PERFECTLY COMPETITIVE INNOVATION

MICHELE BOLDRIN AND DAVID K. LEVINE

ABSTRACT. Based partially on the belief that innovation is not possible under perfect competition, many thousands papers have been written about the nature of innovation under monopoly or oligopoly. In fact, competitive rents can and do sustain innovation in the complete absence of monopoly power. However, little is known about the source and significance of these rents, or about the way in which innovative activity takes place under conditions of competition. We begin to remedy this imbalance by examining the way in which competitive innovators earn rents both in theory and in practice.

Keywords: Economic Theory, Innovation, Perfect Competition, Intellectual Property

JEL Classification: A1; A2

Date: First Version: October 1999, This Version: 9th August 2006.

This paper is based upon work supported by the National Science Foundation under Grants SES 01-14147, and 03-14713. Boldrin also acknowledges research support from the Spanish BEC2002-04294-C02-01 and the University of Minnesota Grants in Aid Program for financial support. Over the last seven years, since the very first draft of this paper was circulated, we benefited from comments and criticisms from more people that we are capable of remembering. We are grateful to each and everyone of them.

Corresponding Author: Michele Boldrin, Dept. of Economics, Washington University in Saint Louis, St Louis, MO, . Email: mboldrin@econ.wustl.edu.

1. Introduction

> *It would be efficient ex post to make the existing discoveries freely available to all producers, but this practice fails to provide the ex ante incentives for further inventions. A tradeoff arises between restrictions on the use of existing ideas and the rewards to inventive activity.* Robert Barro and Xavier Sala-i-Martin (1999)
>
> *Evidently, we must depart from the common practice in neoclassical growth theory of assuming that all firms act as price takers in an environment of perfect competition. Firms must be able to sell their products at prices in excess of unit production costs if they are to recover their up-front outlays on research and development. In other words, some imperfect competition is necessary to support private investments in new technologies.* Gene Grossman and Elhanan Helpman (1994)

The idea that monopoly is the necessary handmaiden of innovation is well established in economics since Schumpeter's [1942] famous book. Analytically, it rests on a fundamental principle that has gained almost universal acceptance: *ideas are public goods*, Arrow [1962], Shell [1966, 1967]. In the macroeconomic literature, the theme is central to the "new growth" theory of Romer [1990] , Grossman and Helpman [1991] and Aghion and Howitt [1992]. Indeed, this idea is such an integral part of conventional wisdom that most textbooks speak in terms similar to the passage from the Barro and Sala-i-Martin textbook quoted above.

That an idea is widely and strongly believed and quoted does not, however, make it true. This is an attempt to cast doubt on the claim that monopoly is necessary for innovation, both as a matter of theory and as a matter of fact. We do not claim complete originality: economists such as Stigler [1956] seem to have explicitly rejected the Schumpeterian claim since its outset, and recent authors such as Irmen and Hellwig [2001], Boldrin and Levine [2002], or Zeira [2006] have produced growth models in which innovation is assumed to take place absent monopoly. However, while many thousands of papers have been written exploring the microeconomics of innovation under monopoly none, to the best of our knowledge, has examined the incentives to innovate "assuming that all firms act as price takers in an environment of perfect competition."

This paper begins filling that gap by providing a menu of models in which innovation happens in competitive equilibrium, which is significant for two reasons. First, innovation under competition has been and

still is pervasive in history and in practice and cannot be understood in the context of a model that rules it out as a matter of principle: what your model cannot see, you cannot measure. Second, there is a vigorous public debate over intellectual property legislation, and an informed debate about policy is not possible if the consequences of eliminating the current system of intellectual property cannot even be conceived theoretically.

The theory of innovation is about the costly creation and adoption of new goods or factors of production. Profit maximizing entrepreneurs will neither create nor adopt if market prices do not compensate them for their total unit cost. If - due to some fixed cost - the unit cost of output exceeds marginal cost while the market price of output equals it, innovation will not be undertaken. It is widely believed that competitive equilibrium always results in prices equal to marginal cost. Hence the belief that competition is inconsistent with innovation. However widespread this belief may be, it is not correct. It is true only in the absence of capacity constraints, which, as we know from the Marshallian model of a competitive industry, is at most a long run tendency. Since innovation is a dynamic process and only a limited number of units of a new good is ever available, as is painfully evident to students of general equilibrium analysis, fixed factors earn competitive rents. These competitive rents can and do compensate innovators for their initial cost. The mechanics of competitive innovation rests on this simple observation, which goes as far back as Marshall's *Principles of Economics*.

Some authors appear to believe, and strenuously argue, that capacity constraints are irrelevant to the theory of innovation because ideas flow costlessly and instantaneously once they are created. The irony of professors - whose substantial salary derives entirely from the difficult and time consuming nature of communicating old ideas - arguing that ideas are freely and costlessly transmitted public goods, should not escape us. We will debunk the empirical validity of this idea in our case studies.

On the theoretical side, we go through a menu of models of competitive innovation, from simple to complex. We also consider such possibilities as unpriced spillover externalities. The key point is that competitive rents are reduced, but still positive. This highlights the issue of appropriability. Competitive equilibrium enables innovators to appropriate some of the social surplus from their innovation, but

not all of it.[1] Hence, ideas that generate a great deal of social surplus relative to the cost of creation will be be invented regardless of whether or not there is monopoly. On the other hand, some marginally socially valuable ideas will not be produced under perfect competition. It is worth noticing, though, that competition over marginal ideas is not likely to be fierce and, to the extent that loss of appropriability is due to the imitative effort of competitors, appropriability for marginal ideas may still be quite high. The welfare tradeoff is complex, since policies that increase appropriability increase other distortions as well. We do not explicitly consider welfare tradeoffs here,[2] but provide a model that allows for such tradeoffs to be studied. It seems clear that, if the point of departure is a model in which perfect competition allows no appropriation, then it is not possible to correctly analyze welfare.

As we have indicated our goal is to argue that monopoly is not necessary for innovation either in theory or in fact. Having examined the theory, we turn, in conclusion, to the evidence. We present a series of case studies of industries in which innovation thrives in the absence of the government enforced monopoly of intellectual property. As a matter of practice, the inventor always has some small degree of monopoly power before his idea becomes widely used. Our goal is to understand both the role of competitive rents and the role of the "first mover" advantage that comes from the initial degree of "natural" monopoly. To achieve this, it seems clear, one needs a theoretical model allowing for both of them.

So let us grab our mallet, enter the glass house of conventional wisdom, and shatter a few illusions.

2. Models of Competitive Innovation

2.1. **Simplest Case: The Tree and its Fruits.** We start with the most elementary situation: an isolated innovator who is contemplating a fixed cost of C to create her new idea. Once the fixed cost is incurred, the innovator will have a single initial copy of a new good - the first copy of the new novel, the first prototype of the new machine, or more broadly the template by which copies may be created, that is: *the seed*. A tree grows from each seed and it is the fruits of the tree that yield consumption, hence we treat seeds and trees as equivalent. Since our interest is in perfect competition, we assume the innovator always takes

[1] Monopoly generally allows a greater amount of appropriability, but still less than full appropriability.

[2] See Boldrin and Levine [2006] for a theoretical and empirical attempt in that direction.

the market price as given; further, from the moment in which she sells any fraction of the tree, she is in immediate and perfect competition with the purchaser(s). That is, any purchaser can make copies of the new commodity using exactly the same technology for producing copies that is available to the original innovator. Indeed, to get the ball rolling, we make the extreme assumption that no additional costs are needed to obtain new seeds from the tree - in the process of growing fruits, additional seeds are also created from which new trees grow that may then be sold.[3]

Suppose, then, that at the cost of C a single initial seed is made available in period $t = 0$. In period t, if k_t trees are available then the fruits of those trees are consumed, yielding a utility of $u(k_t)$, and at the same time reproduce themselves and additional copies, so that βk_t trees will be available at time $t + 1$ where $\beta > 1$. Quah [2002] refers to this as the 24/7 case, meaning that trees reproduce themselves 24/7 regardless of what uses their fruits are put to. As for ownership, we assume that each owner of a tree is endowed the following period with β copies. We assume also an infinitely lived representative consumer with a discount factor of $0 \leq \delta < 1$.

This model is straightforward to analyze. In the first period and subsequently, the price of trees is proportional to marginal utility. Since the initial seed is the only fixed factor, there is constant returns to scale, and there is competition to produce subsequent trees, all competitive rents accrue to the innovator. The number of seeds in circulation grows exponentially: $k_t = \beta^t$, as does the consumption of fruits. The innovator's rents are easily computed to be the discounted present value of the revenue stream from renting trees

$$q_0 = \sum_{t=0}^{\infty} (\beta\delta)^t u'(\beta^t).$$

Unless the single initial tree satiates the market so that $u'(1) = 0$ this competitive rent is always positive, and the innovator will be willing to innovate provided that $q_0/C \geq 1$.

Making the initial innovator behave competitively even in the very first period may be the source of misunderstanding. Since the innovator, by necessity, has a monopoly in the initial period, why do we

[3]The careful reader will notice that this does not imply an unpriced externality, insofar as the copies so created are priced into the value of the original seed. Clearly, these special assumptions have purely the purpose to get started and will be disposed of in the next subsection.

not take account of her incentive to restrict the initial supply of seeds? There are several answers to this question.

- The initial idea or template typically comes in a single indivisible unit, and it may not be feasible or practical to sell less than the single initial unit.
- In many cases, some considered below, there is no incentive to restrict initial supply, in which case the innovator chooses not to use her monopoly power.
- As a matter of practice, any innovator will have some degree of monopoly power early on. Restricting the initial number of units for sale is only one form of exercising such initial monopoly power, and not necessarily the most important one in practice. When we turn to case studies, we will consider how much incentive is due to competitive rents as compared to the choice of the innovator to restrict supply. Naturally our calculation of competitive rent, q_0, gives a lower bound on what the innovator may expect to earn in recompense for incuring the fixed cost.
- In the not uncommon case of simultaneous innovation, the innovator faces competition even in the initial period.
- The first mover advantage in the broad sense of initial monopoly power has been well studied and documented.[4] The goal of our theory is to highlight the role of competitive rents which have not been so thoroughly examined.
- The common assertion is not that monopoly is important in the theory of innovation because innovators necessarily have some monopoly power, but rather that without some monopoly power they would not choose to innovate. To highlight the fallacity of this argument, and to cleanly isolate the role of competitive rents, we choose in examining the theory to focus on the case where there is no monopoly power, even in the initial period.

2.2. **General Case: The Many Things our Minds Invent.** The case of the seed and the fruit trees, in spite of its wide applicability, is obviously very special - no inputs are used other than copies of the new good that is, therefore, both consumption and capital, and there is no tradeoff between consuming the new good and accumulating productive capacity. In general, one likes to distinguish the new productive input - k: the new pharmaceutical plant and its specialized workers - from the new consumption good - c: the new drug - that obtains from it. Specifically, we want to allow for the case in which "innovating"

[4]See, for example, the classical work of Levin et al. [1987], and the vast literature that has followed it confirming its main findings.

means to introduce a new consumption good and to build up the new productive capacity or specialized input capable of producing it, be it physical capital, human capital, or a combination of the two. Studying this more general case does not change, but substantially sharpens, the basic picture. There will still be a positive competitive rent q_0 accruing to the scarce initial fixed factor, and the innovation will still be introduced if $q_0/C \geq 1$.

Suppose that the representative consumer's utility function is

$$\sum_{t=0}^{\infty} \delta^t [u(c_t) - wL_t]$$

, where c_t is the flow of consumption services and L_t is labor, supplied at the constant wage w. The new production factor, k_t, depreciates at a fixed rate so that, without additional reproduction, ζk_t units are available tomorrow; we allow $\zeta > 1$ to include the 24/7 example as a special case.

Productive capacity and labor may be used either in the consumption or in the copying sector. Let k_t^c, ℓ_t^c be the inputs employed in the first and k_t^k, ℓ_t^k those employed in the second sector; both have neoclassical production functions. Consumption is $c_t = F(k_t^c, \ell_t^c)$, and additional productive capacity is $x_t = G(k_t^k, \ell_t^k)$, the stock of which evolves according to $k_{t+1} = \zeta k_t + x_t$. As before we assume perfect competition from the initial period, when our representative innovator enters the market with k_0 units of productive capacity.

In each period $t = 0, 1, ...$, equilibrium solves two maximization problems. First, given k_t, L_t, and x_t, inputs are allocated to sectors in order to

$$\max_{0 \leq k_t^c \leq k_t, 0 \leq \ell_t^c \leq L_t} c_t = F(k_t^c, \ell_t^c)$$

subject to

$$0 \leq x_t \leq G(k_t - k_t^c, L_t - \ell_t^c),$$

The solution to this yields a production possibility frontier

$$c_t = T(k_t, x_t, L_t).$$

Under standard regularity assumptions on F and G, T is increasing in k_t and L_t and decreasing in x_t and is concave. Define also $\overline{x}(k_t) = \lim_{L \to \infty} G(k_t, L)$ - this represent the greatest amount of new productive capacity that can be produced from a given starting capital stock.[5]

[5]We allow the possibility that $\overline{x}(k_t) = \infty$; if G is strictly increasing then this must be the case.

Second, the period labor supply L_t solves

$$\max_{L_t} u\left[T(k_t, x_t, L_t)\right] - wL_t,$$

which has a unique solution $L_t = L(k_t, x_t)$, for given w. Notice that, from the first order condition

$$u'\left[T(k_t, x_t, L_t)\right] \frac{\partial T}{\partial L}(k_t, x_t, L_t) = w,$$

the relation between L_t and either k_t or x_t is ambiguous, as different factor intensity rankings and the possibility of reversal may lead to a non-monotone $L(k_t, x_t)$. Rule out these altogether irrelevant cases by assuming that $T(k_t, x_t, L(k_t, x_t))$ is increasing in k_t, decreasing in x_t, and has a non-negative cross-partial second derivative. Regardless of factor intensities T is always concave and nothing hinges critically on the regularity assumptions we have hereby made. Define the period return function $V(k_t, x_t) = u\left[T(k_t, x_t, L(k_t, x_t))\right]$; this is also increasing in k_t, decreasing in x_t, and strictly concave in both arguments.

Given k_0, the intertemporal competitive equilibrium of this economy is summarized by sequences $\{k_t\}_{t=0}^{\infty}$ solving

$$v(k_0) = \max_{\{k_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \delta^t V(k_t, k_{t+1} - \zeta k_t)$$

subject to

$$\zeta k_t + \overline{x}(k_t) \geq k_{t+1} \geq \zeta k_t.$$

Notice immediately that $q_0 = v'(k_0) > 0$ under exactly the same conditions as in the simple model, that is: as long as consumers are not satiated by the productive capacity available to the innovator in the first period. Our fundamental result is therefore general: limited capacity is a sufficient condition for competitive innovation to be viable. Let us now learn more by studying a number of special cases.

2.2.1. *Genius and its Apprentices.* Consider first the special case in which labor is not an input, so that $c_t = k_t^c$ and $x_t = \beta k_t^k$. This corresponds to our initial model without the 24/7 assumption: if you consume the harvested corn you cannot use it for seeding the fields next season. On a more urban note, if the creative designer (or the great researcher) spends her time training apprentices (or newly minted PhDs) - who will compete with her next period - her production of beautiful new chairs (or brilliant new papers), will be reduced accordingly. The Bellman equation for this optimization problem is

$$v(k) = \max_{0 \leq c \leq k} \left\{u(c) + \delta v((\beta + \zeta)k - \beta c)\right\}.$$

When this problem is decentralized as a competitive equilibrium, the price of consumption in period t is given by $p_t = u'(c_t)$. From the resource constraint

$$c_t = \frac{(\beta + \zeta)k_t - k_{t+1}}{\beta}.$$

When the optimum involves a strictly positive accumulation rate, which is always true for $\beta\delta > 1$, by standard dynamic programming arguments, the price q_t of copies k_t can be computed as

$$q_t = v'(k_t) = p_t \frac{\beta + \zeta}{\beta}.$$

As $p_t > 0$, $q_t > 0$ for all t as long as $\beta + \zeta > 0$. The zero profit condition here implies that q_t decreases at a rate of $1/(\beta + \zeta)$ per period of time and that $q_t \to p_t$ as $\beta \to \infty$, which, nevertheless, does not imply that $q_t(\beta)$ is a decreasing function, as $p_t(\beta)$ may well be increasing as we show later.

2.2.2. *Harvesting Grapes and Delivering DVDs.* Next, let us go back to a 24/7 production technology, but one in which labor is required along with productive capacity to produce consumption. That is, $c_t = F(k_t, L_t)$ and $k_t = \beta^t$. Consider first the case in which copies and labor must be used in fixed proportion - that is $c_t = \min\{k_t, L_t\}$ - similar to, say, the Netflix or the harvesting technologies. While most people are familiar with the way in which grapes are harvested in the vineyard, the same may not be true for the Netflix example. In the latter, DVDs are kept at a central location, every period they are delivered to users who rent them for one period then return them to the central location. Here, labor is needed for delivering and recovering the DVDs. In this case, the optimum is obviously to allow $c_t = k_t = L_t$ until some threshold c^* is passed, then keep c_t constant as k_t continues to (costlessly and irrelevantly) grow. Utility is $u(c_t) - wc_t$, so the threshold is simply where $u'(c^*) = w$.

Since the competitive rent per unit of consumption services in period t is the marginal social value $u'(c_t) - w$, it decreases as productive capacity accumulates period after period and, at the threshold, it falls to zero. After that period no further rents are earned by the innovator: when productive capacity - that is, DVDs of the new movie, or vineyards - is large enough that the marginal utility of consumption is equal to the opportunity cost of the labor necessary to obtain it, innovator's rents have been completely dissipated by free entry and competitive pricing. This situation, in which competitive rents fall to zero in a finite and possibly short period of time, is the closest one gets to conventional wisdom.

By way of contrast, sticking with the 24/7 technology, consider the case in which the innovation is a new piece of capital equipment producing consumption in combination with a labor input, as in $c_t = (k_t)^\alpha (L_t)^{1-\alpha}$. Suppose also that the utility function has the CES form $u(c_t) = c_t^\gamma$ for $1 > \gamma > 0$. Then the period utility is $(k_t)^{\gamma\alpha}(L_t)^{(1-\alpha)\gamma} - wL_t$. The first order condition for the optimum labor supply is

$$(1-\alpha)\gamma(k_t)^{\gamma\alpha}(L_t)^{(1-\alpha)\gamma-1} = w.$$

This can be substituted back into the utility function to find utility as a function of capacity and the wage. We may then easily differentiate this to find the rental price of capital $p_t = du/dk_t$ and the revenue earned in each period by the owner of productive capacity,

$$p_t k_t = \frac{\alpha\gamma}{1-(1-\alpha)\gamma}\left[\frac{(1-\alpha)\gamma}{w}\right]^{\frac{(1-\alpha)\gamma}{1-(1-\alpha)\gamma}} k_t^{\frac{\alpha\gamma}{1-(1-\alpha)\gamma}}.$$

This has the property that as $k_t \to \infty$ the per period revenue becomes infinite. In particular in this case, as the reproduction technology improves and $\beta \to \infty$ the present value competitive rent accruing to the innovator becomes infinite - the opposite of the conventional case.

2.2.3. *Travelpro and Spillover Externalities.* In our basic example, the initial copy owned by the innovator (the "seed") was both necessary and sufficient to create all subsequent future copies of the new good. In those circumstances, we have seen, a competitive innovator earns rents equal to the discounted present value of the whole revenue stream from renting copies of the good. In other words: when what is valuable in the innovation is completely embodied in the initial copy owned by the innovator then the latter appropriates a large share of the social surplus, and there are no externalities.

In the literature on innovation, the idea that there are spillover externalities is widespread. That cheap imitation is possible is undeniable - the best example we know of is the invention, by Travelpro, of the modern wheeled roll-on suitcase with a retractable handle.[6] Obviously such an idea cannot be both useful and secret - and once you see a wheeled roll-on suitcase rolling across the airport terminal it is not difficult to figure out how to make one of your own.

A simple way of capturing this idea is the following. In the elementary model we have assumed that all β copies belong to the owner of the original from which the duplicates are made. With an unpriced spillover externality, some lucky individuals will instead get copies for free. In other words, we may take $\beta = \beta_0 + \beta_S$ where β_0 are the number

[6]We are grateful to Ivan P'ng for bringing this to our attention.

of copies that wind up in the hands of the original owner, and β_S are the number of copies that wind up in the hands of fortunate passers-by. In other words, the competitive rent accruing to the innovator is only

$$q_0 = \sum_{t=0}^{\infty} (\beta_0 \delta)^t u'(\beta^t).$$

Price is driven by the total number of copies β^t, but the innovator collects rents only on those copies β_0^t that have not escaped his control. Unless $\beta_0 = 0$, so that all copies have to be given away to passers-by, this competitive rent is still positive. This result is not terribly surprising: theft after all is a real phenomenon, and yet markets continue to function even though a fraction of goods and services are not sold but are stolen instead.

2.2.4. *Pills and Reverse Engineering.* A more intensely debated case is that of new drugs: after purchasing a few pills, reverse engineering allows you to start producing the same drug at a fraction of the cost shouldered by original creator. According to the standard model, no innovation can possibly take place in such circumstances. According to our model, as we will show momentarily, it depends on how quickly productive capacity can be built up by imitators relative to the rate at which the marginal utility of consumption decreases. Once again, this is an empirical matter that should be carefully investigated, and so far it has not been.

Take, then, a special case of our general model, in which productive capacity requires only labor to be produced, as in $L_t = g(x_t)$, with $g(x_t)$ a monotone increasing and convex cost function. Productive capacity produces the consumption good (the pills) as in $c_t = f(k_t)$, with $f(k)$ a standard neoclassical production function. The representative innovator comes into the market at time $t = 0$ with k_0 - the initial pharmaceutical plant - producing $c_0 = f(k_0)$. As soon as this occurs, the formula for the new drug becomes public knowledge and anyone can build additional productive capacity x_0 using just $L_0 = g(x_0)$ units of labor. This is true for all $t = 0, 1, 2, ...$, hence the law of motion of aggregate capacity is $k_{t+1} = \zeta k_t + x_t$, with $\zeta > 0$. As before, the representative consumer's utility function is $\sum_{t=0}^{\infty} \delta^t [u(c_t) - wL_t]$, and everyone behaves competitively. Given k_0, the intertemporal competitive equilibrium is summarized by sequences of quantities $\{k_t, L_t\}_{t=0}^{\infty}$, solving

$$v(k_t) = \max_{x_t \geq 0} \left\{ u(c_t) - wg(x_t) + \delta v(\zeta k_t + x_t) \right\},$$

and prices $\{p_t, q_t\}_{t=0}^{\infty}$, decentralizing the optimum.[7] The first order condition for an optimum is

$$wg'(x_t) = \delta v'(\zeta k_t + x_t), t = 0, 1, 2, ...,$$

which has a unique monotone decreasing solution $x_t = x^*(\zeta k_t, w)$. The latter is enough to derive all equilibrium quantity and price sequences, for given k_0. In particular, the price of the drug evolves as $p_t = u'(c_t)$, while the market value of a unit of productive capacity is

$$q_t = v'(k_t) = u'(c_t)f'(k_t) + \delta\zeta v'(k_{t+1}).$$

Replacing recursively the first order condition in the latter, we compute the unit price of initial productive capacity,

$$q_0 = \sum_{t=0}^{\infty} (\delta\zeta)^t \, u'(c_t)f'(k_t),$$

which is most certainly positive, and finite if $\delta\zeta < 1$. As in all previous cases, and in spite of the extreme externality we have assumed, competitive innovation is viable if $q_0 k_0 \geq C$, where C is the total fixed cost of inventing the drug and building the initial productive capacity k_0. Alternatively, one can write the price of a unit of k_0 as

$$q_0 = u'(c_0)f'(k_0) + \zeta wg'(x_0).$$

The latter distinguishes the *first mover advantage* - $u'(c_0)f'(k_0)$ - component of competitive rents, from the *cost of imitation* component - $\zeta wg'(x_0)$. It stresses, in particular, that when reverse engineering is costly (as it often is, contrary to widespread fantasies: just think of Aspirin) competitive rents can be quite sizeable. The empirical issue of how strong and quantitatively relevant unpriced spillover externalities are in markets for innovations has scarcely been addressed; our results show it should and suggest a model that can be used to do this.

3. Special Issues

3.1. **Improved Reproduction Technology.** Having established sufficient conditions for competitive innovation to be sustained, let us go back to our elementary model and examine the special case in which

[7] The careful reader will have noticed that, even if there is an externality here, the externality is induced, once and forever, by the arrival of k_0 in the very first period. Once this takes place, everyone knows how to make the drug and the competitive equilibrium from $t = 0$ onward solves the standard optimization problem given in the text.

β, the rate at which copies can be made, increases without bound.[8] This seems to approach the conventional case in which marginal cost falls quickly to zero, and competition is alleged to fail to produce innovations. Even in this special case, conventional wisdom is erroneous for two reasons. First, it ignores the first mover advantage: no matter how good the reproduction technology, getting there first does matter and $q_0 \geq u'(1)$ regardless of β. If, however, the same technological improvement that leads to increased β also reduces the time it takes to make additional copies - that is, periods get shorter increasing the discount factor δ - then $u'(1)$ does effectively fall to zero.

Secondly, and more importantly, conventional wisdom ignores the role that demand elasticity plays even in the extreme case of rapidly expanding production capacity. If the price at which the creator can sell under competition is p, rapid reproduction results in k available copies, and the cost of creating the innovation is C, then the return on the innovation is $R = pk/C$. If k is very large, then p must be small, and so the conventional wisdom holds, certainly $R < 1$, meaning it is unprofitable to innovate under competition. This argument is problematic: it may well be true that p is close to zero, but it is only close to zero because k is large, and there is no theorem that a small number times a large one is small. Indeed pk is simply revenue, and the behavior of revenue for large k depends on the elasticity of demand. When there are many small and large numbers around, we should not be too quick to round off selected ones to zero, which is why, in what follows, we carefully consider also the special case in which the marginal cost of reproduction goes to zero. We compute

$$\frac{dq_0}{d\beta} = (1/\beta) \sum_{t=0}^{\infty} t(\beta\delta)^t u'(\beta^t) \left[1 + \frac{\beta^t u''(\beta^t)}{u'(\beta^t)}\right].$$

This of course can be either positive or negative, depending on whether demand elasticity, $-u''(k)k/u'(k)$, is greater or smaller than one and, more importantly, depending on whether this happens at high or at low levels of k. Let us consider the various cases in turn.

As long as demand is elastic, then naturally as the reproduction rate increases, competitive rents also grow. Notice that this is true regardless of the lead time for making copies - that is, with elastic demand revenue goes to infinity as $\beta \to \infty$ even if $\delta \to 1$. Obviously, this leads to the paradoxical conclusion of infinite rents as it is based on

[8]Purely for reasons of simplicity, we provide here the details only for the elementary model. The main result applies, almost verbatim, to the general case as well.

the implausible assumptions that marginal costs are zero and demand is elastic also at very large levels of consumption. However, the theoretical point remains: with elastic demand competitive rents increase with the reproduction rate, against conventional wisdom.

If, on the other hand, demand is always inelastic, prices may fall to zero sufficiently quickly that revenue falls to zero as well. This latter case supports the conventional theory, which therefore arises as a special case of our model under two particular assumptions: that the reproduction rate approaches infinity and demand is inelastic throughout its range.

Consider, then, the truly relevant case in which $-u''(k)k/u'(k)$ is sometime larger and sometime smaller than one. Inspection of the formula shows that $dq_0/d\beta$ is a weighted average of the sequence $1 + u''(\beta^t)\beta^t/u'(\beta^t)$, $t = 0, 1, 2, ...$, with weights that must be decreasing fast enough for $dq_0/d\beta$ to be finite. With the help of a little algebra, one concludes that competitive rents increase with β as long as demand is elastic initially and inelastic "sufficiently late," that is, at high levels of k. A condition that, altogether, does not sound particularly unrealistic, at least for newly introduced goods.

Finally, notice that if marginal cost is zero, there is no capacity constraint and demand is elastic throughout its range in a static model, then the monopolist would choose to produce arbitrarily large amounts, resulting in an arbitrarily small price. Since we do not observe monopolists doing this, we might be tempted to infer, as some have done, that demand cannot be elastic throughout its range. However, since marginal cost is not zero, there is a capacity constraint, and the world is not static, this inference is not justified.

3.2. **Complementary Sales.** Our model is one of perfect competition *post* innovation. Aside for the case of simultaneous innovation, prior to innovating no lonely innovator can be a price taker, as her innovation will certainly have a non-trivial impact on relevant prices. For example, an innovation that lowers the cost of making cars will certainly have an impact on price of cars; writing a new novel will certainly have an effect on the price at which that novel can be sold. Creation of a new idea or fixed factor will generally affect prices in other markets as well.

The standard case of perfect competition assumes perfect divisibility - that the initial unit may be produced in arbitrarily small quantities. In the theory of innovation, we have dropped that assumption - recognizing that two first halfs of a book are a poor substitute for the whole. With perfect divisibility and perfect competition, we can ignore the effect that production has on the prices of substitutes and complements,

and who owns these other factors does not matter. In the case of innovation, where a new good is produced in a discrete amount, we can no longer safely do so. For example, writing a new novel will have a significant effect on the demand for the author's services on the lecture circuit; writing a new song will have a significant effect on the demand for the singer's live performances; creating a new software package will have a significant effect on the demand for the author as a consultant, and so forth and so on.

For simplicity, we assume there is a single other commodity, the fixed quantity of which we denote by a. We now write the utility of the representative consumer by $u(c_t, a)$. Conditional on the innovation taking place, our analysis of the time path of c_t is not affected by the presence of a. What we are interested in is the net change in the price of a due to innovation, assuming perfect competition post innovation, that is

$$Q = \sum_{t=0}^{\infty} \delta^t \left[u_a(c_t, a) - u_a(0, a) \right] =$$

$$= \sum_{t=0}^{\infty} \delta^t \int_0^{c_t} u_{ac}(c, a) dc.$$

From this we see that if a is a substitute for c_t, so that the cross partial is negative, then innovation causes the price of a to fall, while conversely if a is a complement of c_t, so that the cross partial is positive, then the innovation causes the price of a to rise.

Consider first the case of exogenous ownership. If the innovator does not own (any) a then Q does not matter to the process of innovation. If the innovator owns (some portion of) a then the incentive to innovate is decreased/increased as the complementary good is a substitute/complement. For example, the owner of a car factory will have a reduced incentive to create a new type of car, because this will lower the value of his existing cars, while the writer of a song will have an increased incentive to create a new song, because this will raise the value of his live performances. The case of recorded versus theatrical performances is an interesting case. Overall, the behavior of the existing industry - carefully avoiding DVD releases until after theatrical performances are concluded - seems to suggest that theatrical performances are a substitute for the DVD. Examination of practices in the music industry suggest the opposite is true for music. Notice however, that even under perfect competition, there is no obligation to release a

recordable version of a product, so even without government intervention the movie industry would still be free to release DVDs only after the first theatrical run was complete.[9]

The case of complements is particularly important, since in practice it provides a significant source of competitive rents. We refer to the sale of a complementary product by the innovator as a *complementary sale.* Notice that in the case of complementary sales, raising the efficiency of reproduction (increasing β) always increases Q. In particular, with complementary sales, it is perfectly possible to have innovation under perfect competition even when the reproduced good is given away for free. For most part of the history of the radio and television industries this was exactly the case: the product was given away for free, and profit - substantial profit - came from the complementary sale of advertising.

Ownership, of course, is not exogenous. The owner of a car factory can sell his factory; even such things as the rights to revenue from the live performance of music can be traded in asset markets. This leads to a key point, first pointed out by Hirshleifer [1971]: an innovator has a substantial first mover advantage from his ability to trade in markets on the basis of his inside information about his innovation. The owner of a car factory who invents a better car can sell the factory prior to announcing his invention, for example. Moreover, in asset markets, short-sales may be possible as well - and in the extreme case considered by Hirshleifer in which the innovator is a price-taker in asset markets, she can generate essentially infinite profit through inside knowledge that prices are about to change by a tiny amount. In the Hirshleifer account, the private value of innovation under perfect competition is generally much greater than the social value. In other words if innovators and creators are small enough players in financial markets and liquidity constraints are not too important, then under perfect competition there will be too much rather than too little innovation.

4. Case Studies

In practice, and most often because of legal regulations, there are few markets that operate without some degree of monopoly power. Some of those that more naturally come to mind we have mentioned while introducing one or another version of our general models, which were in fact conceived to match the historical cases of competitive innovation

[9]The theft of master copies to make reproductions (or for any other reason) is covered under ordinary laws concerning theft, independent of any laws on intellectual property.

mentioned there. In this section we examine a few additional cases in which monopoly power is weak or, at least, its legal enforcement via patents and copyright is absent. Obviously this is not an exhaustive list, nor do these markets correspond in all respects to the idealized perfect competition of our model. They do signal that many examples exist of industries in which innovation thrives where there is fierce competition between innovators and imitators, and that competitive rents play an important role in sustaining the incentive to innovate.

4.1. **Apples, Strawberry Fields and MP3 Files.** The first and obvious real world analog to our elementary model lies in the history of agricultural innovations. Until about the late 1960s in the USA, and until much more recently and even nowadays in the rest of the world, new animal species and plant varieties were not patentable and a decentralized and highly competitive nursery and breeding industry existed that operated, country by country, literally like our basic model assumes. The evidence is so abundant and pervasive worldwide that we will not dwell into further details here, but simply provide the reader with a few references to a small portion of the specialized literature - e.g., Barger and Landsberg [1975], Barragan [2006], Campbell and Overton [1991], Harvey and Howlett [1941] McClellan [1997].

Our basic model also captures the reproduction and distribution of music in the form of MP3 electronic files via a peer-to-peer network or, at the other extreme, of written classical music until about the second half of the nineteenth century in most of continental Europe. Because of binding legal restrictions, we have no evidence of the first (even if the P2P experience should suggest that something very similar will probably emerge, should copyright restrictions be lifted) while we have overwhelming evidence of the second. The most brilliantly and carefully documented study by Scherer [2003] is an excellent starting point for the curious reader. In both cases, one initial copy is produced by the performer or the composer or a combination of the two. In the one case (classical music) the original copy of the new good consists of a hand-written manuscript, while in the other it consists of writing and recording the music. This is used to "seed" the peer-to-peer network in an hypothetical competitive market for MP3 music, while it was sold to publishers and impresarios in the actual competitive market of eighteenth and nineteenth century classical music. Subsequently, users download a copy of the MP3 file, and - while listening to it - make their copy available to other users of the network for further downloading. Alternatively, impresarios and musicians both perform the piece in theatres around the world, for a fee, and make additional

copies of the manuscript, which are also sold on the market at a positive price. As time goes on, the number of individuals who have copies of the file, or the manuscript, grows exponentially and the prices drop until it becomes effectively equal to the marginal cost of reproduction.

We do not, of course, observe a thriving market for resale on the P2P networks - the illegality of trading copyrighted music makes it difficult to buy and sell on that market. By way of contrast in the early days of classical music there was no copyright, and copying and resale abounded. As documented in Scherer [2003] the introduction of monopoly through copyright does not seem to have increased the already thriving production of classical music that took place in its absence. We should also mention that Scherer debunks the common myth that composers lived at the whim of a few rich patrons - indeed the greatest production of classical music took place in larger "free" cities of the North, with large markets for music, and no rich and royal patrons.

4.2. **Books.** The creation of new agricultural seeds and plants is, probably, the closest one gets to the case of simultaneous innovation, as more than one breeder or nursery would come up, within a relatively short span of time, with almost equivalent plant varieties. Most innovations, on the other hand, tend to imply some degree of monopoly power due to the first mover advantage. It is important to emphasize the difference between monopoly profits, which are due to artificial scarcity, and competitive rents, which are due to natural scarcity. The former has distortionary welfare effects, the latter does not. Nevertheless, when potential competition is fierce the effects of the initial monopoly are minimal, and the market approaches perfect competition very quickly.

Consider the production of books and literature. Today of course copyright is nearly ubiquitous, with works in the United States copyrighted as a matter of course, unless the author explicitly rejects it. This was not always the case. Without going back to the far past consider the United States during the 19th Century. Until almost the end of the century, foreign works - including books written in England - were not entitled to copyright at all,

> *yet American publishers found it profitable to make arrangements with English authors. Evidence before the 1876-8 Commission shows that English authors sometimes received more from the sale of their books by American publishers, where they had no copyright, than from their royalties in [England]* Arnold Plant [1934]

We should note that in 1850 U.S. population was 23.2 million; in 1851 U.K. population was 27.5 million. Per capita GDP in those same years, in 1996 U.S. dollars, was roughly $1930 in the U.S. and $2838 in the U.K. The literacy rates in both countries were roughly 85%. In other words, the U.S. market was smaller than the U.K. market, but of similar size.

American publishers found it profitable to pay the English authors so that they could get the initial copy ahead of their rivals, earning a substantial profit before copiers would have time to enter the market. That the monopoly distortion from this first mover advantage was small is indicated by the enormous price differential between the sale price of books without copyright in the U.S. and with copyright in the U.K. For example, Dickens' *A Christmas Carol* sold for six cents in the US, while it was priced at roughly two dollars and fifty cents in England.

However relevant this may be for our understanding of how innovations have been historically created, one may wonder whether data from an age of clipper ships and hand presses is relevant to an age of cheap electronic reproduction. The ubiquity of modern copyright makes this question difficult - but not impossible - to answer. Documents produced by the U.S. government are not subject to copyright, and a few have been commercial best sellers. The most significant government best seller of recent years has the rather off-putting title of *The Final Report of the National Commission on Terrorist Attacks Upon the United States*, but it is better known as the *9/11 Commission Report*. The report was released to the public at noon on Thursday July 22, 2004. At that time, it was freely available for downloading from a government website. A printed version of the report published by W.W. Norton simultaneously went on sale in bookstores. Norton had signed an interesting agreement with the government.

> *The 81-year-old publisher struck an unusual publishing deal with the 9/11 commission back in May: Norton agreed to issue the paperback version of the report on the day of its public release. (An indexed hardcover edition will follow.) Norton did not pay for the publishing rights, but had to foot the bill for a rush printing and shipping job; the commission did not hand over the manuscript until the last possible moment, in order to prevent leaks. The company will not reveal how much this cost, or when precisely it obtained the report. But expedited printings always cost extra, making it that much more difficult for Norton to realize a profit. In addition, the commission*

> *and Norton agreed in May on the 568-page tome's rather low cover price of $10, making it that much harder for the publisher to recoup its costs. (Amazon.com is currently selling copies for $8 plus shipping, while visitors to the Government Printing Office bookstore in Washington, D.C. can purchase its version of the report for $8.50.) There is also competition from the commission's Web site, which is offering a downloadable copy of the report for free. And Norton also agreed to provide one free copy to the family of every 9/11 victim.* Brendan Koerner [2004]

To be clear: what Norton received from the government was the right to publish first, and the right to use the word "authorized" in the title. What they did not get was the usual copyright - the right to exclusively publish the book. Because it is a government document, the moment it was released, other individuals, and more important, publishing houses, had the right to buy or download copies and to make and resell additional copies - electronically or in print - at a price of their choosing - in direct competition with Norton. In other words - after the release of the book on July 22, the market became a conventional competitive market.

The right to compete with Norton was not a purely hypothetical one. Another publisher, St. Martin's, in collaboration with the New York Times, released their own version of the report in early August - about two weeks after Norton[10] - and this version contained not only the entire government report - but additional articles and analysis by New York Times reporters. Like the Norton version, this version was also a best seller. In addition, it is estimated that 6.9 million copies of the report were (legally) downloaded over the Internet.[11] Competition, in short, was pretty fierce.

Despite this fierce competition, the evidence suggests that Norton was able to turn a profit. We do not know, unfortunately, how much they would have paid up front to the "author" had the rights to go first been put out to bid. But we do have some idea of how much they made after the fact. First, we know that they sold about 1.1 million copies, and that they charged between a dollar and a dollar fifty more than St. Martin's did. Other publishers also estimated Norton made on the order of a dollar of profit on each copy. Assuming that St.

[10]This date and description of the content is from Wyatt [2004].

[11]May [2005] reports sales and download estimates. The Norton version sold 1.1 million copies and the other publisher St. Martin's sold about 900,000 copies.

Martin's has some idea of how to price a book to avoid losing money, this suggests Norton made at the very least on the order of a million dollars. We also know that their contract with the government called upon them to donate their "profits" to charity - and they did in fact "donate $600,000 to support the study of emergency preparedness and terrorism prevention."[12] Since the entire Hollywood movie industry has managed by creative accounting to avoid earning a profit during its entire history, we can be forgiven if we suspect that Norton earned a bit more than the $600,000 they admitted to.

To put these numbers in perspective, the 9/11 commission report was in paperback and, including free downloads, there seem to be about about 8 million copies in circulation. In contrast the initial print run for *Harry Potter and the Half-Blood Prince* was reported to be 10.8 million hardcover copies.[13] So we can realistically conclude that if J. K. Rowling were forced to publish her book without the benefit of the monopoly conferred by copyright, she might reasonably expect to sell the first copy of the book to a publishing house for several million dollars - or more. This is certainly quite a bit less money than she earns with her current legal monopoly. But it seems likely, given her previous occupation as a French teacher,[14] that it would still give her adequate incentive to produce her great works of literature.

Returning to the 9/11 commission report, the bulk of Norton's profits were from the short-term monopoly of copies during the two weeks prior to entry.[15] Since at most 300,000 people preferred to wait two weeks to purchase a copy for a dollar less than Norton, the deadweight loss from the monopoly is at most $300,000, and a more plausible estimate is to assume a uniform distribution of values, which would make the loss $150,000. If we assume that the 800,000 people who purchased the book for $8.00 during the first two weeks were indifferent to purchasing it two weeks later for $7.00 and discount the value two

[12]This was reported in the Associated Press [2005].

[13]This figure was widely reported. See for example www.veritaserum.com.

[14]Reported in an on-line biography at gaga.essortment.com.

[15]Norton initially printed 600,000 copies, followed according to Koerner [2004] quickly by a print run of an additional 200,000 copies. It appears that these copies were sold prior to the entry by St. Martin's. According to Wyatt [2004] the St. Martin's version was available roughly two weeks after the Norton version. Subsequently, according to May [2005] an additional 300,000 copies were sold by Norton. Since according to best-seller reports in the Washington Post the Norton version outsold the St. Martin's version during the weeks immediately following the release of the St. Martin's version, we assume that the number of people who prefered to wait two weeks was no more than 300,000.

weeks later by a factor[16] of 50%, then they value purchase of the book at $9.00. So the 300,000 people that waited had values between $7.00 and $9.00. A reasonable approximation is to assume that the 800,000 people who did not wait had values ranging uniformly over $9.00 to $15.00. This implies an average per consumer surplus of $4.00 for each 800,000 consumers who bought the $8.00 copy during the first two weeks. In addition we estimate an average per consumer surplus of $1.00 for the remaining 1.2 million copies sold after the first two weeks. So the total consumer surplus is about $4.3 million, about thirty times the deadweight loss from monopoly. By way of contrast with unrestrained monopoly, linear demand, and constant marginal cost, the consumer surplus is equal to the deadweight loss from monopoly - and this is a welfare triangle that does not seem to trouble economists a great deal.

In the end it should be no great surprise that ideas of great social value are going to be produced under competition. The great blockbuster novels; the life-saving drugs - all generate such great surplus relative to the cost of creation that relatively little of that surplus need be captured by the innovator to make it worth her while. And indeed, the great works of Shakespeare and Mozart were created under conditions of perfect competition. What about more socially marginal creations? Naturally, the creator will have to capture a greater share of social surplus if these are to be created. On the other hand, competition is likely to be less fierce. While *Harry Potter and the Half Blood Prince* was scanned and illegally released onto the Internet within hours of appearing in print, we have been unable to find any trace of pirate versions of Sara Rath's opus *Star Lake Saloon and Housekeeping Cottages: A Novel* published the same week.

What is the effect of decreased reproduction costs on marginal creations? A widely held belief seems to be that Internet piracy will eliminate these creations unless the government intervenes, for example, by increasing the penalties for "piracy." Interestingly, the Internet seems to have increased rather than decreased the production of marginal ideas. A case in point is the creation of comic strips with very small audiences. With fixed costs of print runs and distribution - to say nothing of the cost of finding a few interested readers in a population of billions - such comics were never produced prior to the advent of the Internet. Now they are. Realistically, such small scale productions are never going to benefit from copyright or government intervention - it

[16]This is the "impulsive" discount factor used in hyperbolic discounting theory, so it is at the high range of how much we might imagine the near future reasonably being discounted. See Laibson [2005].

would never pay to sue someone over copying a comic strip that few people read, and would be equally hard to get the FBI interested in pursuing copyright violators. So the profits of these small productions come from competitive rents, and a large share of these competitive rents are due to complementary sales, as theorized above. In an earlier version of this paper we joked that novels would still be written under perfect competition as long as authors were able to sell signed copies of t-shirts. It turns out that this is not a joke for small audience comic strips.

> *Rosenberg raves that he has been able to make five times as much off his merchandising as off his subscriptions and that advertising doesn't come close to generating the revenue he gets off t-shirts, noting a profit margin of up to 50%, which would be as much as $9 per item in some cases. Stevens quotes $4-$5 as his margin. Rosenberg further claims to have tripled his 2003 income by switching to t-shirt sales in the last three months of 2003.* Todd Allen [2005]

4.3. **Financial Securities.** Prior to 1998, investment bankers and other firms selling financial securities operated without the "benefit" of IP protection. The rapid pace of innovation in financial securities until the late 1980s is well documented, for example by Tufano [1989]. Tufano estimates that roughly 20% of new security issues involve an "innovative structure." He reports developing a list of some 1836 new securities over a 20 year period and remarks that this

> *severely underestimate[s] the amount of financial innovation as it includes only corporate securities. It excludes the tremendous innovation in exchange traded derivatives, over-the-counter derivative stocks (such as the credit derivatives, equity swaps, weather derivatives, and exotic over-the-counter options), new insurance contracts (such as alternative risk transfer contracts or contingent equity contracts), and new investment management products (such as folioFN or exchange traded funds.)*

Herrera and Schroth [2004] examine the forces underlying this innovation. They show how a market for expertise in the presence of a, fully internalized, learning by doing leads to substantial competitive rents, combined with a small amount of monopoly power that unravels rapidly over time.

4.4. **Pharmaceuticals.** In any discussion of innovation under perfect competition, the pharmaceutical industry quickly comes up. The cost of bringing a new drug to market is large - on the order of $200 million 1989 dollars.[17] Moreover, companies are required to disclose the chemical formula for their products as part of the FDA approval process, and to make available to other manufactors the results of their clinical trials. That is, without patents, this industry would operate under "negative" patent in which the government forces disclosure as a condition of doing business. It is widely perceived that with the elimination of patents in this industry, generics would enter the market at roughly the same time as the original and there would be no profit or rent with which to cover the high cost of creating new drugs.

It turns out that this is far from obvious. From Lanjouw [1999] we can examine the behavior of the Indian pharmaceutical industry. Since 1972 product patents on pharmaceuticals have not been recognized in India. Never-the-less, it takes about 5 years for a new drug to enter the Indian market as a generic following its introduction elsewhere. There are two reasons for this. First, the generic manufacturers generally wait a year to see how the new drug does on the market before making the decision to enter.[18] This highlights a problem with the common view that imitators have an advantage because they only need to imitate successful products. While it is true that few would spend a lot of money imitating a product that sold very little at a low price, by the time the imitator has learned that the original is selling a lot of units at a high price, the innovator has, indeed, sold a lot of units at a high price thereby pocketing quite a bit of competitive rents.

Second, the process of actual imitation and clearing regulatory hurdles takes 3-4 years. Lanjouw conjectures that the amount of time to imitate is short and that the primary delay is the regulatory one, but there is not yet much data that would shed light on this. Regardless, from the creators point of view, in the absence of patent protection - which due to the fact that the patent must be filed long before a drug is approved and marketed lasts only about 10 years - it appears that the innovator of a new product will enjoy a 5 year rather than 10 year monopoly. The evidence suggests that when generics enter the price of

[17]Hansen et al [1991].

[18]Entry of generics following the expiration of a patent in the U.S. is much quicker - according to the CBO [1998] only about a month. However, when the patent expires, the generic manufactor has had nearly 10 years to observe the drug in use and to make plans for entry. For this reason, the Indian market where entry is possible at any time gives a better indication of the effect of abolishing patent protection.

the original does not change much, and the original retains about 50% market share.[19] Suppose a 6% real interest rate and normalize the flow of revenues from monopoly to be 1. Then with 10 years of protection present value is roughly $.6 \times 1 + .4 \times .5 = .80$. With 5 years of protection this becomes $.3 \times 1 + .7 \times .5 = .65$. In other words, even if we continue to enforce the revelation of trade secrets for free in this industry and eliminate patents entirely, the present value of rents would still be 81.25% of what it had been before.

4.5. **Open Source Software.** The market for open source software is the most striking example of competitive innovation. Open source software is characterized by the voluntary renunciation of copyright and patent - buyers are entitled to make their own copies, modified or not, and sell them. In addition there is a voluntary renunciation of trade-secrecy as the original creator publishes the source code - the "blueprint" for writing the software - along with the software itself, and buyers are also entitled to make copies modified or not and sell them.[20]

Two questions arise: First, what exactly are the competitive rents in this market? Second, is the market itself significant - or, as it is sometimes alleged, does the market simply free-ride off the proprietary market making cheap imitations of software that never would have been produced in the first place absent monopoly power?

The source of competitive rents appears to be only partially limited capacity, as the complementary sale of expertise plays a larger role. Hence, this last example combines both aspects of the main model and of the special case we have considered, in which complementary sales play a role. That is, the actual duplication of copies is sufficiently quick and conditions of demand are such that only small rents seem to be obtainable through the actual sale of copies. It is true that, historically, physical copies have been sold for greater than marginal cost. Red Hat is a company that, at one time, sold a distribution of Linux - a modified and customized Linux system with many features, which can be easily installed. Although the underlying Linux system is obtained by Red Hat for free, the customization and testing conducted by Red Hat is costly. Using prices quoted on the Internet on July 10, 2002, Red Hat charged $59.95 for a package containing its system. Because it is based on the underlying GNU/Linux system, competitors can legally duplicate and sell the exact same "Red Hat" system. In

[19]See, for example, CBO [1998].

[20]Some open source software also requires that as a condition of use, buyers make their modifications available under the same terms.

fact, at least two two companies, Hcidesign and Linuxemporium, did exactly this. On July 10, 2002, Hcidesign offered for sale Red Hat Linux 7.2 for a price of $16.00, about 1/3rd of the price charged by Red Hat. Linuxemporium.co.uk offered a similar deal. The striking fact being that Red Hat sold many more $59.95 packages than Hcidesign and Linuxemporium did with $16.00 packages and never represented a dangerous market threat to Red Hat.

However, despite their ability to earn revenue from the sale of copies, Red Hat eventually concluded that they were not selling enough $59.95 copies and switched to a different revenue model. What had previously been sold is now given away for free as "Fedora Core" and is used as a platform to get feedback on features that are incorporated into the commercial system called "Red Hat Enterprise Linux" which is available only by annual subscription at a price that - depending on features - ranged on August 24, 2005 from $349 to $2499. The following blurb from Red Hat promotional material makes it clear what it is that is being paid for

> *Unlimited access to service and support: Subscriptions include ongoing service and support to guarantee your systems remain secure, reliable, and up-to-date. When you have a technical question, you'll speak to Red Hat Certified Software Engineers. Or you can access a self-serve knowledge base of technical information and updates.*

What does this offer that imitators cannot? The answer is obvious: if you have a problem with software, would you prefer to consult with the people who wrote it or the people who copied it? Notice, and this is relevant to our discussion, that the complementary sale here rests on the existence of a fixed factor that is very costly to imitate and duplicate: technical expertise, as embodied in the human capital of Red Hat's employees.

Similar observations are made by Lerner and Tirole [2004] about the financial benefit to individual developers of contributing to open source projects. For example, Apache is the leading webserver on the internet, holding a greater than 65% market share.[21] The team of programmers that develop Apache are ranked according to the significance of their contributions, and hold other jobs. Work by Hann et al [2004] shows that the salaries they receive in these other jobs is heavily influenced by their rank within the Apache organization. In other words, the

[21]See http://news.netcraft.com/archives/web_server_survey.html for current statistics on webserver market share.

"expertise" model is much like that in academia - the software writers write software in order to receive recognition and financial payment for the expertise they demonstrate through their published product. While there is no doubt that some contributions to open source are due to altruism, it is equally certain that this massive industry is in fact largely financed through competitive rents, pretty much like the academic one.[22]

From the perspective of the theory, the relevant competitive rents are generated more by the limited stock of expertise available rather than the direct sales of boxed or downloaded software. Notice that the stock of expertise unfolds over time much as in the two-sector model - originally the writer/creator of the software is the only one with expertise. But as time goes on, others, such as Red Hat Certified Software Engineers are taught the knowledge, and the stock of expertise expands, and the price at which it can be sold drops. Of course, in the meanwhile new innovations are created, and new expertise is generated.

Finally, we turn to the question of just how massive and successful the open source industry is. It could be that it exists only because it is able to free-ride off of the innovations created in the proprietary part of the industry in which the monopoly power of copyright plays a key role. Certainly it is true that Linux is a knock-off of Unix and that Openoffice Writer is a knock-off of Microsoft Word. But this means little, because practically all software, proprietary or not, is an imitation of some other software. Microsoft Windows is an imitation of the Macintosh, which is an imitation of Smalltalk. Microsoft Word is an imitation of Wordperfect, which is an imitation of Wordstar. Microsoft Excel is an imitation of Lotus 1-2-3 which is an imitation of Visicalc. And so forth and so on.

A good example is the webserver.[23] The first webserver was written by Tim Berners-Lee at CERN in 1991 and was followed shortly by the NCSA webserver written by Robert McCool. Neither of these ever saw much commercial use, both were public domain, and both were effectively publicly funded. This initial pattern is similar to the way that basic research, for example, in pharmaceuticals is generally publicly

[22]While we have made repeated and ironic references to the academic industry, we have not bothered to go through the details of showing why the "production of new PhDs by means of old PhDs" business in which we are all involved is very well captured by our general model. Just let capital be the stock of researchers with an economics PhD. and consumption thier teaching of undergraduates and carrying out "prizeable" research. We assume our average reader is pretty familiar with such a competitive industry.

[23]Information about the history of the webserver is from the Wikipedia.

funded. Following this, Netscape corporation introduced a proprietary webserver and at about the same time Apache took over the code from the NCSA webserver. Both of these servers survive today, with the Netscape server having mutated into the the Sun One webserver, and Apache having become the dominant force in the webserver industry. Many new features have been added to these servers since their inception, as well as to the competing Microsoft product - the evidence suggests that Apache has been at least as innovative as the others in introducing new features.

Another interesting case is that of word processing. Many open source alternatives to Microsoft Word exist, including Kword, AbiWord and OpenOffice Writer, the latter being the most widely used. How did the cost of developing this software - financed as it was by an open source model - compare to the cost of developing Microsoft Word? The fact is that most of the cost of writing software is not in the observation that it might be nice to have a button to justify text, or in the algorithms for spacing lines - which were after all developed by Gutenberg back in 1450 - but rather in the detailed implementation and debugging of the computer code. As far as we know, none of these open source projects benefited at all from the work done by Microsoft in developing their detailed computer code. Indeed, it appears as if the development of these open source projects was probably more expensive than the development of Microsoft Word - the single most difficult and expensive programming task faced by the developers of these projects appears to be the need to reverse engineer Microsoft Word documents and provide compatible formatting capability so that documents in Microsoft Word are usable and documents can be exchanged with Microsoft Word. Had these projects gone first, this substantial cost would have been avoided. It is also worth noting that the competitive rents generated by these projects is significantly smaller than they would have been had they hit the market before Microsoft Word did. So it seems reasonable to conclude that perfect competition would have delivered both these programs, as it did, and Microsoft Word as well.

Probably the most innovative program in the last few years is BitTorrent, a program that decentralizes and vastly increases the speed at which very large files can be downloaded off the internet. It is commercially successful in the sense that 50,000 copies a day are downloaded.[24] It is also sufficiently innovative that it is now being imitated - by Microsoft.[25] BitTorrent, however, is open source and, according to their

[24]From the statistics for the project provided by the SourceForge website.

[25]The Microsoft knockoff is called "Avalanche." Nicolai [2005] has details.

website Bram Cohen, the author, maintains the program for a living. His expertise, apparently, is not immediately duplicable.

5. Conclusion

There is one question that we have carefully avoided asking - given that innovation can and does thrive in the absence of monopoly power, and certainly in the absence of the artificial monopoly power imposed by government copyright and patent regulation - what should government policy be? From a social point of view are copyright and patents a good idea? Or perhaps government policy should go the opposite direction - should trade secrecy be abolished? We have no intention of trying to answer those questions here. The trade-off is complicated - certainly competitive rents can and do sustain innovation, and substantial amounts of innovation. But competitive rents are not equal to social value, so in the presence of fixed costs, there will be socially valuable innovations that will not occur. So we cannot say that government sustained monopoly over ideas is necessarily a bad idea. Regardless of what ones priors are on this however, we do not see how it is possible to have a sensible, let alone correct, discussion of policy without understanding first how and why the absence of monopoly by no means implies the absence of innovation.

What we have done, is to present a model which is consistent with plain fact that competitive innovation is pervasive in history and nowadays, in which the size of competitive rents and speed of imitation can be measured and in which, most importantly, the policy relevant questions can be meaningfully asked.

References

[1] Aghion, P. and P. Howitt (1992), "A Model of Growth through Creative Destruction," *Econometrica* **60**, 323-351.

[2] Allen, T. [2005], *Online Comics Vs. Printed Comics: A Study in E-Commerce and the Comparative Economics of Content*, IndignantOnline.

[3] Arrow, K.J. (1962), "Economic Welfare and the Allocation of Resources for Invention,", in Richard Nelson (ed.), *The Rate and Direction of Inventive Activity*, Princeton, NJ: Princeton University Press.

[4] Associated Press (2005), "Publisher of 9/11 Report to Make Donation," *Associated Press*, July 21.

[5] Barger, H. and H.H. Landsberg (1975), *American Agriculture, 1899-1939*, Ayer Publishing reprint of NBER Report (1942).

[6] Barragan-Arce, J. (2006), *Competitive and Directed Innovations in the US Agricultural Sector*, Ph.D. Thesis, Department of Applied Economics, University of Minnesota.

[7] Barro, R.J. and X. Sala-i-Martin (1999), *Economic Growth,* Cambridge, Mass.: MIT Press.
[8] Becker, G., [1971], *Economic Theory* , Knopf Publishing Co.
[9] Boldrin, M. and D. K. Levine (1997), "Growth under Perfect Competition," UCLA and Universidad Carlos III de Madrid, October.
[10] Boldrin, M. and D.K. Levine (2002), "Factor Saving Innovation," *Journal of Economic Theory,* **105**, 18–41.
[11] Boldrin, M. and D.K. Levine (2006), "Market Size and Intellectual Property," mimeo, University of Minnesota and UCLA, February.
[12] Campbell, B. and M. Overton (1991), *Land, Labour and Livestock: Historical Studies in European Agricultural Productivity*, Manchester University Press.
[13] CBO (1998), "How Increased Competition from Generic Drugs has Affected Prices and Returns in the Pharmaceuticals Industry," Congressional Budget Office.
[14] Grossman, G.M. and E. Helpman (1991), "Quality Ladders in the Theory of Growth," *Review of Economic Studies* **58**, 43-61.
[15] Grossman, G. and E. Helpman (1994), "Endogenous Innovation in the Theory of Growth", *The Journal of Economic Perspectives*, **8**, 23-44.
[16] Hann, I., J. Roberts, S. Slaughter, and R. Fielding (2004), "An Empirical Analysis of Economic Returns to Open Source Participation," Carnegie-Mellon University.
[17] Hansen, R. W., H. G. Grabowski and L. Lasagna (1991), "Cost of Innovation in the Pharmaceutical Industry", *Journal of Health Economics* **10**, 107-142.
[18] Harvey, J. G. and F. S. Howlett. (1941), *Modern Fruit Production,* The Macmillan Company, New York.
[19] Hellwig, M. and A. Irmen (2001), "Endogenous Technological Change in a Competitive Economy", *Journal of Economic Theory.* **101**, 1-39.
[20] Herrera, H. and E. Schroth (2004), "Developer's Expertise and the Dynamics of Financial Innovation: Theory and Evidence," FAME-Univ. of Lausanne, Research Paper No. 124, Ocotber; and mimeo, ITAM, Mexico D.F. .
[21] Hui, K.L. and I. P'ng (2003), "Piracy and the Legitimate Demand for Recorded Music," *Contributions to Economics Analysis and Policy* **2**, 11.
[22] Karr, R. (2002), "TechnoPop: The Secret History of Technology and Pop Music," NPR, http://www.npr.org/programs/morning/features/2002/technopop/.
[23] Koerner, B. I. (2004), "Who Gets the 9/11 Report Profits? The publisher's printing at least 800,000 copies. Where will all the money go?" *Slate*, July 27, 2004.
[24] Laibson, D. (2005), "Intertemporal Decision Making," forthcoming in Encyclopedia of Cognitive Science.
[25] Lanjouw, J. O. (1999), "The Introduction of Pharmaceutical Product Patents in India: Heartless Exploitation of the Poor and Suffering?," NBER WP# 6366
[26] Leibowitz, S. (2002), "Record Sales, MP3 Downloads and the Annihilation Hypothesis," mimeo, University of Texas, Dallas.
[27] Lerner, J. and J. Tirole (2004), "The Economics of Technology Sharing: Open Source and Beyond," NBER Working Paper 10956.

[28] Levin, R.C., A.K. Klevorick, R.R. Nelson and S.G. Winter (1987), "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity*, **1987-3**, 783-820.

[29] May, E. R. (2005), "When Government Writes History: The 9-11 Commission Report," *History News Network*, http://hnn.us/articles/11972.html.

[30] McClelland P.D. (1997), *Sowing Modernity: America's First Agricultural Revolution*, Cornell University Press.

[31] Niccolai, J. (2005), "Microsoft readies BitTorrent alternative Avalanche technology could make it easier to distribute big files over the Internet," *IDG News Service*, June 16.

[32] Plant, A. (1934), "The Economic Aspect of Copyright in Books," *Economica*, 167-195.

[33] Quah, D. (2002), "24/7 Competitive Innovation," mimeo, London School of Economics.

[34] Romer, P.M. (1990), "Endogenous Technological Change," *Journal of Political Economy* **98**, S71-S102.

[35] Scherer, F.M. (2003), *Quarter Notes and Bank Notes: The Economics of Music Composition in the Eighteenth and Nineteenth Centuries*, Princeton University Press.

[36] Schumpeter, J. (1942), *Capitalism, Socialism, and Democracy*, Harper Perennial (1962).

[37] Shell, K. (1966), "Toward a Theory of Inventive Activity and Capital Accumulation," *The American Economic Review (Papers and Proceedings)*, **56**, 62-68.

[38] Shell, K. (1967), "A Model of Inventive Activity and Capital Accumulation," in *Essays on the Theory of Optimal Economic Growth* (K. Shell, ed.), Cambridge, Massachusetts: MIT Press, 67-85.

[39] Stigler, G. J. (1956), "Industrial Organization and Economic Practice," in *The State of the Social Sciences* edited by Leonard D. White, University of Chicago Press.

[40] Tofuno, P. (1989), "First Mover Advantages in Financial Innovation," *Journal of Financial Economics*, **3**, 350-370.

[41] Wyatt, E. (2004), "For Publisher of 9/11 Report, a Royalty-Free Windfall," *New York Times*, July 28.

[42] Zeira, J. (2006), "Machines as Engines of Growth," mimeo, Hebrew University, Jerusalem, January.

Department of Economics, Washington University in Saint Louis