934.

The main motivation of the legislators was to allow for changes that had became necessary because of technological progress to ensure that the consumers would fully experience the benefits of it. As the synopsis of the bill says, it is “an original bill to provide for a pro-competitive, de-regulatory national policy framework designed to accelerate rapidly

---

29 In making the connection between the model and the data, this appears as the main advantage.
30 Economides (1999) argues that
31 The deregulation in the banking industry was more of a geographical nature.
32 Andrade et al. (2001)
33 The main source for this section is Economides (1999).
private sector deployment of advanced telecommunications and information technologies and services to all Americans by opening all telecommunication markets to competition”. This seems to corroborate Economides (1999) claim that rapid technological change is the original cause of regulatory change and it gives more content to the empirical assessment that follows.

The main focus of the legislation was to reduce regulatory barriers to entry and competition, to require from the operators the interconnection of networks, the unbundling and the cost-based pricing of leased parts of the network. At the same time, it imposed requirements on the regional Bell operating companies to enter the long distance market.

This change in legislation led to an increase in mergers in the telecommunications industry. Figure 10 shows a merger wave started in 1996, with the value of transactions as a percentage of the market value reaching a peak of 9% by 1999. This seems to

Figure 10: Merger activity in the telecom industry

This change in legislation led to an increase in mergers in the telecommunications industry. Figure 10 shows a merger wave started in 1996, with the value of transactions as a percentage of the market value reaching a peak of 9% by 1999. This seems to

34I compute the market value of all firms in Compustat with SIC codes 4812, 4813, 4822 and 4899, for

34
confirm Andrade et al. (2001) claim that the mergers of the 90s in the telecommunications industry happened as a reaction to changes in legislation.

Economides (1999) argues that some of these mergers were between regional Bell operating companies trying to “maximize their foothold, looking forward to the time when they will be allowed to provide long distance service”. But he also gives an example that seems to suit better the assumptions I made in the model: AT&T acquired TCI to be able to convert the cable access in a multiservice link to residential customers.

Figure 11 shows that after 1996 the telecommunications industry also rose in importance in what concerns the total merger activity when comparing with others industries: though it had been important since the beginning of the decade, it achieved the first position among all industries in 1996 and 1999.

Figure 11: Ranking of telecom sector in terms of merger activity

the relevant years. For the years 1991 to 1995, I divide the total value of transactions listed in Mergerstat by the market value (so the percentage value is an upper bound as not all the sellers listed in Mergerstat show in Compustat). For the years 1996-2001 I consider only the transactions of that Mergerstat classifies as Public and Divestitures (I do not include Private and Foreign). This information was not available for dates prior to 1996. See the appendix A for more details.
8.2 Data

I focus my attention on mergers in which both the target and the acquirer are publicly traded firms. The procedure I adopt is as follows. I select from Compustat all the firms with the SIC codes the Mergerstat aggregates under the denomination “Communications” (4812, 4813, 4822, 4899) that were deleted from Compustat after 1995. Of these, I only consider those that were deleted due to “acquisition or merger” and for which the computation of the market value\(^{35}\) is feasible for the year prior to deletion (52 observations). Using Lexis-Nexis and Mergerstat I match the targets with the acquirers. As before, I only consider those acquirers that appear on Compustat and for which the computation of the market value the year before the deletion of the target is possible (48 observations). Then, I delete all those transactions that do not show up in Mergerstat (44 observations). This database represents 63% to 75% of the value of the mergers in the telecommunications sector that took place between 1996 and 2000, though only 7% to 43% of the transactions.\(^{36}\) Table 1 presents some brief description of this database.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of transactions</th>
<th>Value of transactions (in Million US $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>5</td>
<td>36,614.6</td>
</tr>
<tr>
<td>1997</td>
<td>8</td>
<td>38,488.4</td>
</tr>
<tr>
<td>1998</td>
<td>13</td>
<td>144,393.5</td>
</tr>
<tr>
<td>1999</td>
<td>10</td>
<td>111,703.1</td>
</tr>
<tr>
<td>2000</td>
<td>8</td>
<td>51,634.9</td>
</tr>
</tbody>
</table>

Notes: The sample of mergers is obtained from Compustat and Mergerstat. All mergers must have Compustat information the year prior to de-listing. The year refers to the year of announcement of the merger as described in Mergerstat. The values of the transactions are reported in Mergerstat.

Mergerstat considers the mergers in which at least one of the parts is a US entity. There may be other reasons (also of regulatory nature) that make foreign firms acquire US firms or US firms takeover foreign firms. Therefore, from the baseline database I build a database in which the parts are only US entities, i.e., I exclude mergers with at least one

\(^{35}\)Market value is obtained using the book assets, subtracting the common equity at par and adding it at market value: Market Value = book assets + market equity - book equity. This is the same procedure followed, among others, by Andrade et al. (2001). In the appendix I give more details on this.

\(^{36}\)In the Appendix D I explain this in more detail.
of the parts being an ADR or a Canadian firm (40 observations). Though the model seems more suited to mergers of entities from the same political region, I do all the empirical assessments with both databases.\footnote{In both databases there are six inter-industry mergers. The empirical results do not change if I exclude these observations from the sample.}

## 8.3 Evidence

One of the strongest results of the model is that of proposition 2: the acquirers have low \( m \) and the acquirers have high \( m \). Therefore, one would expect that in the transition from a world in which mergers are not allowed (or are difficult to do) to a world in which mergers are allowed, the value of the firms that eventually will take part in the merger rises, anticipating the gains of merging. If one interprets \( m \) as the productivity of the firm, one should expect the value (or \( q \)) of the firms with low and high productivity goes up, and the value (or \( q \)) of the firms that will not merge, those with median productivity, remains unaffected when deregulation is announced.

The bill that would become the Telecommunications Act of 1996 originated from the Senate Commerce, Science and Transportation Committee and the first session took place in March 30, 1995. However, by November 1994, several newspapers (e.g., USA Today, The Washington Post) anticipated that in year of 1995 the Senate would be working in a telecommunications bill that would impose fewer regulations. It is reasonable to assume that by late 1994 and early 1995 the economic agents had already incorporated the information that a telecommunications act would soon come out of the Congress. This anticipation may also be the reason of why mergers in this industry substantially reduced in 1994 and 1995 (see figures 10 and 11).

Total factor productivity is not available from Compustat, so I look at output per employee as a measure of productivity. My model abstracts from entry and exit, hence I look at firms in the telecommunications sector (as defined above) in the end of 1993 that are still in business in 1995. I consider only those for which the computation of the market value is possible and the number of employees is available (75 observations). I order these firms by increasing order of productivity by the end of 1993 (output per employee) and I
divide the group in tiers. I compute the $qs$ of these firms at the beginning of 1994, 1995 and 1996. I take the growth rates and I average within tiers. Tables 2 and 3 present the results.

Table 2 – Average of growth rates of $q$ from beginning of 1994 to beginning of 1995

<table>
<thead>
<tr>
<th>Productivity Tier</th>
<th>Average*</th>
<th>Test of differences in means with respect to middle tier**</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>-24.4%</td>
<td>1.3%</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.47)</td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>-25.7%</td>
<td>-</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>-10.4%</td>
<td>15.3%</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * The p-value of a t-test of mean equal to zero is in parentheses.
** I test the hypothesis that $x-y$ has a mean of zero against the alternative hypothesis that mean$(x-y)>0$. The p-value of a t-test is in parentheses. It is robust to the assumption of different variances.

Table 3 – Average of growth rates of $q$ from beginning of 1994 to beginning of 1996

<table>
<thead>
<tr>
<th>Productivity Tier</th>
<th>Average*</th>
<th>Test of differences in means with respect to middle tier**</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>-6.1%</td>
<td>11.1%</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>(0.84)</td>
<td>(0.36)</td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>-17.2%</td>
<td>-</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>-0.4%</td>
<td>16.8%</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>(0.96)</td>
<td>(0.08)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * The p-value of a t-test of mean equal to zero is in parentheses.
** I test the hypothesis that $x-y$ has a mean of zero against the alternative hypothesis that mean$(x-y)>0$. The p-value of a t-test is in parentheses. It is robust to the assumption of different variances.

A literal interpretation of the model results would be that the averages of the low and high productivity tiers are positive and that of the middle productivity tier is zero. However, one needs to be aware that there may be some other factors driving the behavior of the stock prices (e.g., aggregate shocks). Therefore, the “correct” prediction of the model is that low and high productivity tiers have averages larger than that of the middle tier. Tables 2 and 3 show the prediction of the model is true for the high productivity tier but it is not valid for the lower productivity tier. This is intriguing as some literature (e.g. Andrade et al. 2001) documents the fact that targets are usually the ones that extract more value from the merger.
In table 4 I look at the change in $qs$ from 1995 to 1996. As I did before, I look at firms in the telecommunications sector (as defined above) in the end of 1994 that are still in business in 1995. Of those, I consider only those for which the computation of the market value is possible and the number of employees is available (97 observations). I order these firms by increasing order of productivity by the end of 1994 (output per employee) and I divide the group in tiers. I compute the $qs$ of these firms at the beginning of 1995 and 1996. I take the growth rates and I average within tiers.

Though the sign of the averages for low and high productivity tiers is as predicted by the model (but not significant in the case of the low productivity tier), the relevant test is the one that looks at the differences between these averages, as some other factors may be driving the behavior of stocks (e.g., aggregate shocks). As predicted by the model, I cannot reject that the average of growth rates of $q$ for the low and high tiers is larger than that for the middle productivity tier.38

<table>
<thead>
<tr>
<th>Productivity Tier</th>
<th>Average*</th>
<th>Test of differences in means with respect to middle tier**</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>20.1%</td>
<td>26.7%</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>-6.6%</td>
<td>-</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>10.4%</td>
<td>17%</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.03)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * The p-value of a t-test of mean equal to zero is in parentheses. ** I test the hypothesis that $x-y$ has a mean of zero against the alternative hypothesis that $\text{mean}(x-y)>0$. The p-value of a t-test is in parentheses. It is robust to the assumption of different variances.

Most of the literature on mergers looks at the Tobin’s $q$ as a measure of managerial performance. Lang et al (1989) and Servaes (1991) found that the $q$ of acquirer is larger than the $q$ of the acquirer. Andrade et al. (2001) calculate that this is true in approximately 66% of the mergers that occurred between 1973 and 1998. In table 5 I summarize the statistics of the samples under study. Coincidently, also in this sample the $q$ of the acquirer is larger than that of the target in 66% of the cases.

38If I look at the same 75 firms of the previous sample between 1995 and 1996, I cannot reject that the average growth rates of $q$ are equal for all tiers. This suggests that this result is driven by changes in the $q$ of the firms that entered in 1995.
The model is silent about the fraction of times the \( q \) of target is smaller than that of the acquirer. It identifies both the distribution function of managerial abilities and the learning function as the primitives that determine this fraction and it says that the \( q \) of the acquirer tends to be larger than that of the target. Nevertheless, it predicts that the relative \( q \) must rise during the merger wave, that is, if I regress the relative \( q \) on time I must observe a positive and significant coefficient

\[
\frac{q^S}{q^B} = \alpha + \beta t, \tag{45}
\]

that is, \( \beta \) is positive. Whether I use as a reference the year of announcement of the merger or the year of de-listing of the target firm, the regression confirms the theory. As table 6 shows, the evidence is stronger for the baseline database than for the US database.

The model also has predictions about the ratio \( \frac{m^S(t)A(t)}{m^B(t)} \). If one thinks of this ratio as a
ratio of relative productivities of target to acquirer, the model predicts that the coefficient of a regression of this ratio on time is positive and significant.

\[
\frac{\text{productivity of target}}{\text{productivity of acquirer}} = \gamma + \lambda t, \quad (46)
\]

that is, \( \lambda \) is positive.

As I described above, I use output per employee as a measure of productivity. Not all firms in Compustat have available data on the number of employees, so the size of the samples reduces to 41 and 37, for the baseline and the US database, respectively.

Surprisingly, the results are strong. In tables 7 and 8, I present the results of a regression of relative productivity of target to acquirer on time (it can be either year of announcement of merger or year of de-listing of the target); in table 7 I consider the full samples, whereas in table 8 I run the same regressions but without an outlier. From the results on the tables it is clear that this outlier observation plays an important role, both in the magnitude of the coefficients as in their statistical significance and in the fit of the regressions. Nevertheless, the result is robust for the sample that consists of US entities only.

<table>
<thead>
<tr>
<th>Table 7: Dynamics of relative productivities (target to acquirer)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mergerstat</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Year</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
</tr>
<tr>
<td>( N )</td>
</tr>
</tbody>
</table>

*Notes: The table represents estimates for equation (40), either with the year of announcement or the year of de-listing as independent variable. The \( p \)-values are in parentheses.*
Table 8: Dynamics of relative productivities (target to acquirer) without outlier

<table>
<thead>
<tr>
<th></th>
<th>Mergerstat</th>
<th>Mergerstat US only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year of</td>
<td>Year of de-</td>
</tr>
<tr>
<td></td>
<td>announcement</td>
<td>listing</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>Year</td>
<td>0.108</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>4.5%</td>
<td>3.7%</td>
</tr>
<tr>
<td>N</td>
<td>40</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table represents estimates for equation (41), either with the year of announcement or the year of de-listing as independent variable. I removed an outlier observation from the benchmark databases (relative productivity of around 8.5). The $p$-values are in parentheses.

One could also be willing to interpret the managerial ability as a biunivocal measure of size as Lucas (1978) did.\textsuperscript{39} Then the conclusions would be that small and large firms see their value (or $q$) rise upon deregulation (whereas middle size firms are unaffected) and the relative size of target to acquirer increases during the merger wave. However, these data do not support this interpretation.

9 Conclusion

In this paper, I explain mergers and merger waves using an equilibrium model of a market for organization capital. A technological and/or deregulatory shock delivers a higher level of profits in the long run but leads to reductions in profits in the short run. The short run losses and the long run gains differ across firms due to heterogeneity in managerial ability. If the knowledge accumulated about the new technology (organization capital) is transferable through the ownership of the firms, mergers happen.

I show that a model of this kind is able to match some of the stylized facts about mergers, namely that mergers occur in waves, the value created by the merger is larger the

\textsuperscript{39}See footnote 5. In such a set-up, the size as measured by the number of employees is $l = m \left( \frac{1-\alpha}{\omega} \right)^{\frac{1}{\alpha}}$ or by the sales (or output) is $m \left( \frac{1-\alpha}{\omega} \right)^{\frac{1-\alpha}{\alpha}}$ is an increasing (and linear) function of the managerial ability.
larger the differences in managerial abilities and the \( q \) of acquirer tends to be larger than that of the target. At the same time, this model has some unique empirical predictions: it predicts that the ratio of the \( q \) of target to the \( q \) of acquirer and the ratio of productivities of target and acquirer rise during the merger wave.

I compare these empirical implications of the model with one of the most important merger waves of the 90s, that of the telecommunications sector. The data seems to confirm the theoretical implications of the model, namely, the dynamic behavior of the relative \( q \) and the relative productivity.

This paper leaves some open questions for future research. On the theoretical side, it would be interesting to see what happens if I relax the assumption that the pace of learning is equal for every firm. This could be done in two different ways: one way would be to make learning a function of the managerial ability (increasing? decreasing?); another way would be to introduce a different characteristic, a learning ability. Perhaps this last approach is more realistic but I believe the solution to the competitive equilibrium, namely the price function, would be very hard to obtain.

On the empirical side, it would be interesting to try to understand to what types of industries this model fits better. I showed one industry for which this type of capital may be important (or, at least, I cannot reject it is not). Organization capital may not be important or transferable in all industries. Probably it is more important in horizontal mergers than in conglomerate formations. In the formulation I proposed, the organization is not visible, as it is in the work of Prescott and Visscher (1980). A better understanding of what is organization capital probably would help to answer this question.
Appendix A - Organization capital is non-tradable

There may be cases in which the intangible capital of the firm coincides with the human capital of its workers, that is, it is imbedded in the individuals that make part of the firm, but not in the firm itself. In these cases, the potential acquirer is always at risk of observing the employees of the target firm running away after the takeover takes place.40

In this section I consider the general case in which firms are not able to transfer their organization capital either because it “vanishes” after the transaction or because the frictions associated with the transaction totally destroy it.41 In other words, the information each firm gathers about the new technology is not transferable. In an environment like this there is no market for organization capital.

The problem each firm faces is

\[
V(m) = \max_{t_a} \int_0^{t_a} e^{-rt} md t + \int_0^\infty e^{-rt} A(t) md t,
\]

where \( t_a \), the decision variable, is the time of adoption of the new technology. The FOC of this problem is

\[
e^{-rt_a} m \left[ 1 - r \int_0^\infty e^{-rt} A(t) dt \right],
\]

and its sign depends on the sign of \( 1 - r \int_0^\infty e^{-rt} A(t) dt \). That is, if

\[
\int_0^{t_a} e^{-rt} md t > e^{-rt_a} \int_0^\infty e^{-rt} A(t) md t \Leftrightarrow 1 - r \int_0^\infty e^{-rt} A(t) dt > 0
\]

\( FOC > 0 \), so \( t_a = \infty \), i.e., it is better to never adopt the new technology; if

\[
\int_0^{t_a} e^{-rt} md t < e^{-rt_a} \int_0^\infty e^{-rt} A(t) md t \Leftrightarrow 1 - r \int_0^\infty e^{-rt} A(t) dt < 0
\]

40 Brealey and Myers (2000) give an example of a Portuguese bank that bought an investment management company and saw the whole team of the target firm move out and build a rival company with a very similar name to that of the target.

41 In the article “After the deal” published in January 7, 2001 issue of The Economist it is argued that there is a “soft trap” in the mergers that can destroy value. This section is an extreme interpretation of the view presented in that article.
\( FOC < 0 \) and so \( t_a = 0 \), it is better to adopt the new technology immediately. The case in which \( FOC = 0 \) is uninteresting because of the indeterminacy (and indifference) result, so I overlook it.

The problem of the firm, *adopt it now or never adopt it*, is better described by the problem

\[
\max \left\{ \int_0^\infty e^{-rt} A(t) \, dt, \int_0^\infty e^{-rt} \, dt \right\},
\]

or equivalently

\[
m \max \left\{ \int_0^\infty e^{-rt} A(t) \, dt, \frac{1}{r} \right\}
\]

or

\[
\max \left\{ \int_0^\infty e^{-rt} A(t) \, dt, \frac{1}{r} \right\}.
\]

The managerial ability plays no role whatsoever. As I assume the new technology is available to all firms, if \( \int_0^\infty e^{-rt} A(t) \, dt > \frac{1}{r} \), all firms adopt it; if \( \int_0^\infty e^{-rt} A(t) \, dt < \frac{1}{r} \), all firms stick to the old technology. Clearly, this is a consequence of the particular form of the profit function.

### Appendix B - Proofs

#### Proof of Lemma 1

The value of being a seller is

\[
V^S(m) = \int_0^{\phi(m)} e^{-rt} A(t) \, dt + e^{-r\phi(m)} q(\phi(m)) ,
\]

where

\[
\psi(m) = \arg \max_{t^s} \int_0^{t^s} e^{-rt} A(t) \, dt + e^{-rt^s} q(t^s).
\]
The value of being a buyer is

\[ V^B (m) = \int_0^{\psi(m)} e^{-rt} m \, dt + \int_0^{\infty} e^{-rt} A(t) \, m \, dt - e^{-r\psi(m)} q(\psi(m)), \]

where

\[ \psi(m) = \arg \max_{t^d} \int_0^{t^d} e^{-rt} m \, dt + \int_0^{\infty} e^{-rt} A(t) \, m \, dt - e^{-rt^d} q(t^d). \]

The decision a firm \( m \) makes is

\[ V(m) = \max \{ V^S (m), V^B (m) \}. \]

Define \( \bar{m} \) as the firm that is just indifferent between being a buyer and a seller, i.e.,

\[ V(\bar{m}) = V^S (\bar{m}) = V^B (\bar{m}). \]

This implies that

\[ \int_0^{\phi(\bar{m})} e^{-rt} A(t) \, \bar{m} \, dt + e^{-r\phi(\bar{m})} q(\phi(\bar{m})) = \int_0^{\psi(\bar{m})} e^{-rt} \bar{m} \, dt + \int_0^{\infty} e^{-rt} A(t) \, \bar{m} \, dt - e^{-r\psi(\bar{m})} q(\psi(\bar{m})), \]

or

\[ e^{-r\phi(\bar{m})} q(\phi(\bar{m})) + e^{-r\psi(\bar{m})} q(\psi(\bar{m})) = \left[ \int_0^{\psi(\bar{m})} e^{-rt} \, dt + \int_0^{\infty} e^{-rt} A(t) \, dt - \int_0^{\phi(\bar{m})} e^{-rt} A(t) \, dt \right] \bar{m}. \]

By the envelope theorem\(^{42}\)

\[ \frac{\partial V^S (\bar{m})}{\partial m} = \int_0^{\phi(\bar{m})} e^{-rt} A(t) \, dt. \]

\[ \frac{\partial V^B (\bar{m})}{\partial m} = \int_0^{\psi(\bar{m})} e^{-rt} \, dt + \int_0^{\infty} e^{-rt} A(t) \, dt, \]

\(^{42}\)Here I abuse notation: \( \frac{\partial V^S (q)}{\partial m} \) means \( \frac{\partial V^S (q)}{\partial m} \) evaluated at \( m = \bar{m} \).
If either \( q(\phi(\bar{m})) > 0 \) or \( q(\psi(\bar{m})) > 0 \) (in Proposition 1, I show this is indeed the case), then

\[
\frac{\partial V^B(\bar{m})}{\partial m} > \frac{\partial V^S(\bar{m})}{\partial m}.
\]

The value functions are convex in \( m \):

\[
\frac{\partial^2 V^S(\bar{m})}{\partial m^2} = \begin{cases} 
\phi'(m) e^{-r\phi(m)} A(\phi(m)) > 0 & \text{if } \phi(m) < +\infty \\
0 & \text{if } \phi(m) = +\infty
\end{cases},
\]

\[
\frac{\partial^2 V^B(\bar{m})}{\partial m^2} = \begin{cases} 
\psi'(m) e^{-r\psi(m)} [1 - A(\psi(m))] > 0 & \text{if } \psi(m) < +\infty \\
0 & \text{if } \psi(m) = +\infty
\end{cases}.
\]

Therefore it must be that \( \bar{m} \) is unique and so for \( m < \bar{m}, \ V^B(m) < V^S(m) \), and for \( m \geq \bar{m}, \ V^B(m) \geq V^S(m) \), that is,

\[ S = [m_l, \bar{m}] \text{ and } B = [\bar{m}, m_h]. \]

Hence, there exists a unique cut-off level in the managerial ability that divides firms in sellers and buyers. \( \blacksquare \)

**Proof of Lemma 2**

The lowest managerial ability buyer is \( \bar{m} \). If \( \lim_{t \to \infty} q(t) \geq 0 \), then \( \psi(\bar{m}) < +\infty \), because otherwise \( \bar{m} \) would always be better off by adopting the technology at date zero and never sell it\(^{43} \), that is, \( V^B(\bar{m}) < V^S(\bar{m}) \). But this contradicts the definition of \( \bar{m} \).

As \( \psi(\bar{m}) < +\infty \) then it must be that \( \exists \ _m \in S \), such that \( \phi(\_m) = \psi(\bar{m}) = \bar{t} \) and so

\[
\frac{\bar{m}}{_m} = \frac{A(\bar{t})}{A(\bar{t}) - 1}.
\]

This means that \( _m < \bar{m} \). As \( \psi'(m) < 0 \), \( \bar{t} = \psi(\bar{m}) \) is the latest (finite) date at which transactions occur. This means that for markets to clear it must be that \( m \in (_m, \bar{m}) \).

\(^{43}\)This follows from the assumption \( r \int_0^\infty e^{-rt} A(t) \, dt > 1 \).
choose never to sell the technology they adopted at time zero, i.e., \( \phi(m) = +\infty \) for \( m \in \left( m_-, \bar{m} \right) \).

**Proof of Lemma 3**

Define \( \hat{G}(\cdot) \) as

\[
\hat{G}(x) = G(T(x,1)) = \int_{0}^{T(x,1)} e^{-rt} (1 - A(\tau)) \, xd\tau - \int_{T(x,1)}^{\infty} e^{-rt} A(\tau) \, sd\tau,
\]

whose first and second derivative are, respectively,

\[
\hat{G}'(x) = \int_{0}^{T(x,1)} e^{-rt} (1 - A(\tau)) \, d\tau
\]

and

\[
\hat{G}''(x) = \frac{dT(x,1)}{dx} e^{-rT(x,1)} (1 - A(T(x,1))).
\]

As \( \frac{dT(x,1)}{dx} < 0 \) and \( A(T(x,1)) > 1 \), for all \( x \), the function \( \hat{G}(\cdot) \) is globally convex.

The minimum of this function is unique and the minimizer \( x_{\text{min}} \) satisfies the equation

\[
\int_{0}^{T(x_{\text{min}},1)} e^{-rt} (1 - A(\tau)) \, d\tau = 0.
\]

The value of \( x \) that attains the zero of the function, \( x_{0} \), verifies the equality

\[
\int_{0}^{T(x_{0},1)} e^{-rt} (1 - A(\tau)) \, x_{0}d\tau = \int_{T(x_{0},1)}^{\infty} e^{-rt} A(\tau) \, sd\tau
\]

and this implies that

\[
\int_{0}^{T(x_{0},1)} e^{-rt} (1 - A(\tau)) \, d\tau > \int_{0}^{T(x_{\text{min}},1)} e^{-rt} (1 - A(\tau)) \, d\tau = 0
\]

and so \( x_{0} > x_{\text{min}} \) and \( \hat{G}''(x_{0}) > 0 \).

This means that the value of \( x \) that attains the zero must be to the right of the minimizer. As the function \( \hat{G}(\cdot) \) is globally convex, this value is therefore *unique*. ■
Proof of Proposition 5

The \( q \) of the target by the date of acquisition is \( q^S(t) = V^S(t) = e^{rt}p(t) = \int_t^\infty e^{-r(\tau-t)}A(\tau)s(\tau)d\tau \). The \( q \) of the acquirer by the date of acquisition is \( q^B(t) = e^{rt}\int_t^\infty e^{-r(\tau-t)}A(\tau)(b(\tau) - s(\tau))d\tau = \int_t^\infty e^{-r(\tau-t)}A(\tau)b(\tau)d\tau - q^S(t) \). From here it follows that \( q^B(t) > q^S(t) \Leftrightarrow b(t) > \frac{2\int_t^\infty e^{-r\tau}A(\tau)s(\tau)d\tau}{\int_t^\infty e^{-r\tau}A(\tau)d\tau} \). As \( b'(t) < 0 \) and

\[
\frac{\partial}{\partial t}\left(\frac{\int_t^\infty e^{-r\tau}A(\tau)s(\tau)d\tau}{\int_t^\infty e^{-r\tau}A(\tau)d\tau}\right) = \frac{-e^{-rt}A(t)[s(t)\int_t^\infty e^{-r\tau}A(\tau)d\tau] - \int_t^\infty e^{-r\tau}A(\tau)s(\tau)d\tau}{\left[\int_t^\infty e^{-r\tau}A(\tau)d\tau\right]^2} > 0,
\]

because \( s'(t) \geq 0 \). \( \blacksquare \)

Appendix C - About the Mergerstat

This appendix describes the composition and the methodology of Mergerstat.\( ^{44} \) This is important to understand figures 1 and 2 and tables 2 and 3.

“Mergerstat compiles statistics on publicly announced mergers, acquisitions and divestitures involving operating units. Mergerstat tracks formal transfers of ownership of at least 10% of a company’s equity where the purchase price is at least \( \$1,000,000 \) and where at least one the parts is a US entity”.

“Divestitures: sales of corporate units, unit management buyouts and minority equity interest purchases, both foreign and domestic, public and private”.\( ^{45} \)

“Publicly traded sellers: acquisitions of publicly traded companies, tender offers and public statutory mergers, but excludes any minority equity interest purchases and acquisitions of foreign companies”.

“Privately owned sellers: acquisitions of privately held companies except those considered divestitures and those of foreign companies”.

“Foreign sellers: acquisitions of foreign companies except foreign divestitures, unit management buyouts and minority equity interest purchases”.

The universe of the database does not coincide totally with any of these. For example, not all the foreign sellers are represented in Compustat and the management buyouts

\( ^{44} \) In what follows, all the quotations are from the Mergerstat review (any year between 1997 and 2001).

\( ^{45} \) Minority equity interest purchase: the acquisition of 10 to 50% of a company is a minority equity purchase.
are not registered in Compustat as an acquisition or merger. Therefore, the numbers represented in the graphs and tables are only an approximation to the true value. Specially in what concerns Table 3, the values seem to be pretty accurate as changes in the universe under consideration do not change the values by much.

Appendix D - Constructing the database and the variables

I used the following footnotes and data items of Compustat to build the database and the relevant variables:

Table App. D. 1 – Compustat Codes

<table>
<thead>
<tr>
<th>Compustat footnote</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFTNT 33</td>
<td>Month of deletion</td>
</tr>
<tr>
<td>AFTNT 34</td>
<td>Year of deletion</td>
</tr>
<tr>
<td>AFTNT 35</td>
<td>Reason of deletion</td>
</tr>
<tr>
<td></td>
<td>(code 1: acquisition or merger)</td>
</tr>
</tbody>
</table>

Table App. D. 2 – Construction of variables

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Compustat data item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Assets</td>
<td>6</td>
</tr>
<tr>
<td>Sales</td>
<td>12</td>
</tr>
<tr>
<td>Number of employees</td>
<td>29</td>
</tr>
<tr>
<td>Productivity</td>
<td>12/29</td>
</tr>
<tr>
<td>Book equity</td>
<td>60</td>
</tr>
<tr>
<td>Market equity</td>
<td>199*25</td>
</tr>
<tr>
<td>Market value</td>
<td>6+199*25-60</td>
</tr>
<tr>
<td>Tobin's q</td>
<td>(6+199*25-60)/6</td>
</tr>
</tbody>
</table>

I did the matching between target and acquirer using Lexis-Nexis and Mergerstat review. Then I looked up in Compustat for the data concerning the acquirer. Here it was important to keep track of changes in name and SMBL (ticker symbol). I did this using Lexis-Nexis. The list of mergers is available upon request.

The baseline database represents from 7% to 43% of the transactions and from 63% to 75% of the value of the mergers that took place between 1996 and 2000. The range depends on whether I take as the valid universe all the transactions in which the seller is not privately owned (lower bound) or the case in which the seller is a public firm (upper bound). None of these universes is the true universe of the transactions registered in this
database, but the true relative weight of these lies somewhere between the lower and the upper bound of the ranges described in Table App. D. 3 and in Table App. D. 4.

Table App. D. 3 – Fraction of telecom mergers registered in Mergerstat

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Except Private</th>
<th>Except Private &amp; Foreign</th>
<th>Public</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>4.1%</td>
<td>11.4%</td>
<td>17.24%</td>
<td>50.0%</td>
<td>11-50%</td>
</tr>
<tr>
<td>1997</td>
<td>4.8%</td>
<td>9.6%</td>
<td>14.81%</td>
<td>40.0%</td>
<td>10-40%</td>
</tr>
<tr>
<td>1998</td>
<td>5.3%</td>
<td>12.1%</td>
<td>18.84%</td>
<td>61.9%</td>
<td>12-62%</td>
</tr>
<tr>
<td>1999</td>
<td>2.2%</td>
<td>4.8%</td>
<td>10.42%</td>
<td>31.3%</td>
<td>5-31%</td>
</tr>
<tr>
<td>2000</td>
<td>1.7%</td>
<td>4.7%</td>
<td>8.42%</td>
<td>42.1%</td>
<td>5-42%</td>
</tr>
<tr>
<td>Total</td>
<td>3%</td>
<td>7.2%</td>
<td>12.8%</td>
<td>43.1%</td>
<td>7-43%</td>
</tr>
</tbody>
</table>

Source: Mergerstat. The year refers to year of announcement of the merger.

Table App. D. 4 – Fraction of value of telecom mergers registered in Mergerstat

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Except Private</th>
<th>Except Private &amp; Foreign</th>
<th>Public</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>47.8%</td>
<td>48.6%</td>
<td>49.0%</td>
<td>49.6%</td>
<td>49-50%</td>
</tr>
<tr>
<td>1997</td>
<td>70.1%</td>
<td>70.6%</td>
<td>77.2%</td>
<td>91.0%</td>
<td>71-91%</td>
</tr>
<tr>
<td>1998</td>
<td>84.8%</td>
<td>86.3%</td>
<td>89.9%</td>
<td>97.0%</td>
<td>86-97%</td>
</tr>
<tr>
<td>1999</td>
<td>36.6%</td>
<td>37.9%</td>
<td>40.2%</td>
<td>46.5%</td>
<td>38-47%</td>
</tr>
<tr>
<td>2000</td>
<td>60.6%</td>
<td>69.2%</td>
<td>75.2%</td>
<td>90.0%</td>
<td>70-90%</td>
</tr>
<tr>
<td>Average</td>
<td>60%</td>
<td>62.5%</td>
<td>66.3%</td>
<td>74.8%</td>
<td>63-75%</td>
</tr>
</tbody>
</table>

Source: Mergerstat. The year refers to year of announcement of the merger.
References


