Entrepreneurship and Innovation: The (Micro) Theory of Price and Profit

William J. Baumol[1]

Abstract. This article is a step toward inclusion of the entrepreneurs, with their critical role in innovation and growth, into elementary mainstream microtheory. It describes how market pressures tend to *enforce* discriminatory pricing even in the basic Schumpeterian model. Moreover, the discriminatory prices the market imposes on those engaged in the innovation process can even be expected to satisfy the requirements of Ramsey optimality. The paper also proposes an explanation of the negative economic profits of invention and innovative entrepreneurship that empirical studies report. This may also help to account for the substantial share of breakthrough innovations apparently provided by small entrepreneurial enterprises. For the low wages of inventors and entrepreneurs can make it profitable for the large firms to outsource radical innovation to the small firms, the former specializing instead in R&D directed to incremental improvement.

Economic growth of the past two centuries, in which entrepreneurs arguably played a critical role, can be estimated to have increased real U.S. per-capita income by an astonishing 2,000 percent or even more. It seems clear that even elimination of most of the static "market failures" in the usual list would be incapable of making a nearly comparable contribution. Yet in the microtheoretical materials of our mainstream journals and textbooks, much space is devoted to these market failures, while entrepreneurs are the invisible men. They may be listed as a fourth "factor of production," but little more on the subject is usually even mentioned. As Joseph

[1] Harold Price Professor of Entrepreneurship and Academic Director, Berkley Center for Entrepreneurial Studies, New York University; and Senior Economist and Professor Emeritus, Princeton University. I am very grateful to Elizabeth Bailey, Alan Blinder, Diego Comin, Robert Hahn, Michael Levine, Sidney Ludvigson, Melissa Schilling, Richard Sylla and, as usual, to Sue Anne Batey Blackman for very helpful comments and suggestive ideas. Special thanks must be reserved for Bronwyn Hall, whose superior knowledge of some of the most critical facts contributed greatly to evaluation of the substance of this piece and many of whose comments are quoted verbatim in this article. I am also deeply indebted to the Ewing Marion Kauffman Foundation for its generous support of this work. A portion of this material was previously presented at an American Enterprise Institute-Brookings Institution meeting where I received extremely helpful comments.

Schumpeter (1911) reminds us, the appropriate cliché is that here we have a performance of Hamlet from which the Prince of Denmark is absent.

This article is meant as a step toward filling that significant gap. And in thus working toward incorporation of the entrepreneur into established value theory, I have deliberately sought to rely largely on the most elementary and common microeconomic tools, so as not to preclude entry of the analysis even into the basic textbooks.[2] But despite its simplicity the model does provide a number of surprises. For example, it describes how *competitive* market pressures tend to *enforce* discriminatory pricing in this arena, leaving the actor with no other option, and show that this is implicitly at the heart of the Schumpeterian model. Although the form taken by Schumpeterian discrimination seems different from that in usual models of differential pricing, we will see that the formal analysis of these forms of price discrimination is just the same. The analysis also indicates that discriminatory pricing, rather than being (as is often asserted) an exercise of monopoly power, is actually *imposed* by competitive pressures.[3] Moreover, the discriminatory prices the market thus enforces on entrepreneurs and others engaged in innovation can even be expected to satisfy the requirements of Ramsey optimality.

But perhaps the most surprising conclusion is that it is possible at all to construct such an elementary value theory. A number of writers, including the present author, have

[2] This paper does not seek to offer anything like a full description of the entrepreneur's activities, which are evidently more complex than anything described here. Rather, it is to investigate how the entrepreneur deals with the issues examined in standard microtheory. For good reasons, value and distribution microtheory does abstract from the particular tasks carried out by an input. A plumber repairs pipes, but distribution theory, ignores this and focuses on his prices.

[3] Of course, we all recognize this for the case of the airlines, which are subject to some forms of vigorous competition that drives economic profits to zero, and even lower, in markets characterized by extreme price discrimination. But the literature does not seejm to recognize how widespread such arrangements are. For a general discussion of price discrimination in the presence of competition and the implications for anti trust law see Baumol and Swanson (1993), with the analysis already having received explicit approval of the U.S. Supreme Court 547 U.S.____(2006).

offered reasons why it could not be done. I have elsewhere argued, first, that because innovation is by definition the ultimate heterogeneous product, generalizations about it are very difficult, perhaps almost impossible. Second, in mainstream economics, microtheory is generally composed of static equilibrium models in which, structurally, nothing is changing, thus, by definition, excluding the entrepreneur. The entrepreneur is absent from such a model because she does not belong there. Standard theory of the firm focuses on repetitious decisions in an enterprise that is already present and fully grown. In such a scenario, the entrepreneur has already completed his job and left for other places where his talents for firm-creation can be used.

I will be dealing throughout this article with innovative rather than what I have called "replicative" entrepreneurs—that is, I will concentrate on persons who launch a firm with the aid of a new product or some other innovation, rather than those who open a company similar to a multitude of already extant enterprises. The theory offered here also rests on the fact that much of today's innovative activity is carried out by the economy's giant corporations, although, contrary to Schumpeter's prediction, this has not marginalized the role of the independent entrepreneur. I will argue that what has happened, rather, is that the two have taken on critical complementary roles in the economy's innovative activity. I will also show that the pricing models applicable to the two types of innovative enterprise are very similar in formal structure.

To test whether the material that follows actually constitutes the beginnings of a systematic microanalysis of the inputs of innovation—the entrepreneur, the independent inventor and the large firm with significant R&D activity—we must find out whether the analysis provides answers (preferably ones that can be expressed operationally in mathematical or at least graphic terms) to four questions: (1) How is the supply of the inputs determined?[4] (2) What are the earnings of these inputs and how are they determined? (3) What determines the prices of the products of these inputs and what are the properties of these prices? (4) What are the welfare implications of these supplies, earnings and profits? I hope to show that the material provided here offers systematic answers to these questions, and thus indeed does constitute the elementary theoretical microanalysis I am seeking. Although the answers are undoubtedly far from complete, I trust that they provide a sufficient foundation for further exploration by others.

The Background: Division of Innovative Tasks between Small and Large Firms

As a foundation for the analysis, I must briefly outline the striking evidence indicating that private innovative activity has been divided by market forces between small firms and large firms, with each tending to specialize in a different part of the task (see Baumol, 2002a, for more detail). Even though the preponderance of private expenditure on research and development (R&D) is provided by the giant business enterprises, a major share of the innovative breakthroughs of recent centuries stems from firms of very modest size. These radical inventions then have been sold, leased or otherwise put into the hands of the giant companies, which have proceeded to develop them—adding capacity, reliability, user-friendliness and marketability—to turn them into the novel consumer products that have transformed our lives. I have referred to this

[4] As will be noted again later, since the entrepreneur is self-employed, the demand for her services is not directly market determined. Nevertheless, some use of a demand relationship will be made below.

division of labor as the "David-Goliath partnership," the value of whose combined contributions clearly exceed the sum of the parts.[5]

The small enterprises continue to make critical contributions to the market economies' unprecedented growth and innovation accomplishments. Without breakthroughs such as the airplane, FM radio and the personal computer, all introduced by small firms or a few individuals working independently, life in our economies would be very different. Moreover, without these breakthrough inventions, the big companies would not have the items upon which to confer their improvements, each of which by itself is unexciting, but when accumulated can constitute enormous advances. This is typified by the evolution of the Wright brothers' primitive aircraft into the Boeing 777 or the steady improvement of the computer chip, whose computational speed has grown over three decades by some 3 million percent. Such incremental innovation by large firms is typically carried out in a routinized manner, avoiding the great risks of breakthrough invention, with management often deciding not only the R&D budget and personnel, but even what the firm's R&D division should next seek to invent.[6]

---

[5] The empirical evidence on the David-Goliath partnership hypothesis is quite ambiguous. For an excellent compendium of evidence on this matter and references to the literature, see Audretsch (2006) and particularly the illuminating analyses in the articles by Audretsch and Acs reproduced in that book.

[6] This story is consistent with a recent Booz Allen Hamilton study of "…1,000 publicly held companies around the world that spent the most on R&D in 2004," which reported that the odds of success "...increase when senior management makes sure that there is clear customer demand [for the innovation to be worked upon] and a profitable way to bring the innovation to market before any project gets the go-ahead" (Paul B. Brown, "R.&D. under the Microscope," *New York Times*, December 24, 2005, p. C5, reporting on Barry Jarulzelski, Kevin Dehoff, and Rakesh Bordia, "The Booz Allen Hamilton Global Innovation 1000: Money Isn't Everything," *Strategy+Business,* Issue 41, pp. 1-14, New York, NY: Booz Allen Hamilton, available at http://www.strategy-business.com. However, it should be made clear that here and elsewhere there remains considerable disagreement in the empirical studies on matters related to entrepreneurship. For example, Chandy and Tellis (2000) conclude on the basis of their empirical data on 64 "radical innovations" that "contrary to conventional wisdom…today's incumbents and large firms account for many radical innovations, especially since World War II" (p. 12), and "Large firms account for a substantially larger proportion of radical innovations relative to their number in the economy" (p. 11).

Innovators' Price Discrimination: Several Subvariants

I begin my price discrimination argument in general terms that apply to the innovative entrepreneur and his inventor partner (who can, of course, be the same person). And the general form of the argument, as will be shown later, applies equally to the innovative large enterprises. For all this I must distinguish three variants of ordinary price discrimination: interpersonal, interlocational, and intertemporal. The most commonly discussed variant—interpersonal price discrimination—is illustrated by student discounts, which divide customers into separate classes to which a commodity is sold at different prices. Interlocational (geographic), discrimination occurs when prices differ from one city or one airline route to another, even with equal supply costs. Intertemporal discrimination is illustrated by reduced prices in the annual post-Christmas sales. Firms may use several of these, as when a retailer declares Tuesdays to be "senior citizen discount days."

Noteworthy is the common analysis that applies to all these pricing practices. Profit maximization requires equal marginal revenues in all the market segments; prices will be highest in submarkets where the firm's product demand is most inelastic; and, other things being equal, the absolute elasticity will be highest in submarkets with the largest number of competitors, whether the discrimination is interpersonal, interlocational or intertemporal. Oligopolistic innovators most often employ prices that are interpersonally or interlocationally discriminatory, while for the Schumpeterian entrepreneur-inventor partners the intertemporal variant lies at the heart of the story.

I will next show how, contrary to widespread assertion, competition characteristically tends to enforce discriminatory price taking, particularly where firms are motivated repeatedly to incur sunk outlays, as is common in R&D activity.

Mandatory Price Discrimination in Competitive Markets

The analysis of the discriminatory pricing scenario that, I claim, applies to the independent entrepreneur rests on five, fairly common assumptions:

> (1) Firms can enter and exit markets at low cost and with little delay, doing so whenever there are profits to be earned—the markets are “contestable” with zero Stiglerian barriers to entry, so that expected equilibrium profits must also be zero.[7]
>
> (2) Customers are divisible into different submarkets with different, finite, demand elasticities (negatively sloped demand curves) for the firm.
>
> (3) If such a customer group is offered a lower price than another, it is not feasible, or it is at least difficult, for the former to resell purchases to the latter.[8]
>
> (4) The average cost curves of the firms are roughly U-shaped, at least in the relevant portion of the loci.
>
> (5) The firm sinks R&D outlays repeatedly, expenditures that do not enter marginal costs since total R&D cost does not rise when users increase.

The second and third premises are, of course, the standard assumptions of any price discrimination model. The first assumption—zero entry barriers and entry whenever profits are available in a market—is, of course, an extreme case. Indeed, interpreted

---

[7] Of course, there are also innovative firms that do possess market or monopoly power, and for them one cannot presume that they will earn zero profits. These cases clearly require further investigation. But even an industry in which entry into the *production process* is difficult may not be safe from relatively easy entry by inventors and entrepreneurs. Even Microsoft appears to be concerned about such entry. And where entry is effectively impeded, the firms will still want discriminatory prices to enhance profits.

[8] Such segmentation is virtually automatic in the case of intertemporal discrimination, in which high prices are succeeded by price reductions, because there is no way someone who buys at the subsequent high prices can resell the purchase in the earlier period—bygones are bygones. Of course, in other cases such segmentation of customers incurs costs that are often substantial, although sometimes it is easy and almost automatic (e.g., senior citizen discounts for theater admission). Where the cost of segmentation is significant, a model can determine the profit-maximizing level of such expenditure by the firm.

strictly, it rules out sunk costs, for the need to incur such costs constitutes a barrier to entry, imposing risk costs upon the entrant from which the incumbent is immune.

But the sunk costs important here are not once-and-for-all outlays. Rather, they are continuing expenses as when competition forces the firm in a high-tech industry repeatedly to sink money in R&D. As we know, in these industries, a firm that introduces a superior product or production process cannot rest on its oars. For is engaged in a kind of arms race with its rivals, with expenditure on further innovation the prime weapon, and success in the race ultimately a matter of life and death for the firm.

The sunk costs that are said by traditional theory not to matter for an incumbent firm's decisions, but do constitute barriers for the potential entrant, are the once-and-for-all expenditures made by the incumbent in the past and not repeated, but are not yet incurred by the potential entrant. For the old firm this is the ancient history that no current decision can change, whereas the sunk outlays relevant here are those that are still to be incurred and will repeatedly be sunk again for the foreseeable future. It is these expectable and recurring sunk outlays that most directly drive the firm to discriminatory pricing. They are not barriers to entry in Stigler's pertinent sense, because they are equal burdens for the entrants and the incumbents. So they offer no substantial competitive advantage and, hence, no monopoly power to an incumbent firm.

The five premises just described are sufficient for my central result:

Proposition 1. Entry enforces discrimination. In a market with no barriers to entry by enterpreneurs, the firm's equilibrium economic profits will be zero. But if, in addition, the entrepreneurs can separate their customers into distinct submarkets with different demand elasticities of the firm's submarket demand curves, and the firm can prevent its product from being transferred from one customer to another, the normal assumptions of the theory of the firm require discriminatory prices to avoid losses, so that equilibrium will entail such prices. This result is applicable widely but holds in particular for the products of an innovative entrepreneur.

It should be noted that this result rejects the widespread view that monopoly power is required for prices to be discriminatory. However,, the assumed differences in demand elasticities in the different submarkets do preclude *perfect* competition, with its universally horizontal supplier demand curves. In particular, the analysis fits in with the case,, common in practice, where competing innovations introduced by rival entrepreneurs are not identical, and are not valued equally by their buyers. Then effective competition will force the prices of these products to follow a Ricardian rent pattern, with nonzero rents to the owners of the intellectual property and profits net of rents equal to zero. In equilibrium, buyers will divide their purchases among the competing innovations so that the prices of items considered inferior are sufficiently lower to compensate for the lower valuation of those items.

It should be clear, intuitively that the first three of my five assumptions yield Proposition 1. For if freedom of entry forces the firm to charge its profit-maximizing prices in order to break even, and those prices permit it to do so, then those are the prices it will have to select in equilibrium. But if there exist discriminatory prices that yield profits higher than those that are possible under uniform pricing then, *tautologically*, the (zero profit) equilibrium profit-maximizing prices must be discriminatory. Moreover, if the maximum is unique, the firm will be a price taker, with a unique vector of discriminatory prices, dictated element-by-element by the market. Note that I have not yet used the U-shaped average cost premise, which will be employed in discussing existence and stability of the competitive, discriminatory equilibrium.[9]

Need Entry Undercut Discriminatory Prices, and Can Discriminators Coexist?

The preceding arguments raise at least two questions: Will not the assumed easy entry force prices toward uniformity? And will the discriminating firm be able to take over the market and evolve into a monopolist? The general answer to both questions is that sometimes either can happen, but not normally in the situation described by the model. To deal with all this most directly but heuristically, I translate the model into elementary diagrams. For simplicity, assume that there are two customer groups, each with downward-sloping and linear demand curves, $p_i = AR = a_i - b_i y_i$ $(i = 1, 2)$. Their marginal revenues will also be linear, with slopes $-2b_i$ (Figures 1a and 1b) for submarkets 1 and 2, respectively. As usual, to find the profit-maximizing decisions for the firm as a function of its total output in the two submarkets, we add the marginal revenue curves horizontally, to obtain the kinked marginal revenue curve for the firm as a whole (MRm in Figure 1c), since profit maximization at any total output level requires marginal

[9] Here it is important to emphasize that the equilibrium considered here is not part of a static depiction of the behavior of the firm and the market. First, there is no guarantee that the equilibrium will ever be attained even approximately, though even those we can describe as "mainstream Austrians" speak of "…those important tendencies which markets display toward continual discovery and exploitation of pure profit opportunities thus tending to nudge the market in the equilibrium direction." (Kirzner, 1997, p. 73). Moreover, the equilibrium entails a choice of level of R&D expenditure by the firm, an outlay whose purpose is, surely, to *prevent* stationarity.

revenues in the two markets to be equal.

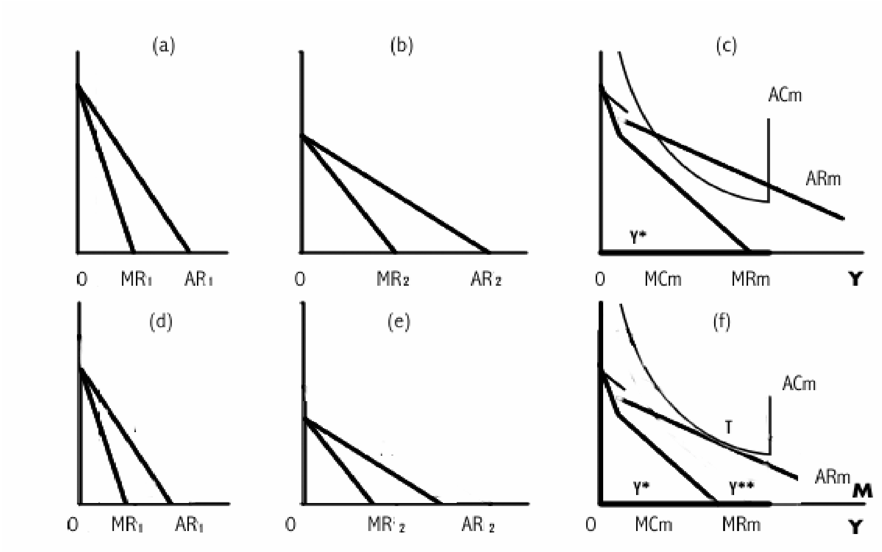

Figure 1

From the marginal revenue curve for the two submarkets together, we can derive a curve of average revenue, AR_m, for the firm as a function of its total output when the price in each submarket maximizes the total revenue derived from that level of y. To the left of y*, the firm's average revenue curve will have the usual relationship to submarket 1's MR curve, both starting at the same point and MR having twice the downward slope of AR. But the firm's average revenue curve has a downward discontinuity at y*, as shown, because at higher values of y the firm obtains part of its revenue from a submarket with lower average returns than the initially-served submarket, reducing the average of the returns from the two sources. Thereafter, to the right of y*, the average revenue curve will be linear with a slope whose absolute value is smaller than that of either submarket.

Now, suppose that part of the average cost curve for the combined market, AC_m, lies below the market AR locus, as in Figure 1c. Then the availability of positive profits and unimpeded entry will force a downward (leftward) shift in one or both of the submarket demand curves, which will proceed to the point of tangency, T, between the combined AR and the AC curves, if such a point exists (Figure 1f).[10]

With appropriate curvatures of the average cost and revenue loci, there will be a unique tangency at output, y**, and at any other output level the firm will incur losses. So at any output other than the equilibrium tangency point in Figure 1f, the firm will be unable to survive. This is where the fourth assumption, the U-shaped average cost curve for the firm, comes into play. The diagrams show that with linear submarket demand curves, this is sufficient (but not necessary) for that result. What is required, clearly, is that the curvature (convexity) of the AC curve be greater than that of the AR curve; that the former contain a downward sloping segment whose slope (if continuous) decreases in absolute value as y increases from a level initially higher than that of the corresponding point on the total market AR locus; and that the slope of AC exceed that of AR.

This yields answers to the two questions at the beginning of this section. First, we see that nothing in our equilibrium analysis precludes existence of multiple competing firms. The assumed cost and demand relationships prevent expansion of the firm's output above the equilibrium amount. Consequently, total market demand for the product may well be many times as large as the equilibrium output of our firm, so there can coexist a plurality of such firms with similar demands, costs and outputs, in a stable industry equilibrium. Second, in such an equilibrium it is not possible for the discriminatory

[10] In Figure 1, for illustration, the AC curve is that for a total cost that is fixed, and capacity is absolutely limited. However, none of the discussion depends on these premises.

pricing to be undermined by “cream-skimming” entry that attacks only the most lucrative submarkets. For, as depicted in the graphs, any firm with the same revenue and cost possibilities will incur losses if it serves fewer submarkets than are served in our firm’s equilibrium. It follows that, even with the competitive pressures that characterize a market with absolute freedom of entry and exit, prices need not be driven toward uniformity. While successful entry into our industry is possible, at least in the unique equilibrium case, it will only be an entrant opening for business fully grown to the equilibrium output who will be able to survive.

The discussion also has implications about existence and uniqueness of the discriminatory price equilibrium. What has been shown here indicates that with little more added (such as continuity of the relevant portions of the derivatives of the AC and AR loci), we can expect such an equilibrium to exist. Similarly, aside from the obvious possibility of a multiplicity of tangency points, made implausible by the U-shaped AC curve assumption, uniqueness would appear to be assured. But all this calls for further exploration.

Discriminatory Price Makers or Price Takers?

The discussion also shows that ease of entry deprives the firm of choice in the setting of prices. If the profit-maximizing price vector for the firm is unique, and that price vector and no other yields zero profits, then the firm will have no choice. Any enduring deviation from that vector must be suicidal. The firm will, effectively, be a price taker, though not one that follows a publicly posted price in a market like that for pork bellies.[11]

[11] Because only the price-taker side of the firm’s activities has been discussed, there is little room for insights from game theory.

This is not mere theory. In practice, marginally surviving firms scramble for every perceived source of potential revenue, and are driven to price discrimination. Neither impecunious theaters nor marginal airlines can tolerate empty seats. Each seeks desperately to fill them with students, leisure travelers and others who will pay a price that contributes *something* exceeding marginal cost while having to impose high prices on last-minute business customers. Experience may enable these firms to approximate the most lucrative prices and, if they fail, they will be replaced by rivals.

If the constraint on profit imposed by entry is potent, the only way for the firm with large fixed and continuing sunk costs to survive will be actually to engage in price discrimination of the most sophisticated variety that is workable. It follows that, where entry is easy, *price discrimination is not to be taken as an unequivocal manifestation of monopoly power*.[12] The firm with discriminatory prices may or may not be competitive.

Ramsey Optimality of the Price Taker's Discriminatory Prices

I come next to my second central result:

Proposition 2. Ramsey Optimality. If the profit-maximizing equilibrium that yields zero profit with discriminatory pricing is unique, then it is a Ramsey optimum.

That is, the equilibrium will entail the vector of prices that is Pareto optimal, subject to the constraint that the (expected) economic profits of the firm (and the industry) are zero. The argument is almost trivial. If this is the only set of prices that meet the zero-profit requirement, then no other prices can satisfy the Ramsey constraint.

---

[12] On this point, I have my only disagreement with Hausman and Mackie-Mason's excellent and illuminating article, when they speak of, "…the necessary monopoly power for price discrimination to take place" (1988, p. 245 fn). For the origin of the argument that discriminatory pricing need not require monopoly power, see Levine (2002).

Hence, there are no other zero-profit prices that can add to consumers' plus producers' surplus, or that benefit some individuals without harming anyone.[13] That is all there is to the argument. However, we can do a little better than this, deriving the requirements for the equilibrium and showing that they lead to the usual Ramsey formulas. For simplicity, I deal with the case where the marginal costs of serving consumers in the different submarkets are all the same and equal to C', while the demands of the different submarkets are independent, so that all cross elasticities of demand are zero. Then, using obvious notation, equilibrium requires

(1) $$\text{Max } \Sigma\, p_i y_i - C(y_1,\ldots,y_n)$$

Subject to

(2) $$\Sigma\, p_i y_i - C(y_1,\ldots,y_n) \leq k \text{ (where we will select } k = 0).$$

The Lagrangian is

(3) $$L = (1- r)\,[\Sigma\, p_i y_i - C(y_1,\ldots y_n)] + rk,$$ where r is the Lagrange multiplier.

But for any $y_i > 0$, the first-order (Kuhn- Tucker) conditions for maximization of (3) obviously include

(4) $$(1- r)\, p_i\, [1 + (y_i/p_i)\, dp_i/dy_i\,] = (1- r)\, p_i\, [\, 1 - 1/E] = (1-r)\, C' \text{ or}$$

(5) $$(p_i - C')/\, p_i = 1/E,$$

which is, of course, the most elementary form of Ramsey equation.

[13] Of course some innovations are contributed by firms that possess monopoly power, and for them the zero-profit constraint patently does not hold. If for them all the Proposition 1 assumptions except the easy-entry premise are satisfied, their economic profits will generally be positive. Discriminatory prices will still serve their interests, and the profit maximizing discriminatory prices will still meet the requirements of Ramsey second-best welfare optimality, as shown by the argument in the text with k > 0.

Innovative Entrepreneurial Firms and Intertemporal Price Discrimination

The preceding discussion applies generally to all competitive markets that satisfy the five premises. Let us see now how it applies, in particular, to independent inventors and their entrepreneur partners. The same story is *precisely* what the Schumpeterian scenario shows. Only the way the tale is told disguises its price discrimination side. For the form it takes entails what was described above as *intertemporal* price discrimination. In essence, in that scenario the submarkets occupy different dates rather than different locations. The entrepreneur and inventor recoup most of their outlays of money and effort in the early submarket, soon after introduction of an innovation. Then there is no competition, so the seller's demand curve is inelastic and the profit-maximizing price is high. But in later submarkets, when imitators appear, presumably offering imperfect substitute products, the (entrepreneur) initial seller's demand curve grows more elastic, even if the rivals' products are not identical, and price must fall. To get back to more lucrative operation she must invest in further innovation, and so she is driven repeatedly to sink costs in the process. Obviously, there is no difference between the formal analysis of the discrimination among customer groups described in the preceding sections and the discrimination among time periods described by Schumpeter. For this story, Figure 1 remains totally unchanged except that now Figure 1a must be interpreted to represent the firm's year t submarket, while Figure 1b now pertains to year t+1 submarket, rather than to a set of submarkets with different customers, all in year t. So Propositions 1 and 2 remain applicable to Schumpeter's model. Moreover, as we will soon see, in this respect, the analysis of the independent entrepreneur's pricing under market pressures is structurally similar to that of the innovative oligopoly firm.

Profits and Supply of Innovative and Entrepreneurial Services[14]

It is clear that Schumpeter called for positive economic profits to the entrepreneur, at least in any growing economy: "Without development there is no profit, and without profit there is no development" (1911, 1936, p. 184). The analysis here, up to this point, in contrast, implies that such profits will be zero. But the empirical evidence indicates clearly that we were both wrong. The expectable economic profits of the representative entrepreneur in reality are apparently distinctly negative. How can this be? And what are the implications?

Before turning to answers offered by theory, I briefly describe some of the evidence. Thus, for example, Freeman (1978) and Benz and Frey (2004) show that the average earnings of self-employed individuals are significantly lower than those of employees with similar qualifications, and the same is presumably true, in particular, of the self-employed innovative entrepreneurs. There are at least two studies that support this hypothesis for innovative entrepreneurs. Astebro (2003) reports on the basis of a sample of 1,091 inventions that, "The average IRR [internal rate of return] on a portfolio investment in these inventions is 11.4 percent. This is higher than the risk-free rate but lower than the long-run return on high-risk securities and the long-run return on early-stage venture capital funds…the distribution of return is skew; only between 7-9 percent reach the market. Of the 75 inventions that did, six received returns above 1400 percent,

[14] Because we are at this point focusing on self-employed inventors and entrepreneurs, the demand function is not directly applicable (but see below). However, as I will argue next, the structure of supply in this market is observable and distinctive. Here it may be noted that the special attributes of *supply* rather than demand are focused upon throughout basic distribution theory. The distinctive supply conditions of the different factors of production are often employed as the primary attribute that distinguishes one from another. Land is fixed in supply. Physical capital is a produced good so that the profit calculation determines its quantity, while the supply of labor is largely determined exogenously. In contrast, at least in marginal productivity theory, there is virtually no difference among the factors in the way their demands are determined.

60 percent obtained negative returns and the median was negative" (p. 226). Perhaps even more striking is the recent work of Nordhaus (2004), whose calculations show how little of the efficiency rent goes to the innovator: "Using data from the U.S. non-farm business section, I estimate that innovators are able to capture about 2.2 percent of the total surplus from innovation. This number results from a low rate of initial appropriability (estimated to be around 7 percent) along with a high rate of depreciation of Schumpeterian profits (judged to be around 20 percent per year)....the rate of profit on the replacement cost of capital over the 1948-2001 period is estimated to be 0.19 percent per year" (p. 34).[15]

Two mechanisms may explain the dismal economic profits of entrepreneurs and the failure of innovative and entrepreneurial effort supply to dry up as a consequence. They can be referred to as: (1) the superstar reward structure; (2) the psychic rewards to innovative activity, and I will discuss them next.

Superstar Market Reward Structure, or the Multimillion-Dollar Lottery. The most obvious incentive to which one can attribute the willingness of independent inventors and their entrepreneur partners to pursue more radical ideas, despite their low expected return, is, of course, the great wealth and enormous prestige that success promises. There are the inventor-entrepreneurs who are enduring legends: Eli Whitney, James Watt, Elias Singer, Thomas Edison, the Wright Brothers, etc., etc. The enormous prestige and great financial rewards, *along with their rarity,* transform the innovative entrepreneur's activities into a lottery that offers just a few mega-prizes. It is an occupation with a few superstar positions, and the prize is available only to those who

[15] Using a cruder and more-intuitive approach, the present author also reached a very low figure for the returns to innovation that are not dissipated in spillovers (see Baumol, 2002b, pp. 134-5).

provide *breakthrough* innovations. A technological contribution that permits humanity to fly or to send messages through the air can elicit headlines, but a minor improvement in automobile door handles is hardly likely to compete. And just as multimillion-dollar lotteries have a greater attraction than the thousand-dollar lottery at the local club (even when the latter's terms are far better actuarially), the pursuit of breakthrough innovations surely has a very special attraction to the entrepreneur.

In part, the willingness of innovators, like the buyers of lottery tickets, to accept the biased terms is probably attributable to over-optimism[16] or to sheer miscalculation, like that of the buyers of mega-lottery tickets. But that is hardly the end of the story

Monetary Compensation, Psychic Compensation. Both activities—innovative entrepreneurship and purchase of lottery tickets—also provide an important payoff of a second sort. Each offers distinct psychic rewards, and not only to those who have already achieved success or who even have a real and substantial likelihood of success. The *prospects* of glory, of wealth and fame are something of value even if they never materialize. They are, indeed, the stuff that dreams are made of. And for the entrepreneur, contemplation of imagined success is only part of the psychic reward. The biographies of the great inventors bring out the fascination their work elicited, their moments of

[16] The available evidence certainly supports the notion that entrepreneurs are unrealistically optimistic. Sixty-eight percent of respondents to a survey of American entrepreneurs by Cooper et al. (1988) thought the odds of their businesses succeeding were better than for others in the same sector, while only five percent thought that they were worse. Recently, Puri and Robinson (2005) measured optimism as the difference between self-reported life expectancy and the survey respondents' statistical life expectancy based on smoking-, age-, education-, race- and gender-corrected mortality tables. Puri and Robinson find all respondents over-optimistically expect to live longer than the life tables suggest, but that entrepreneurs are substantially more likely to think they will live longer. These authors also establish that optimism is significantly positively associated with the propensity to be an entrepreneur." (Parker, 2006, p. 2) On the other hand, Xu and Ruef (2004) report that though "Entrepreneurs have long been assumed to be more risk-tolerant than the general population….our empirical results consistently show that nascent entrepreneurs are more risk-averse than non-entrepreneurs."
-

triumph, and even the pleasure of puzzle-solving and experimentation, though punctured by frustration and exhaustion.

So, while the representative entrepreneur may indeed be underpaid in terms of financial reward alone, his *total* payoff may be closer to what economic theory leads us to expect, though part of the payoff takes a form other than money. It is as though he were being paid off in two different currencies: partly in dollars, partly in euros. In equilibrium, such two-coin payees can clearly expect fewer dollars than someone similarly engaged whose contract calls for payment only in that one currency.[17]

The story recurs throughout the economy. The fact that multimillion-dollar lotteries are openly unfair actuarially does not mean that the purchasers who endure long queues to grab up the tickets are irrational. They may in fact receive an adequate payment in another currency: the psychic rewards. That same argument also helps to explain why, despite the rigors of their training and their performances, the earnings of professional dancers are so meager.[18] One can easily think of other examples.

The market mechanism enforces this, as Adam Smith pointed out: Given two occupations, one very distasteful and the other a source of great pleasure, if other things including payoffs and ability requirements were equal, we must expect the workforce to

[17] This suggests one way in which it may sometimes be possible to place a monetary value on psychological enjoyment and even esthetic pleasure. A similar situation has been noted in other arenas. For example, there are data showing that the average financial return to investment in works of art is significantly lower than the return to investment in bonds, the difference being interpreted as the financial valuation of the aesthetic yield of painting ownership (see Frey and Pommerehne, 1989).

[18] Other examples include scientific research and academic occupations. It may also hold for the self-employed in their enjoyment of freedom from control by superiors (see Hamilton, 2000; and Benz and Frey, 2004). This phenomenon and its relation to the work of innovators has long been recognized: "The knowledge of the man of science, indispensable as it is to the development of industry, circulates with ease and rapidity from one nation to all the rest. And men of science have themselves an interest in its diffusion; for upon that diffusion they rest their hopes of fortune, and, what is more prized by them, of reputation too" (J.-B. Say, 1819, 1834, p. 82).

shun the one and flock to the other, driving wages up in the former and depressing them in the latter as a garden-variety manifestation of supply and demand.[19]

On the Longer-Run Supply of Innovative Entrepreneurship

I turn next to the long-run supply of *productive* entrepreneurial activity, as distinguished from that of entrepreneurship that is engaged in activities that are unproductive or even destructive. In the nonmathematical growth literature, it is often asserted that the appearance of an abundant supply of entrepreneurs effectively stimulates growth, while shrinkage in the economy's cadre of entrepreneurs is a significant impediment. But this appearance and disappearance of the entrepreneurs is left as a mystery, related to cultural developments and changes in other psychological and sociological influences, more or less vague in character. As I have described earlier (Baumol, 2002b, Chapter 5), a review of the historical evidence suggests an alternative explanation, less magical or science-fictional in character. This holds that entrepreneurs do not suddenly appear from nowhere or just as mysteriously vanish. Rather, entrepreneurs are always with us, but as the structure of the rewards offered in the economy changes, the entrepreneurs switch the locus of their activity, moving into arenas where the payoff prospects have become most attractive. And in doing so, they may move into activities that are most generally recognized as entrepreneurial *and productive*, exchanging them for other activities that also require considerable enterprising talent but are often quite distant from the production of goods and services, and often even impede

[19] "The wages of labour vary with the ease or hardship, the cleanliness or dirtiness, the honourableness or dishonourableness of the employment…. A journeyman weaver earns less than a journeyman blacksmith. His work is not always easier, but it is much cleanlier…..The exorbitant rewards of players, opera-singers and opera-dances, &c. are founded upon these two principles: the rarity and beauty of the talents, and the discredit of employing them in this manner. It seems absurd at first that we should despise their person, and yet reward their talents with the most profuse liberality. While we do the one, however, we must of necessity do the other" (Smith, 1776, Book I, Chapter X, Part I).

production. Just as technological change led workers and engineers to reallocate themselves from canal building to railroad construction, and then on to still more-modern enterprises, the entrepreneurs have also reallocated themselves in accord with institutionally induced changes in the structure of payoffs to the different occupations. Thus, when there is a pertinent change in the institutions that govern relative rewards, entrepreneurs will be induced to shift their activities among the affected fields of endeavor and so the set of productive entrepreneurs will appear to expand or contract autonomously. In sum:

Proposition 3. Entrepreneurship is a resource that is subject to reallocation between productive and unproductive activities. The supply of *productive* entrepreneurs is heavily influenced by the institutional arrangements that determine the relative payoffs to the two types of activity.[20]

That is the essence of the longer run supply-side of the theory of entrepreneurship.

Oligopolistic "Red-Queen" Innovation Games and Mandatory Price Discrimination

As asserted earlier, breakthrough innovations characteristically still have a long way to go before they achieve their market and utility potential, when they leave these initial providers. Preponderantly, innovations tend to be taken over and developed by the

[20] The role of institutions in influencing the supply of entrepreneurs is, of course, not new, and has been emphasized by a number of authors, notably North (see North and Thomas, 1973, and North, 1990). What may be new about Proposition 3 is the assertion that institutional changes do not do this primarily by inducing creation of a body of new entrepreneurs where there were few before, but by enticing enterprising individuals away from their previous unproductive activities and leading them to transfer to productive undertakings. In a personal communication, Richard Sylla gives a striking example of this critical conclusion: "Meiji Japan… reformers commuted peasant rice payments to Samurai into government bonds, giving the Samurai government bonds and taxing the peasants in money to pay interest on the bonds. The Samurai were encouraged to become investors… and bankers… and Japan with a modern financial system suddenly left the rest of Asia in the dust and caught up with the West. The Samurai leaving fighting and becoming bankers must surely be a classic example of an institutional change enticing enterprising individuals away from previous unproductive activities and leading them to transfer to productive undertakings.'"

“Goliath” firms. So, to complete the story, we must turn to the R&D-related activities of those large enterprises.

Elsewhere, I have described the R&D activities of the oligopoly firms in the high-tech industries as an “arms race” (Baumol, 2002b, especially Ch. 3). Following Schumpeter, I emphasized that these firms now regard such things as price to be weapons of *secondary* importance in their competitive battles. The primary weapon has clearly become the new or improved product that these firms race to introduce before a rival can bring to market an alternative more attractive to consumers. No firm in this position dares to fall behind in the race to create new and better products, because protracted failure to do so can be fatal. Just as Lewis Carroll’s Red Queen in *Through the Looking Glass* had to run as fast as she could in order to stay in the same place, so must each firm constantly come up with new products in order to preserve its position in the market. This Red-Queen game is a key attribute of the advanced economies that helps to account for their continuing outpouring of innovations. It is clearly far more powerful than just a high monetary reward, which allows the recipient to rest on his laurels and withdraw from further innovative activity. An innovation arms race permits no rest, lest competitors outperform and thereby kill off the enterprise of anyone who becomes inactive.

It is noteworthy that this incredibly effective and, indeed, ingenious arrangement is yet another product of the market mechanism, not consciously invented, designed or imposed by any person or group. Unlike the institutional developments that first induced the expansion of productive entrepreneurial activities, largely ascribable to historical accident, the Red-Queen game attribute of the innovation process is automatically introduced by oligopolistic competition and is therefore yet another of the contributions of the market mechanism itself.

All this yields three critical observations: First, the oligopolistic high-tech firm finds itself unable to avoid frequent and substantial reinvestment in R&D, without which it cannot survive, but it also cannot survive if its pricing approach precludes recoupment of those outlays. Second, as I have argued,[21] relative ease of entry into the innovation process and vigorous oligopolistic competition may well condemn even many of the large oligopolistic firms to expect to earn near-zero economic profits on their R&D investments (as empirical evidence appears to confirm),[22] but with the opportunity to recoup their repeatedly sunk R&D outlays, because entry will not occur with any lower level of earnings. Third, because the firm's R&D outlay underlying an innovation need not be increased by an expansion in the number of customers for the product, none of this R&D expense will enter marginal cost, so that marginal-cost pricing will generally not be viable for such an enterprise. But before dealing further with such pricing issues, let us consider the amount of R&D expenditure to which the innovative oligopolists are driven by market forces.

---

[21] In Baumol (2002b), I cite evidence (p. 40 fn) that, on average and taking account of both the spectacular successes and spectacular failures, the earnings of all computer-related activities together have yielded something close to the normal (competitive) rate of profit. Bronwyn Hall, who is as well-informed on the subject as anyone, reports, on the basis of work underlying an article in the course of completion, that "It does not yet answer [the] question precisely, [but]…I think it does contain evidence that the return to R&D in large corporations has been driven towards normal returns over the past 20 years" (personal communication, October 2006).

[22] Of course, some such firms possess market power that can be substantial. This can clearly open the way to positive profits. But those firms, too, will be enabled to enhance their profits if they can adopt discriminatory pricing and, as shown above, if those prices do indeed maximize profits they will be still be Ramsey optimal, but now subject to the constraint that they yield the available monopoly profits.

A Kinked-Revenue-Curve Model of Spending on Innovation

Despite the zero-profit prospect, oligopolists are driven to keep up their R&D spending by an understandable preference for *zero* economic profits—that yield only normal competitive returns—over *negative* profits, investor flight and insolvency. We can make the technological scenario underlying the investment decision more explicit with the help of a microeconomic construct very similar to the old kinked-demand-curve model that was proposed to explain why prices tend to be "sticky" in oligopoly markets. The underlying mechanism in that model is an asymmetry in the firm's expectations about its competitors. The firm hesitates to lower its price for fear that its rivals will match the price cut, so that the firm will end up with few new customers, but dramatically reduced revenues. But it also fears that if it increases its price, the others will *not* follow, so that it will be left all by itself with an overpriced product. It is concluded that a firm with such beliefs will set its price at the prevailing industry level, no more and no less, and leave it there unless the situation changes drastically.

The essence of the analogous process of R&D level determination is the fact that in a market in which innovation is the prime instrument of competition, the firm's expected profit depends on the quality of its anticipated new products and processes *relative to those of its competitors,* which, in turn, depend on the relative amounts the firms spend on R&D. For simplicity, I assume expected total revenue functions linear in the single variable ([R&D]w – [R&D]c), where the first of these terms is the R&D outlay of our maximizing firm, w, and the second term is that of its representative competitor, c. The total cost function is also assumed linear, but its variable is only our firm's own R&D expenditure. Then we have the two functions, total revenue = TR([R&D]w –

[R&D]c), and total cost = TC([R&D]w). We graph this under two extreme and alternative assumptions about competitor behaviour: *Case 1:* [R&D]c = k (constant), so the rival does not change its behavior no matter what our maximizing firm does, and *Case 2:* [R&D]c = [R&D]w, so that the rival imitates exactly, whatever level of spending our firm, w, selects.

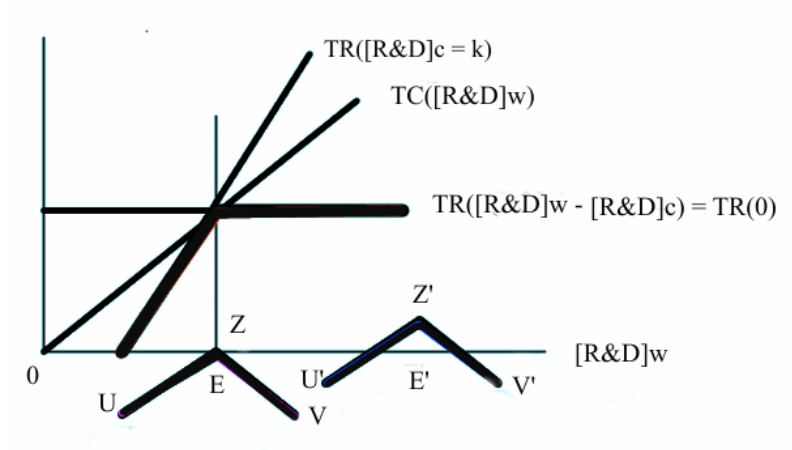

Figure 2

Our firm's total *cost* will be upward sloping (Figure 2), no matter what its rival’s decision rule. TCw is drawn through the origin, but that is irrelevant for the analysis.[23]

However, there are two alternative total *revenue* curves. The straight line labeled TR([R&D]c = k), as noted, corresponds to the case where competitors do not respond to our firm's spending decision. Then our firm's (expected) total revenue locus is obviously also upward sloping, as in the figure—the more our firm spends, the more its product can be expected to improve relative to that from the unchanged outlays of its rivals.

But we must also consider the case where rivals precisely follow our firm's expenditure level. This total revenue curve, labeled TR({[R&D]w – [R&D]}c = 0) is a

[23] These properties are, of course, oversimplifications, because the firm's total production cost will vary with its sales volume, which must vary, in turn, with its own and its competitor's R&D outlays. But this does not affect the structure of the analysis or its conclusions, and we can avoid the problem by assuming that all costs other than those of research and development are zero.

horizontal line, because its one variable takes the same value (zero) throughout. Let point E on the horizontal axis be our firm's initial R&D spending, above which all three lines meet, indicating that the firm earns normal (zero) economic profit. Following the usual logic of kinked-curve scenarios, it is plausible (and supported by some empirical evidence) that, if our firm's management considers whether to change the amount of its R&D spending, it will reason as follows: If we raise our expenditure and our rivals don't, they will be threatened with severe loss of customers; so we can expect them quickly to match our spending increase, and we will therefore lose money because we will receive only our old revenues despite our increased outlays. But if we spend less, we will also lose money because our rivals will be tempted not to cut, thereby stealing our customers, so our revenues will fall sharply, probably more than costs.[24]

Two conclusions follow. First, the actually expected total revenue curve (the heavy kinked line), UZV, will be composed of two segments, each a part of the other two TR lines. To the left of point E, segment UZ corresponds to the revenue curve with the rivals' R&D spending constant, while the segment, ZV, of the expected revenue curve to the right of E follows the curve where rivals match our firm's outlay rise, so spending by all firms remains equal.

The second conclusion is that our firm, like its rivals, will be driven to continue unchanged its current level of spending on R&D, which we can call the "industry spending norm," because either a rise in our expenditure or a fall will not just reduce our profits but actually cut profits below the competitive level. This is shown by the kinked total profit curve whose peak is E. This profit curve just touches the horizontal axis at Z,

[24] This scenario is consistent with the 2005 Booz Allen Hamilton study of the 1,000 biggest R&D spender firms in the world in 2004 (cited in footnote 6 above) that reported, "While spending more does not necessarily help, spending too little will hurt."

at the current spending point, E, since competition prevents our firm from earning more than zero economic profits, and the TP line lies below that axis everywhere else. The TP curve is, of course, obtained by plotting the vertical distance between the kinked thick total revenue curve, the revenue expectation locus, and the total cost curve.

That is the stationary equilibrium story of the determination of the amount of R&D spending by the firm in an innovation arm's race. But it is only a snapshot picture. Moreover, it gives little substantive information because it does not indicate how equilibrium point, E, is selected but only, whatever its magnitude, that it will be sticky.

To find out more, we must proceed from our snapshot to a motion picture of which the portion of Figure 2 that has already been described is only a single frame. Motion *will* occur if our firm happens to hit upon a particularly promising invention that justifies an increase in its R&D spending from its previously sticky level. Then the total profit curve will shift upward and to the right, to the kinked line U'Z'V' whose peak, Z', lies above point E' in the figure. To protect their market positions, our firm's rivals will soon be forced to raise their spending, too, so the industry norm will rise from E to E' in the graph. *This result clearly is a ratchet mechanism.*[25] It allows spending to increase when changing conditions make it profitable, but permits no reductions, because reductions cut profit to negative levels.[26]

---

[25] Bronwyn Hall comments "[T]he ratchet effect…certainly seems to be there….I looked at the median and aggregate R&D intensity in a few quasi-2-digit industries for the past 30 years….I have not studied them carefully (in particular, the pharma sector seems to be sensitive to the presence of a lot of biotech firms that have little sales, so the median fluctuates a lot)….Another support of the ratchet model that you did not mention: Many firms seem to choose their R&D budget by looking at industry R&D to sales ratios and matching. That is, the relevant variable is R&D intensity rather than level" (personal correspondence, 2006). I may add that I agree fully with the required correction of my text called for by this last observation.

[26] This statement somewhat exaggerates the effectiveness of ratchets in preventing the economy from *ever* sliding backward in its R&D expenditures. After all, even in machinery, ratchets sometimes slip. Firms

The analysis, if valid, predicts how R&D investment will change over time. But because of this mobility, the analysis does not identify any one level as the unique stationary equilibrium. Indeed, it cannot do so, because no one such equilibrium exists, except very temporarily and, in effect, chosen by accident.

Thus we conclude:

Proposition 4. Quantity of R&D funding by an oligopolist in an innovation arms race. Oligopolistic competition that uses innovation as its primary weapon forces the firm's R&D expenditure to follow the current industry spending stamdard; but occasionally a firm will exceed this R&D outlay, thereby forcing its rivals to raise their spending..

Such ratcheting is a critical part of the mechanism that produces the extraordinary growth records of free-enterprise economies and differentiates them from all other known economic systems. Competitive pressures force firms to run as fast as they can in the innovation race just to keep up with the others.

Outsourcing of Breakthroughs: Low Cost of Entrepreneurs' Psychic Benefits

There is an additional important implication for innovative oligopolists' behavior in our earlier entrepreneurial remuneration story. So far, I have not suggested any role of the market mechanism in the disproportionate allocation of breakthrough innovation activity to the independent inventors and entrepreneurs rather than the corporate giants. But there is such a role. The underlying mechanism is found in the fact that psychic benefits are a very tangible reward to the entrepreneur innovation supplier, but are

---

may, for example, be forced to cut back their R&D expenditure if business is extremely bad. Or they can simply make mistakes in planning how much to spend on investment, or they may be discouraged by repeated failures of their research division to come up with saleable products. The economy's ratchets are indeed imperfect, but they are there. They cannot completely prevent backsliding in R&D expenditure, but they can be a powerful influence that is effective in resisting such retreats.

generally *costless to the innovation buyer*. So an innovative entrepreneur who on average receives great pleasure but meager financial rewards from the activity may nevertheless be richly rewarded overall. But the low *financial* payment means that innovations obtained from this source are cheap in monetary terms, giving these suppliers a marked competitive advantage. So the independent innovative entrepreneur will tend to be the more economical provider of breakthrough innovation to the economy. This means that the low-cost psychic reward component of the independent innovator's compensation will make it more economical for the large firm, in considering its make-or-buy options, to acquire its breakthroughs more generally from others, rather than seeking to provide them in-house.[27] Firms are forced to do so for fear that if they do not, their rivals will. This, then, suggests one market-based reason[28](that is not mere happenstance) why a disproportionate share of radical innovation stems from the independent entrepreneur.[29]

[27] Independent entrepreneurs who happen to care only about monetary payout may also choose to purchase intellectual property for the low price demanded by the inventor. Such transactions probably do occur, perhaps not infrequently. However, that does not exclude the large high-tech firm from this activity.

[28] Bronwyn Hall comments: "…it is a longstanding stylized fact in innovation that radical change usually comes from small firms/individuals and incremental change from large firms. There does not seem to be a single explanation. The idea that this is because inventors are willing to accept a lower average rate of return makes sense, but I tend to think there are other reasons. [I agree—see below, WB] An important one is the nature of technology trajectories and sunk R&D costs—large firms are good at what they currently do (that is how they got large) and face much lower costs than new entrants when improving existing products, because they don't have to sink the learning costs to develop their current level of capability again. New entrants can imitate (*pace* patents) but a pretty substantial investment is necessary to reach the same capability. That explains why large firms are the experts at incremental improvement.

On the other side (radical invention), I had some conversations with biotech/pharmaceutical-financial types which suggested to me that large established firms find it more difficult to shut off risky R&D projects, even when they are not viewed as successful. This leads to a preference for alliances with biotechs rather than bringing them in-house. The reasons have to do with the demoralizing effect on the scientists involved, which is contagious, and the inertia built into a larger organization. Terminating a particular alliance contract seemed easier. I do not know how widespread this type of problem is. I do agree that in the case of really radical innovations, the motivator is not expected return—there is some other utility component that is not financial. Or there is over-optimism. Or both." (personal communication, 2006)

[29] Why does this low-wage competitive advantage of the independent innovator-entrepreneur not extend also to the cumulative incremental improvements that are a giant firm specialty? At least part of the answer is the greater complexity and investment cost characteristic of the latter. A Boeing 777 is obviously far more complicated than the primitive device the Wright brothers made airborne at Kitty Hawk, and the

There are important additional reasons for the large firms to outsource innovative activity that is targeted at breakthroughs, because that is a way of shifting the most extreme risks to the independent entrepreneur. Normally, one would think that it is the large firms that are in a better position to deal with risk than is the independent entrepreneur, because they are apt to have larger reserves to fall back on, and can afford to undertake a number of projects simultaneously, so that the law of large numbers would be on their side. But the evidence suggests that their deeper pockets are not enough to induce them to focus on revolutionary invention.[30]

There are a number of additional incentives for large firms to devote themselves primarily to incremental innovations that build upon their current product technology rather than aiming at breakthrough ideas.[31] First, large firms are likely to have invested heavily in complementary assets and skills that make their existing technologies more valuable (or economical). This makes the payoff to incremental innovation in current lines of business more attractive for the large firm than for a new entity without such investments.

---

transformation of the Boeing 747 into the Boeing 777 entailed an army of engineers and designers and an expenditure that made the outlays of the Wrights dramatically insignificant by comparison. This, too, is not accidental. By its very nature, this revolutionary invention, like so many before it, grew ever more complex as it was repeatedly modified and improved. Thus, the independent innovator was and continues to be at a marked disadvantage in the financing of incremental improvements of inventions that have reached an advanced stage of sophistication.

[30] Schramm points out a variant of the cost-saving and risk-avoidance explanations for large firm outsourcing of breakthrough invention: "Large U.S. firms today effectively outsource much of their research and development to start-ups. Rather than take on all of the effort and risk of developing an idea internally, they help a new firm do so via strategic investments and working partnerships. There are many twists on this strategy. Intel, for example, tries to build markets for its chips by investing in companies that develop new systems and products that will use the chips; it has invested in more than a thousand such start-ups" (2006, p. 157).

[31] For a review of some of the research on the areas of advantage and disadvantage of large firms relative to small firms in the innovation process, see Schilling, 2005, particularly Chapter 10.

Second, the large firm is likely to have made significant strategic commitments to customers and suppliers that to a degree dictate what the firm must invest in: it must continue to invest in delivery of the goods and services to which it is already committed and upon whose revenues the firm relies. Third, large firms face large ongoing operating costs that are not easily or quickly reduced. This creates a significant source of inertia for the firm, and makes it much riskier for a large firm to alter its technological course than for a new or small firm without such commitments.

Third, in the competitive race to hone their competencies and become more efficient than their peers, large firms have typically refined a set of employee incentives and norms that induce employees to behave in a way that maximizes their effectiveness and efficiency in the firm's *current* activities. But these very incentives and norms simultaneously make it difficult to evaluate and reward employees for creating new competencies or creating radically new technology. Consequently, these incentives and norms can make the pursuit of radical innovation unattractive for employees of the large firm. Related to this, if creative individuals within the firm believe they have recognized a breakthrough technological opportunity with enormous profit potential, they may prefer to pursue this opportunity outside the firm where they can appropriate more of the earnings for themselves. This can lead individuals with entrepreneurial tendencies to leave their large employer to pursue technological opportunities on their own.

Finally, large firms choose to focus on incremental innovation because work on breakthrough ideas clearly is beset by great uncertainty. One probably has to have a touch of madness to devote most of one's time and resources to the untried prospects that a breakthrough innovation unavoidably entails. Indeed, as we have seen, the evidence suggests that innovative entrepreneurs characteristically are self-selected risk lovers, that is, persons attracted by the prospect of magnificent prizes, who disregard the low probability of their attainment. This is a much more extreme type of predisposition than that of the lottery ticket buyer who, after all, risks little if the few tickets he purchases do not pay off. Rather, the inventor is apt to risk all, sometimes the welfare of his family, perhaps led on by the belief that success in his enterprise is not just a matter of chance, but depends on his own ability and persistence.

The large firm, in contrast, is characteristically driven to seek to avoid risks or at least to minimize them, because failure in the innovation process can even threaten survival and, where entry is easy, profits and reserves are apt to be low, making it even more difficult to cover the cost of projects that prove disappointing.[32] The bottom line is that market forces drive the large innovative firms to risk-avoidance to the extent that is permitted by innovative activity, and this, too, forces them toward focus on incremental activity, leaving the enormously risky pursuit of breakthroughs to the independent inventor and his entrepreneur partner. We can sum this up as:

Proposition 5. Market-Driven Assignment of the Tasks of Innovation. The assignment of a large share of breakthrough innovations to independent inventors and entrepreneurs and of incremental improvement to large firms is driven by market forces such as the effects of psychic rewards on financial costs, the relatively high investment cost of incremental improvements, and the different risk-bearing propensities of the two groups.

[32] "In established businesses, innovation is mostly shaped through small, incremental steps of additional features to augment basic functionalities. With short product lifecycles, time to recoup R&D investments is limited…. Success is relatively predictable through the execution of well-defined innovation processes and in-depth knowledge of their markets in the respective business units" (Ad Huijser, Ph.D., former executive vice president and chief technology officer, Royal Philips Electronics, The Hague, September 2003).

There seems to be no reason to expect the market forces just described to be very transitory. If they are indeed enduring, it follows that we can expect the current division of innovative labor between small and large firms to continue.

There are implications of this division of the task for pricing and revenues that are also of some interest and not immediately obvious. Here, the most elementary supply-demand diagram tells the story. We assume a population of prospective entrepreneurs and also assume initially that their products are demanded by a set of vigorously

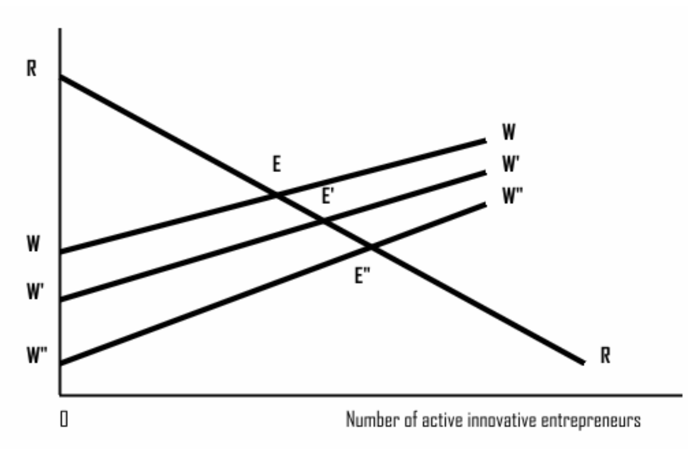

Figure 3

competing oligopolists whose profits are driven to zero. The horizontal axis in Figure 3 indicates the number of entrepreneurs/inventors employed in this market. The vertical axis reports *monetary payments exclusively.* The downward sloping curve, RR, represents the average revenues the oligopolists obtain per (self) employed entrepreneur and inventor. There are also three upward-sloping earnings curves, representing money earnings per entrepreneur and inventor as a function of the level of employment in three sets of circumstances. WW represents the situation in which only money payments are valued by these labor suppliers, and there is no psychic income or underestimation of risk. The next lower curve, W'W', represents the case where psychic benefits constitute part of labor earnings, an amount indicated by the vertical distance between WW and W'W'. Finally, W"W", the lowest of these curves, corresponds to the case where the entrepreneurs are characterized both by psychic benefits and underestimated uncertainties. With zero profits, the three corresponding equilibrium points are E, E' and E", respectively.

Four conclusions, that may not have been obvious before, follow immediately:

(1) At E', the entrepreneurs and inventors can be said to obtain full compensation for their efforts, as much as they do at E. But in the former, part of the compensation takes the form of psychic benefits.

(2) However, at E'', they are underpaid relative to what they receive at E because they lose out as a result of their underestimation of risk and so receive too little insurance contribution from buyers of their products.

(3) Although costs of the large firms are reduced from the height of E to that of E' or E'', the profits of these firms do not rise because the assumed competition will force them to pass on all the savings to the final-product consumers. This is the counterpart, on the cost side, of Proposition 1, which focused on the revenue side of the story. Here, we deal with firms are that are forced by competition to take every opportunity to reduce costs, but any decreased cost only prevents them from earning negative profits.

(4) The comparative static reduction in per entrepreneur payment by the large-firm customers will be less than the reduction in per-entrepreneur earnings at a given level of employment, that is, the vertical distance between curve W"W" and WW will be more than the difference in height between Points E'' and E. The explanation is the same as the corresponding result for the shifting and incidence of taxation under perfect

competition, for the reduction in money wages at issue is formally equivalent to a tax cut. These results undergo only one change if the competition of the oligopolists is insufficient to eliminate any market power they possess. In that case, the relevant shift is in their marginal cost curves rather than their average cost curves, and because their profits are not driven to zero they may be able to keep some of any reduction in money wages for themselves.

Conclusion: Notes on the Value Theory of Entrepreneurship and Innovation

The preceding discussion amounts to the construction of a model that can determine the prices of the services of entrepreneurs, independent inventors, and the large firms with substantial R&D activities. With entry into innovative activities largely unobstructed, and the big firms driven by an innovative arms race to constantly repeated investment in the innovation process, the analysis indicates that these large firms will be driven toward normal profits on their R&D outlays, and the quest for survival will prevent them from abandoning the activity. In addition, given the low marginal cost of use of the resulting information, they will be forced to adopt the market-determined discriminatory prices that maximize their profits in order to recoup their R&D investments. And, by Proposition 2, those prices will tend to be Ramsey optimal.

A very similar story holds for the independent entrepreneurs and inventors, and entails a slight reinterpretation of Schumpeter's scenario. Moreover, here and elsewhere (see Baumol 2002b, Chapter 6), I have uncovered some insights on the welfare properties of the supplies, earnings levels and prices of the innovative inputs. All of this implies that the analysis offered here does constitute a substantial beginning of the relevant value theory. These constructs seem to represent progress toward restoration of the entrepreneurs to their rightful place in value theory—making them visible, once again, as they clearly need to be for an operational microeconomic theory of innovation and growth. The neophyte economics student need no longer be fed exclusively on a diet of static theoretical constructs, whose relevance seems questionable to many of these textbook readers. Rather, the microeconomist's standard tools can be employed to yield insights on innovation and growth, which surely merit the role of primary goals of long-run economic policy.

Bibliography

Astebro, Thomas. 2003. "The Return to Independent Invention: Evidence of Unrealistic Optimism, Risk Seeking or Skewness Loving." *The Economic Journal*, January: 226-238.

Audretsch, David B. 2006. *Entrepreneurshio, Innovation and Economic Growth,* Cheltenham, UK: Edward Elgar.

Baumol, William J. 2002a. "Entrepreneurship, Innovation and Growth: The David-Goliath Symbiosis." *Journal of Entrepreneurial Finance and Business Ventures* 7 (2). Fall: 1-10.

________________. 2002b. *The Free-Market Innovation Machine: Analyzing the Growth Miracle of Capitalism.* Princeton, NJ: Princeton University Press.

________________ and Daniel G. Swanson, "The New Economy and Ubiquitous Competitive Price Discrimination: Identifying Defensible Criteria of Market Power." 2003. *Antitrust Law Journal,* 70, 661-685.

Benz, Matthias, and Bruno S. Frey. 2004. "Being Independent Raises Happiness at Work." *Swedish Economic Policy Review* 11: 95-134.

Chandy, R. K., and G. J. Tellis. 2000. "The Incumbent's Curse? Incumbency, Size and Radical Product Innovation." *Journal of Marketing* 64 (3). July: 1-17.

Cooper, A.C., C.Y. Woo and W.C. Dunkelberg. 1988. "Entrepreneurs' Perceived Chance of Success." *Journal of Business Venturing* 3 (3): 97-108.

Freeman, Richard B. 1978. "Job Satisfaction as an Economic Variable." *American Economic Review* 68 (2): 135-141.

Frey, Bruno S. and W.W. Pommerehne. 1989. *Muses and Markets: Explorations in the Economics of the Arts.* Oxford, UK: Basil Blackwell.

Hamilton, Barton H. 2000. "Does entrepreneurship pay? An empirical analysis of the returns to self-employment." *Journal of Political Economy* 108 (3): 604-631.

Hausman, Jerry A., and Jeffrey K. Mackie-Mason. 1988. "Price Discrimination and Patent Policy." *RAND Journal of Economics* 19: 253-256.

Kirzner, Israel M. 1979. *Perception, Opportunity and Profit.* Chicago, IL: University of Chicago Press.

______________. 1997. "Entrepreneurial Discovery and the Competitive Market Process: an Austrian Approach." *Journal of Economic Literature* XXXV. March: 60-85.

Levine, Michael E. 2002. "Price Discrimination without Market Power," *Yale Journal on Regulation* 19 (1). Winter: 1-36.

Marx, Karl. 1867, 1915. *Capital.* Vol. I. Chicago: Charles H. Kerr & Company.

Nordhaus, William D. 2004. "Schumpeterian Profits in the American Economy: Theory and Measurement." Cambridge, MA: National Bureau of Economic Research, Working Paper 10433.

North, Douglass C. 1990. *Institutions, Institutional Change, and Economic Performance.* Cambridge, UK and New York, NY: Cambridge University Press.

_______________, and Robert P. Thomas. 1973. *The Rise of the Western World: A New Economic History.* Cambridge, UK, and New York: Cambridge University Press.

Say, Jean-Baptiste. 1834. *A Treatise on Political Economy.* Philadelphia: Claxton, Remsen and Haffelfinger (French original, 1819).

Schilling, Melissa A. 2005. *Strategic Management of Technological Innovation*. New York: McGraw Hill.

Schramm, Carl J. 2006. *The Imperial Imperative.* New York: Harper Collins.

Schumpeter, Joseph A. 1911. *The Theory of Economic Development.* Cambridge, MA: Harvard University Press (English translation, 1934).

Smith, Adam, 1776. *An Inquiry Into the Nature and Causes of the Wealth of Nations,* London.

Supreme Court of the United States. 2006. Illinois Tool Works v. Independent Ink, 547 U.S., 1-17.

Xu, Hongue, and Martin Ruef. 2004. “The Myth of the Risk-Tolerant Entrepreneur.” *Strategic Organization* 2 (4): 331-355.