Technology Uncertainty, Sunk Costs, and Industry Shakeout

Luís Cabral*
New York University

June 2009

Abstract

I propose a novel model of industry shakeout: Because of capacity sunk costs and the fear of backing the wrong technology, each firm initially invests up to a very small capacity, leading to a large number of initial competitors. As the dust settles and a dominant technology emerges, each surviving firm expands to its long-term optimal capacity, which results in a reduction in the number of firms (notwithstanding the increase in total market output).

*Stern School of Business, New York University; lcabral@stern.nyu.edu. I thank Hugo Hopenhayn and Boyan Jovanovic for comments and suggestions. The usual disclaimer applies.

1 Introduction

A common feature of the evolution of new industries is the occurrence of a shakeout: a significant decrease in the number of active firms that takes place during a phase of market expansion. Industry shakeouts have been documented for a variety of industries, from automobiles to computers (Gort and Klepper, 1982; Klepper and Graddy, 1990; Filson, 2001). While the theoretical explanation for industry shakeouts varies, most models have one element in common: at some point in the industry life cycle, the best technology's marginal cost schedule decreases, that is, technology evolution calls for larger firm sizes. If the increase in firm size is greater than the rate of market expansion, then a shakeout must take place (Gort and Klepper, 1982; Hopenhayn, 1993; Jovanovic and MacDonald, 1994; Klepper, 1996; Klepper and Miller, 1996).[1]

I propose a complementary explanation for the drastic reduction in the number of active firms. My theory relies on the interaction of two well documented stylized facts about industry evolution: technology uncertainty and investment sunk costs. Before introducing my theory in greater detail, I describe these stylized facts in greater detail.

The first stylized fact is that the early evolution of an industry is a time of great uncertainty, which may be gradually or rapidly resolved. Initially, firms invest in creating different product variants; and as consumers experiment with the various offerings, a winner eventually emerges — a dominant design — at which point firms start focusing on how better to produce it (Utterback and Albernathy, 1975; Abernathy and Utterback, 1978; see also Mueller and Tilton, 1969, Tushman and Anderson, 1986). However, the resolution of uncertainty need not be limited to consumer preferences: Often, multi-

1. A different stream of literature interprets shakeouts as overshooting in the entry process when there is limited coordination among potential entrants. Building on the work of Dixit and Shapiro (1986) and Cabral (1993), Klepper and Miller (1995) test the explanatory power of this approach, finding the empirical support to be mixed. On this approach, see also Vettas (2000).

ple competing technologies battle for dominance until a de facto standard emerges (Dosi, 1982; Anderson and Tushman, 1990; Schilling, 1998).

The second stylized fact is that many of the firms' product and process investments are to some extent sunk (Dixit and Pinkyck, 1994; Lambson, 1991; Cabral, 1995); that is, these investments become worthless if the firm decides to leave the industry (if, for example, it happens to back the wrong horse).

In this paper, I put these two ideas together and show that they lead to a very natural theory of industry shakeout. The idea is that firms initially invest up to small capacity levels; and once uncertainty has been resolved, most or all of the firms that made the wrong investment leave the industry, whereas the remaining ones expand their capacity to its optimal long-run level. Although market output is smaller in the initial phase of industry expansion, each firm's output is much smaller, which in turn implies the number of firms is greater. This is the essence of an industry shakeout: whereas market output increases, the number of active firms decreases.

Lest this may seem a trivial point, I note (and below show formally) that this reasoning depends critically on investment cost sunkness: If investments were not sunk, then the resolution of technology uncertainty would simply lead firms to switch production technology from the "orphaned" ones to the winning technology design.

I see the effect of sunk costs as complementary to the effect of technology change described in the previous literature. An industry shakeout requires that either cost functions change (leading to lower marginal cost) or capacity costs are sunk. Or both. In fact, the reality of most industries probably reflects both theories. The relative contribution of each theory is then a matter of empirical research. One possible test for the contribution of the my theory is that the greater the degree of investment sunkness in a given industry, the greater the likelihood (and the greater the extent) of an industry shakeout.

The remainder of the paper is structured as follows. In Section 2, I lay down my basic model of industry evolution. In Section 3, I present the main result, stating when and why an industry shakeout takes place, in particular the role of capacity sunk costs in producing such shakeout. Section 4 concludes the paper.

2 A model of industry evolution

Consider a competitive industry in discrete time ($t = 1, 2, \ldots$). In each period, there is a demand $D(p)$ for the industry's homogeneous product. By assuming product homogeneity, I am ignoring all technology uncertainty related to product design, focusing instead on technology uncertainty. However, as I will argue later, the thrust of my argument is valid in both cases.

I assume market demand has the following properties: if $p \le p^*$, then demand is given by Q^*; if $p^* < p \le p^{**}$, then demand is given by $Q^{**} < Q^*$; and if $p > p^{**}$ then demand is zero. This stylized representation of demand encapsulates the idea that there are two distinct market segments: the enthusiasts, who are willing to pay up to a high amount (p^{**}), and the mass market, where consumers are only willing to pay up to p^*.

Next I assume there are two production technologies, 1 and 2. I will also refer to these as the small-scale and the mass-market production technologies, respectively. For a setup cost of ϕ_i, the firm is able to produce at zero additional cost up to a capacity κ_i, where $i = 1, 2$ and $\phi_1 < \phi_2$ and $\kappa_1 < \kappa_2$.

We now come to the crucial assumptions regarding technology evolution and sunk costs. At $t = 2$, one of the two initial technologies is rendered useless. Firms that invested in the "good" technology are able to continue using it. Firms that previously opted for the "bad" technology must either exit or switch technologies. The implications of technology uncertainty depend crucially on whether investment costs are sunk or not. With sunk costs, the set up cost ϕ_i cannot be recovered: whatever amount the firm spends in period one in a "bad" technology is a foregone investment. If investment costs are not sunk, however, then the firm effectively pays $(1-\delta)\,\phi_i$ in each period in order to use technology i. The thrust of my argument is that, if technology investment is not a sunk cost, then firms opt for a mass-market technology from the get-go; whereas under sunk costs firms prefer a niche technology until technology uncertainty is resolved.

I make a number of assumptions regarding parameter values.

Assumption 1 $\quad p^*\,\kappa_1 < (1-\delta)\,(\epsilon+\phi_1)$

This assumption essentially implies that the niche-production technology is not profitable at the mass market price. This assumption is necessary to ensure that, after $t=2$ (sunk costs) and after $t=1$ (no sunk costs), it is optimal to choose a mass-market technology.

Assumption 2 $\quad p^*\,\kappa_2 < \epsilon+\phi_2$

This assumption implies that, at the mass-market price level and assuming production for one period only, the mass-market production technology is not profitable. This assumption is necessary to ensure that, under sunk costs, it is optimal to start-up with the niche production technology.

Assumption 3 $\quad \frac{(1-\delta)\,(\epsilon+\phi_2)}{\kappa_2} < p^* < \frac{\epsilon+\phi_1}{\kappa_1} - \frac{\delta}{2}\,\frac{\phi_2}{\kappa_2}$

This assumption implies that the mass-market price is neither too low nor too high, in which case we obtain a corner solution (whereby either the niche or the mass-market technology is always chosen).

Assumption 4 $\quad \frac{\kappa_2-\kappa_1}{\kappa_1} > 2\,\frac{Q^*}{Q^{**}}$

This assumption states that the proportional difference between the niche and the mass-market production technologies is very significant with respect to the difference between the niche and the mass-market segments of market demand.

These assumptions define an open set of parameter values, and so my results are not knife-edged, that is, do not depend on a very specific (measure zero) set of parameter values. Specifically, consider for example the case when $p^* = 1, Q^* = 4000, Q^{**} = 2000, \delta = .9, \kappa_1 = 10, \kappa_2 = 80, \phi_1 = 120, \phi_2 = 300, \epsilon = 400$. Then the inequalities of Assumptions 1–4 become $10 < 52$, $80 < 700$, $.875 < 1 < 50.3125$, and $7 > 4$.

3 Industry shakeout

As Jovanovic and MacDonald (1995) state, a theory that explains the evolution of industries like the U.S. automobile tire industry (the particular industry they focus on), must address the following time-series features: "a decline in price, an increase in average and total output, and a nonmonotone time path for firm numbers." My main result states that such an equilibrium exists, and moreover, the above features only take place if capacity costs are sunk.

Theorem 1 (sunk costs and industry shakeout) *Under A1–A4, there exists a unique equilibrium. If investment costs are not sunk, then* $n_1 = n_2 = n_t$*; if investment costs are sunk, then* $n_1 > n_2 = n_t$ *even though* $Q_1 < Q_2 = Q_t \quad (t \geq 2)$.

The proof is presented in the Appendix. Figure 1 illustrates the evolution of the number of competitors. If there are no sunk costs, then firms invest to their optimal long-run level from the get-go. The resolution of uncertainty regarding technology only has the effect of inducing some firms to switch from one technology to another; but the optimal output level per firm always remains the same: κ_2. If technology investment costs are sunk, however, then firms prefer to invest initially in the niche technology (firm capacity κ_1). As a result, the total number of firms is greater. When technology uncertainty is resolved, some of the firms who backed the wrong technology exit the industry, leading to a shakeout.

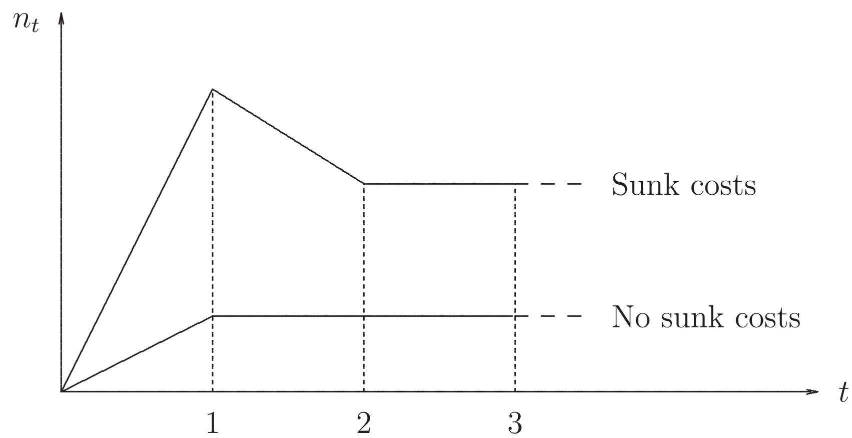

Figure 1: Number of active firms with and without sunk costs.

How is it possible that, under sunk costs, firms who chose the wrong technology are indifferent between exiting and remaining active; whereas, under no sunk costs, all firms that initially entered strictly prefer to remain active? The answer is that market price under sunk costs, which starts off at a very high level, falls to a very low level at $t = 2$, lower than under no sunk costs. Why does price fall down by so much? Because of the large number of entrants in the first period, who have already spend the entry cost and are eager to continue competing.

Notice that, even though market conditions with and without sunk costs are the same after $t = 2$, the number of active firms is permanently different between the two cases. The reason is that, because the initial entry cost is sunk, there is some degree of path dependence: a good fraction of the many initial entrants under the sunk cost case remain active in future periods.

4 Final remarks

I propose what I think is a very natural theory of industry shakeout: Because of capacity sunk costs and the fear of backing the wrong technology, each firm initially invests up to a very small capacity, leading to a large number of initial competitors. As the dust settles down and a dominant technology emerges, each surviving firm expands to its long-term optimal capacity, which in turn results in a shakeout.

As I mentioned earlier, one need not consider different theories as alternative, competing views. In fact, the idea of sunk costs is perfectly compatible with the "implementation race" described by Gort and Klepper (1982) and Jovanovic and MacDonald (1994). Ultimately, the merit of each explanation is an issue to be decided by empirical evidence. Take for example the automobile industry. Mass production started with the the 1901 Curved Dashed Oldsmobile. However, it was not until the Ford Motor Company introduced the Model T in 1908 that it really took shape. The timing of the switch to mass production is roughly consistent with the resolution of one important element of technological uncertainty: engine technology. In fact, by the early 1900s, gasoline cars started to outsell both steam and electric engines.

This is hardly a proof of the idea of this paper, but it seems broadly consistent with it. I conjecture that other industries would produce similar evidence.

Appendix

Proof of Theorem 1: I first show that the following is an equilibrium. Under no sunk costs, firms choose the mass-production technology from the beginning; let price be p_0 under this equilibrium and suppose that $p_0 < p^*$ (which I will confirm later). Under sunk costs, firms choose the small-scale production technology in the first period and the mass-production technology in future periods; let p_1, p_2 be the equilibrium price at $t = 1$ and $t \geq 2$, respectively, and suppose that $p_1 > p^*$ and $p_2 < p^*$ (both of which I will confirm later). After showing that this is an equilibrium, I show that it induces a shakeout under sunk costs. I also show that the equilibrium is unique.

■ **No sunk costs.** Consider first the case of no sunk costs. The zero-profit condition for an entrant in the first period is given by

$$\frac{\kappa_2 \, p_0}{1-\delta} - \phi_2 = \epsilon$$

which implies

$$p_0 = \frac{(1-\delta)\,(\epsilon + \phi_2)}{\kappa_2} \tag{1}$$

The condition that none of the entrants wants to switch to the niche technology is

$$\frac{\kappa_2 \, p_0}{1-\delta} - \phi_2 \geq \frac{\kappa_1 \, p^*}{1-\delta} - \phi_1$$

To understand this inequality, notice that, by reducing output from κ_2 to κ_1, a deviant firm would induce an increase in price from p_0 to p^*. Substituting (1) for p_0 and simplifying we get

$$p^* \leq \frac{(1-\delta)\,(\epsilon + \phi_1)}{\kappa_1} \tag{2}$$

■ **Sunk costs.** Consider now the case with sunk costs. The zero-profit condition for an entrant in the first period is given by

$$\kappa_1 \, p_1 - \phi_1 + \frac{\delta}{2} \, \frac{\kappa_1 \, p_2}{1-\delta} = \epsilon \tag{3}$$

The reason for this is that, after the first period, (a) if the firm chose the wrong technology (probability $\frac{1}{2}$) it will be indifferent between remaining active and exiting; (b) if the firm chose the right technology (probability $\frac{1}{2}$), then it will earn a stream of profits $\frac{\kappa_1 \, p_2}{1-\delta}$.

At $t = 2$, first-period entrants who chose the wrong technology must be indifferent between remaining active (and investing in the mass-production technology) and exiting:

$$\frac{\kappa_2 \, p_2}{1-\delta} - \phi_2 = 0 \tag{4}$$

or simply

$$p_2 = \frac{(1-\delta) \, \phi_2}{\kappa_2} \tag{5}$$

Substituting for p_2 in (3) and simplifying,

$$p_1 = \frac{\epsilon + \phi_1}{\kappa_1} - \frac{\delta}{2} \, \frac{\phi_2}{\kappa_2} \tag{6}$$

A sufficient condition that a first-period entrant does not prefer to switch to a mass-production technology is

$$\kappa_1 \, p_1 - \phi_1 + \frac{\delta}{2} \, \frac{\kappa_1 \, p_2}{1-\delta} \geq \kappa_2 \, p^* - \phi_2 + \frac{\delta}{2} \, \frac{\kappa_2 \, p_2}{1-\delta} \tag{7}$$

To understand this inequality, notice that, by increasing output from κ_1 to κ_2, a deviant firm would induce an decrease in price from p_1 to p^*. The condition is sufficient because, for $t \geq 2$, equilibrium price under deviation would be lower than p_2. Using (3) and (5) to simplify the left-hand side and the right-hand side of (7), respectively, we get

$$p^* \leq \frac{\epsilon + \left(1 - \frac{\delta}{2}\right) \phi_2}{\kappa_2} \tag{8}$$

■ **Equilibrium prices.** To be consistent with our assumptions regarding the demand function and the equilibrium, we require that $p_1 > p^*$, and both $p_0 < p^*$ and $p_2 < p^*$. Equations (1) and (5) imply that $p_0 > p_2$, and so the condition $p_0 < p^*$ implies the condition $p_2 < p^*$. We are thus left with the condition $p_1 > p^*$, which from (6) becomes

$$p^* < \frac{\epsilon + \phi_1}{\kappa_1} - \frac{\delta}{2}\,\frac{\phi_2}{\kappa_2} \tag{9}$$

and the condition $p_0 < p^*$, which from (1) becomes

$$p^* > \frac{(1-\delta)\,(\epsilon + \phi_2)}{\kappa_2} \tag{10}$$

We thus have four equilibrium inequalities: (2), which states that, under no switching costs, firms prefer the mass-production technology; (8), which states that, under switching costs, firms prefer the small-scale-production technology; (9), which ensures that niche-production technology leads to a price higher than p^*; and (10), which ensures that mass-production technology leads to a price lower than p^*. These inequalities correspond to Assumptions 1–4.

■ **Uniqueness.** To show uniqueness, I still need to check that there exists no alternative equilibrium whereby, even without sunk costs, firms would invest in the niche technology and equilibrium price is high. Suppose that is the case. The the zero-profit condition implies that

$$\frac{\kappa_1\,p_0'}{1-\delta} - \kappa_1 = \epsilon$$

which implies

$$p_0' = \frac{(1-\delta)\,(\epsilon + \phi_1)}{\kappa_1} \tag{11}$$

The condition that none of the entrants wants to switch to the mass production technology is

$$\frac{\kappa_2\,p^*}{1-\delta} - \phi_2 \le \frac{\kappa_1\,p_0'}{1-\delta} - \phi_1$$

Substituting (11) for p_0' and simplifying, we get

$$p^* \leq \frac{(1-\delta)\,(\epsilon+\phi_2)}{\kappa_2}$$

which contradicts Assumption 3.

■ **Shakeout.** Finally, I need to show that the above equilibrium does correspond to a shakeout, that is, under sunk costs the number of firms decreases from $t = 1$ to $t = 2$. First notice that

$$n_1 = \frac{Q^{**}}{\kappa_1}$$

Regarding n_2, notice that firms which chose the right technology in the first period keep working with it at capacity κ_1, whereas some of the firms that initially chose the wrong technology switch to the right one and at the mass-production level. It follows that

$$n_2 = \frac{n_1}{2} + \frac{Q^* - \frac{1}{2}\,Q^{**}}{\kappa_2}$$

Substituting the above equations for n_1 and n_2 in the condition for a shake-out, $n_2 < n_1$, and simplifying, we obtain Assumption 4. ■

References

Abernathy, William J, and James M Utterback (1978), "Patterns of Industrial Innovation," *Technology Review* **80**, 41–47.

Anderson, Philip, and Michael L Tushman (1990), "Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change," *Administrative Science Quarterly* **35**, 604–633.

Cabral, Luís (1993), "Experience Advantages and Entry Dynamics," *Journal of Economic Theory* **59**, 403–416.

Cabral, Luís (1995), "Sunk Costs, Firm Size and Firm Growth," *Journal of Industrial Economics* **43**, 161–172.

Dixit, Avinash, and Robert Pindyck (1994), *Investment Under Uncertainty*, Princeton, NJ: Princeton University Press.

Dixit, Avinash, and Carl Shapiro (1986), "Entry Dynamics With Mixed Strategies," in L. Thomas (Ed), *The Economics of Strategic Planning*, Lexington, Massachusetts: Lexington Books.

Dosi, Giovanni (1982), "Technological Paradigms and Technological Trajectories," *Research Policy* **11**, 147–162.

Eggers, Jamie P (2007), "Falling Flat: Failed Technologies and Investment under Uncertainty," NYU Stern.

Filson, Darren (2002), "Product and Process Innovations in the Life Cycle of an Industry," *Journal of Economic Behavior and Organization* **49**, 97–112.

Gort, Michael, and Steven Klepper (1982), "Time Paths in the Diffusion of Product Innovations," *Economic Journal* **92**, 630–653.

Hopenhayn, Hugo (1993), "The Shakeout," University of Rochester.

Jovanovic, Boyan, and Glenn M MacDonald (1994), "The Life Cycle of a Competitive Industry," *Journal of Political Economy* **102**, 322–347.

Lieberman, Marvin B, and D B Montgomery (1988), "First-mover Advantages," *Strategic Management Journal* **9**, 41–58.

Mueller, Dennis C, and John Tilton (1969), "Research and Development Costs as a Barrier to Entry," *Canadian Journal of Economics* **2**, 570–579.

Klepper, Steven (1996), "Entry, Exit, Growth, and Innovation over the Product Life Cycle," *American Economic Review* **86**, 562–583.

Klepper, Steven, and Elizabeth Graddy (1990), "The Evolution of New Industries and the Determinants of Market Structure," *RAND Journal of Economics* **21**, 27–44.

Klepper, Steven, and John Miller (1996), "Entry, Exit, and Shakeouts in the United States in New Manufactured Products," *International Journal of Industrial Organization* **13**, 413–614.

Lambson, Val E. (1991), "Industry Evolution With Sunk Costs and Uncertain Market Conditions," *International Journal of Industrial Organization* **9**, 171–196.

Petrakis, Emmanuel, Eric Rasmusen, and Roy Santanu (1997), "The Learning Curve in a Competitive Industry," *RAND Journal of Economics* **28**, 248–268.

Schilling, Melissa A (1998), "Technological Lockout: An Integrative Model of the Economic and Strategic Factors Driving Technological Success and Failure," *Academy of Management Review* **23**, 267–284.

Tushman, Michael, and Philip Anderson (1986), "Technological Discontinuities and Organizational Environment," *Administrative Science Quarterly* **31**, 439–456.

Utterback, James M (1994), *Mastering the Dynamics of Innovation*, Boston, MA: Harvard Business School Press.

Utterback, James M, and William J Abernathy (1975), "A Dynamic Model of Product and Process Innovation," *Omega* **3**, 639–656.

Vittas, Nikolaos (2000), "On Entry, Exit, and Coordination with Mixed Strategies," *European Economic Review* **44**, 1557–1576.

Wernerfelt, Birger, and Aneel Karnani (1987), "Competitive Strategy Under Uncertainty," *Strategic Management Journal* **8**, 187–194.